Proceedings SOR

Rupnik V. and L. Bogataj (Editors): The 1st Symposium on Operational Research, SOR'93. Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 1993, 310 pp.

Rupnik V. and M. Bogataj (Editors): The 2nd International Symposium on Operational Research in Slovenia, SOR'94. Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 1994, 275 pp.

Rupnik V. and M. Bogataj (Editors): The 3rd International Symposium on Operational Research in Slovenia, SOR'95. Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 1995, 175 pp.

Rupnik V., L. Zadnik Stirn and S. Drobne (Editors.): The 4th International Symposium on Operational Research in Slovenia, SOR'97. Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 1997, 366 pp. ISBN 961-6165-05-4.

Rupnik V., L. Zadnik Stirn and S. Drobne (Editors.): The 5th International Symposium on Operational Research SOR '99, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 1999, 300 pp. ISBN 961-6165-08-9.

Lenart L., L. Zadnik Stirn and S. Drobne (Editors.): The 6th International Symposium on Operational Research SOR '01, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2001, 403 pp. ISBN 961-6165-12-7.

Zadnik Stirn L., M. Bastič and S. Drobne (Editors): The 7th International Symposium on Operational Research SOR'03, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2003, 424 pp. ISBN 961-6165-15-1.

Zadnik Stirn L. and S. Drobne (Editors): The 8th International Symposium on Operational Research SOR'05, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2005, 426 pp. ISBN 961-6165-20-8.

Zadnik Stirn L. and S. Drobne (Editors): The 9th International Symposium on Operational Research SOR'07, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2007, 460 pp. ISBN 978-961-6165-25-9.

Zadnik Stirn L., J. Žerovnik, S. Drobne and A. Lisec (Editors): The 10th International Symposium on Operational Research SOR'09, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2009, 604 pp. ISBN 978-961-6165-30-3.

Zadnik Stirn L., J. Žerovnik, J. Povh, S. Drobne and A. Lisec (Editors): The 11th International Symposium on Operational Research SOR'11, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2011, 358 pp. ISBN 978-961-6165-35-8.

Zadnik Stirn L., J. Žerovnik, J. Povh, S. Drobne and A. Lisec (Editors): The 12th International Symposium on Operational Research SOR'13, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2013, 390 pp. ISBN 978-961-6165-40-2.

Zadnik Stirn L., J. Žerovnik, M. Kljajić Borštnar, S. Drobne (Editors): The 13th International Symposium on Operational Research SOR'15, Proceedings. Ljubljana: Slovenian Society Informatika, Section for Operational Research, 2015, 559 pp. ISBN 978-961-6165-45-7.

Proceedings SOR'17

Proceedings of the 14th International Symposium on OPERATIONAL RESEARCH

SOR '17

Bled, Slovenia
September 27-29, 2017

Edited by:
L. Zadnik Stirn • M. Kljajić Borštnar • J. Žerovnik • S. Drobne

SOR '17 Proceedings

The 14th International Symposium on Operational Research in Slovenia

Bled, SLOVENIA, September 27 - 29, 2017

Edited by:

L. Zadnik Stirn, **M. Kljajić Borštar**, **J. Žerovnik** and S. Drobne

Slovenian Society INFORMATIKA (SDI)
Section for Operational Research (SOR)

Program Committee:

L. Zadnik Stirn, University of Ljubljana, Biotechnical Faculty, Ljubljana, Slovenia, Chair
M. Kljajić Borštnar, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia, Co-Chair
Z. Babić, University of Split, Faculty of Economics, Department for Quantitative Methods, Split, Croatia
M. Bajec, University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia
M. Bastič, University of Maribor, Faculty of Business and Economics, Maribor, Slovenia
M. Bogataj, The Mediterranean Institute for Advanced Studies, Šempeter pri Gorici, Slovenia
M. Bohanec, Jožef Stefan Institute, Department of Knowledge Technologies, Ljubljana, Slovenia
D. Bokal, University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor, Slovenia
A. Brezavšek, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
S. Cabello, University of Ljubljana, Faculty of Mathematics and Physics, Ljubljana, Slovenia
K. Cechlarova, P. J. Šafarik University, Faculty of Science, Košice, Slovakia
S. Cozzini, CNR/IOM and eXact-lab srl, Trieste, Italy
T. Csendes, University of Szeged, Department of Applied Informatics, Szeged, Hungary
V. Čančer, University of Maribor, Faculty of Business and Economics, Maribor, Slovenia
S. Drobne, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana, Slovenia
K. Dumičić, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia
L. Ferbar Tratar, University of Ljubljana, Faculty of Economics, Ljubljana, Slovenia
W. Gutjahr, University of Vienna, Department of Statistics and Decision Support Systems, Vienna, Austria
N. Gvozdenović, University of Novi Sad, Faculty of Economics, Subotica, Serbia
J. Jablonsky, University of Economics, Faculty of Informatics and Statistics, Prague, Czech Republic
S. Klavžar, University of Ljubljana, Faculty of Mathematics and Physics, Ljubljana, Slovenia
D. Kofjač, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
J. Kušar, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia
A. Lisec, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana, Slovenia
Z. Lukač, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia
L. Neralić, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia
M. Pejić Bach, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia
U. Pferschy, University of Graz, Department of Statistics and Operations Research, Graz, Austria
J. Povh, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia
V. Rajkovič, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia
M. S. Rauner, University of Vienna, Dept. of Innovation and Technology Management, Vienna, Austria
B. Rodič, Faculty of Information Studies, Novo mesto, Slovenia
R. Sotirov, Department of Econometrics and Operations Research, Tilburg University, The Netherlands
K. Šorić, Zagreb School of Economics and Management, Zagreb, Croatia
D. Škulj, University of Ljubljana, Faculty of Social Sciences, Ljubljana, Slovenia
O. Tang, Linköping University, Department of Management and Engineering, Linköping, Sweden
T. Trzaskalik, University of Economics, Department of Operational Research Katowice, Poland
G. W. Weber, Middle East Technical University, Institute of Applied Mathematics, Ankara, Turkey
M. Zekić Sušac, University of Osijek, Faculty of Economics, Croatia
J. Žerovnik, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia

Organizing Committee:

S. Drobne, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Ljubljana, Slovenia, Chair
P. Gorjanc, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia, Co-Chair
N. Fileš, Slovenian Society Informatika, Ljubljana
M. Kljajić Borštnar, University of Maribor, Faculty of Organizational Science, Kranj, Slovenia
J. Povh, Faculty of Information Studies, Novo mesto, Slovenia
L. Zadnik Stirn, University of Ljubljana, Biotechnical Faculty, Ljubljana, Slovenia

The 14th International Symposium on Operational Research in Slovenia - SOR '17
Bled, SLOVENIA, September 27 - 29, 2017

Chairs:

D. Battini, University of Padua, Department of Management and Engineering, Padua, Italy

M. Bogataj, CERRISK - Zavod INRISK, Ljubljana, Slovenia

M. Bohanec, Jožef Stefan Institute, Department of Knowledge Technologies, Ljubljana, Slovenia

D. Bokal, University of Maribor; Faculty of Natural Sciences and Mathematics, Maribor, Slovenia

V. Rajković, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia

V. Čančer, University of Maribor, Faculty of Economics and Business, Maribor, Slovenia

K. Dumičić, University of Zagreb, Faculty of Economics and Business, Zagreb, Croatia

R. Gumzej, University of Maribor, Faculty of Logistics, Celje, Slovenia

T. Illés, Budapest University of Technology and Economics, Budapest, Hungary

M. Kljajić Borštnar, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia

D. Kofjač, University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia

Z. Lukać, University of Zagreb, Department of Mathematics, Zagreb, Croatia

B. Novkovska, University of Tourism and Management, Faculty of Economics, Skopje, Republic of Macedonia

M. Pejić-Bach, University of Zagreb, Faculty Economics and Business, Zagreb, Croatia

S. Pivac, University of Split, Faculty of Economics, Split, Croatia

J. Povh, University of Ljubljana, Faculty of Mechanical Engineering, Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

J. Šilc, Jožef Stefan Institute, Computer Systems Department, Ljubljana, Slovenia

V. Vukašinović, Jožef Stefan Institute, Computer Systems Department, Ljubljana, Slovenia

L. Zadnik Stirn, University of Ljubljana, Biotechical Faculty, Ljubljana, Slovenia

M. Zekić-Susac, University of Josip Juraj Strossmayer in Osijek, Faculty of Economics, Osijek, Croatia

J. Žerovnik, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia

Preface

This volume, Proceedings of The 14th International Symposium on Operations Research, called SOR'17, contains papers presented at SOR'17 (http://sor17.fov.uni-mb.si/) that was organized by Slovenian Society INFORMATIKA (SDI), Section for Operations Research (SOR) and University of Maribor, Faculty of Organizational Sciences, Kranj, Slovenia, held in Bled, Slovenia, from September 27 to September 29, 2017. The volume contains blindly reviewed papers or abstracts of talks presented at the symposium.

The opening address at SOR'17 was given by Prof. Dr. Lidija Zadnik Stirn, President of the Slovenian Section of Operations Research, Mr. Niko Schlamberger, President of the Slovenian Society Informatika, Prof. Dr. Iztok Podbregar, Dean of the Faculty of Organizational Sciences, University of Maribor, Dr. Sarah Fores, The Association of European Operational Research Societies (EURO), and presidents/representatives of a number of Operations Research Societies from abroad.

SOR'17 is the scientific event in the area of operations research, another one in the traditional series of the biannual international OR conferences, organized in Slovenia by SDI-SOR. It is a continuity of thirteen previous symposia. The main objective of SOR'17 is to advance knowledge, interest and education in OR in Slovenia, in Europe and worldwide in order to build the intellectual and social capital that are essential in maintaining the identity of OR, especially at a time when interdisciplinary collaboration is proclaimed as significantly important in resolving problems facing the current challenging times. Further, by joining IFORS and EURO, the SDI-SOR agreed to work together with diverse disciplines, i.e. to balance the depth of theoretical knowledge in OR and the understanding of theory, methods and problems in other areas within and beyond OR. We believe that SOR'17 creates the advantage of these objectives, contributes to the quality and reputation of OR by presenting and exchanging new developments, opinions, experiences in the OR theory and practice.

SOR'17 was highlighted by a distinguished set of five keynote speakers. The first part of the Proceedings SOR'17 comprises invited abstracts and papers, presented by five outstanding scientists: Dr. Serge Bogaerts, Managing Director of PRACE aisbl, PRACE, the Partnership for Advanced Computing in Europe, Brussels, Belgium, Prof. Dr. Ulrike Leopold-Wildburger, University of Graz, Austria, Prof. Dr. Matjaž Perc, University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor, Slovenia, Prof. Dr. Luk N. Van Wassenhove, INSEAD, Fontainebleau Cedex, France, and Prof. Dr. Marijana Zekić-Sušac, University of Josip Juraj Strossmayer in Osijek, Faculty of Economics, Osijek, Croatia.

Proceedings includes 97 papers or abstracts written by 184 authors. Most of the authors of the contributed papers came from Slovenia (50), then from Croatia (47), Czech Republic (14), Turkey (14), Slovak Republic (10), Italy (6), Hungary (6), Norway (5), Spain (5), Greece (3), Israel (3), Russia (3), Belgium (2), Bosnia and Herzegovina (2), Brazil (2), Poland (2), Macedonia (1), Algeria (1), Austria (1), Denmark (1), France (1), Montenegro (1), Nigeria (1), Serbia (1), and United Kingdom (1). The papers published in the Proceedings are divided into Plenary Lectures (4 abstracts and 1 paper), seven special sessions: Advances in Modelling and Statistical Research of the Western Balkan Countries in the Times of Economic Crisis (8 papers), High-Performance Computing and Big Data and General OR Topics (8), Logistics (5), MCDM – Software and Applications (6), Metaheuristic Optimization (5), MRP and Related Systems Approach to Industrial Engineering and Services (5), Systems Optimization and Control with Applications (5), and

eight sessions: Econometric Models and Statistics (10), Environment and Human Resources (6), Finance and Investments (5), Location and Transport, Graphs and their Applications (6), Machine Learning (4), Mathematical Programming and Optimization (8), Multiple Criteria Decision Making (4), and OR Perspectives: Where we have been, where we can go (3).

The Proceedings of the previous thirteen International Symposia on Operations Research organized by the Slovenian Section of Operations Research, that are listed at https://www.drustvo-informatika.si/sekcije/sor/sor-publikacijepublications/, are indexed in the following secondary and tertiary publications: Current Mathematical Publications, Mathematical Review, Zentralblatt fuer Mathematik/Mathematics Abstracts, MATH on STN International and CompactMath, INSPEC. The Proceedings SOR'17 are expected to be covered by the same bibliographic databases.

The success of the scientific events at SOR'17 and the present proceedings should be seen as a result of joint effort. On behalf of the organizers we would like to express our sincere thanks to all who have supported us in preparing the event. We would not have succeeded in attracting so many distinguished speakers from all over the world without the engagement and the advice of active members of the Slovenian Section of Operations Research. Many thanks to them. Further, we would like to express our deepest gratitude to prominent keynote speakers, to the members of the Program and Organizing Committees, to the referees who raised the quality of the SOR'17 by their useful suggestions, section's chairs, and to all the numerous people - far too many to be listed here individually - who helped in carrying out The 14th International Symposium on Operations Research SOR'17 and in putting together these Proceedings. Last but not least, we appreciate the authors' efforts in preparing and presenting the papers, which made The 14th Symposium on Operations Research SOR'17 successful.

We would like to express a special gratitude to The Partnership for Advanced Computing in Europe (PRACE) for a financial support.

Bled, September 27, 2017

Lidija Zadnik Stirn
Mirjana Kljajić Borštnar
Janez Žerovnik
Samo Drobne
(Editors)

Contents

Special Session 2: High-Performance Computing and Big Data and General OR Topics 71

Special Session 3: Logistics 121

Special Session 4: MCDM – Software and Applications 153

Special Session 5: Metaheuristic Optimization   193

Special Session 6: MRP and Related Systems Approach to Industrial Engineering and Services   227

Session 5: Machine Learning 453

Session 6: Mathematical Programming and Optimization 479

Session 7: Multiple Criteria Decision Making 521

Session 8: OR Perspectives: Where we have been, where we can go *547*

APPENDIX *567*

Author index

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Plenary Lectures

PRACE, THE RESEARCH INFRASTRUCTURE SUPPORTING SCIENCE AND INDUSTRY THROUGH PROVISION OF HIGH PERFORMANCE COMPUTING RESOURCES

Serge Bogaerts
Managing Director of PRACE aisbl

Abstract: PRACE (www.prace-ri.eu), the Partnership for Advanced Computing in Europe, provides access to Europe's world-class High Performance Computing Research Infrastructure (RI), enabling scientists and researchers from academia and industry to carry out complex and excellent experiments and simulations that address society's grand challenges. This keynote will present the organisation of this RI and describe the services that researchers can get access to, with a special focus on the evolution brought by the recently started second phase of the organisation. We review some success stories made possible by the infrastructure to trigger some new ideas by researchers to make innovative use of this tool. Some key aspects for improving the chance to be awarded resources will also be covered.

OPERATIONS RESEARCH AND BEHAVIORAL ECONOMICS

Ulrike Leopold-Wildburger
University of Graz, Universitaetsstrasse 15, Graz, Austria
Email: ulrike.leopold@uni-graz.at
Homepage: www.uni-graz.at.sor

Abstract: While operations research represents the field of a science for delivering better decisions using optimal (or near-optimal) solutions to complex decision-making problems our behaviour in practical applications quite often has to deal with non-fully rational decision makers. The tension between the two scopes shall be worked out and supported by a series of examples. Coming from the field of OR we are aware that that techniques such as mathematical modelling, statistical analysis, and mathematical optimization are engaged in applications of advanced analytical methods with the aim to make better decisions. However, in everyday life OR is not executed in its pure version but often connected with other fields and disciplines, as psychology and behavioural sciences, integrating neuroscience and microeconomics.

Some characteristic examples from the field of game theory will be prepared and checked with the actual behaviour of decision makers in specific economic situations. We will deal with topics as cooperation, fairness and honesty and we will try to compare theoretical concepts with empirical data.

TRANSITIONS TOWARDS COOPERATION IN HUMAN SOCIETIES

Matjaž Perc
University of Maribor
Faculty of Natural Sciences and Mathematics
matjaz.perc@uni-mb.si
http://www.matjazperc.com/

Abstract: Cooperation in nature is a challenge to Darwin's theory of evolution, and it is fundamental for the understanding of the main evolutionary transitions that led from single-cell organisms to complex animal and human societies. If only the fittest survive, why should one perform an altruistic act that is costly to perform but benefits somebody else? Why should we care for and contribute to the public good if freeriders can enjoy the same benefits for free? Recent research indicates that a comprehensive answer to these questions requires that we look beyond the individual and focus on the collective behavior that emerges as a result of the interactions among individuals, groups, and societies. Although undoubtedly driven also by culture and cognition, cooperation in human societies is just as well an emergent, collective phenomenon in a complex system.

Nonequilibrium statistical physics, in particular the collective behaviour of interacting particles near phase transitions, has been recognized as valuable for understanding counterintuitive evolutionary outcomes in structured populations. However, unlike pairwise interactions among particles that typically govern solid-state physics systems, interactions among humans often involve group interactions, and they also involve a larger number of possible states even for the most simplified description of reality. The complexity of the problem is further amplified by the inevitable interactions among groups and societies, which can give rise to interdependencies that often induce cascading failures and accelerate transitions towards defection. When studying cooperation in human societies, it is therefore important to consider not only the fact that the range of interactions among people is limited and thus best described by networks, but also that these networks change over time and are often interdependent.

Ultimately, the goal is to develop a predictive, computational theory that will allow us to better understand the rich variety of phenomena that rely on large-scale cooperative efforts. From the mitigation of social crisis and inequality to the preservation of natural resources for next generations, by having a firm theoretical grip on human cooperation we can hope to engineer better social systems and develop more efficient policies for a sustainable and better future.

HUMANITARIAN OPERATIONS:
PAST ACHIEVEMENTS, FUTURE CHALLENGES
AND RESEARCH OPPORTUNITIES

Luk N. Van Wassenhove
INSEAD
Boulevard de Constance, 77305 Fontainebleau Cedex, France
email: luk.van-wassenhove@insead.edu

Abstract: The Humanitarian caseload is increasing while funding is substantially decreasing. The nature and impact of crises is also changing. This creates serious tensions and a need to reflect on recent advances in research. What has been relevant and impactful? Are we dealing with the major issues and providing answers to practitioners? At the same time, technology now allows us to analyse situations and do things that were impossible only half a dozen years ago. How can these new technologies be better integrated and used? How can developments from other disciplines allow humanitarian practitioners to do more with less in these resource-scarce times? The field of humanitarian operations is in great flux and organizations need help in facing the rapidly changing world of providing aid to the ones in need.

The presentation will not attempt to cover all developments for the simple reason that I am not competent for that. However, I can share with you my own research work and experience in collaborating with a set of humanitarian organizations. I hope to convince you that this is an exciting field of research with lots of relevant multi-disciplinary problems where we can make a substantial impact.

MACHINE LEARNING IN ENERGY CONSUMPTION MANAGEMENT

Marijana Zekić-Sušac, PhD, Full professor
University of Josip Juraj Strossmayer in Osijek, Faculty of Economics
Trg Lj. Gaja 7, 31000 Osijek, Croatia
marijana@efos.hr

Abstract: Energy management is one of the hot topics among researchers. How to consume less energy from non-renewable resources is among the biggest challenges and could be approached by different methods. Previous research in the area of energy management shows that various deterministic and stochastic methods have been used for predicting energy consumption. Statistical methods such as autoregressive moving averages (ARMA), cycle analysis or multiple regression are among the most frequent ones, while recent papers reveal that machine learning methods such as neural networks, support vector machines and others show more accuracy in prediction. The advantages of machine learning over standard statistical methods are in the fact that they do not have strong requirements regards stationarity and interdependence of input data, they are robust, and show more success in short-term prognoses. In this paper several machine learning methods will be tested to analyze the influence of specific characteristics of buildings and implemented measures on energy consumption. The extracted knowledge could serve in the process of decision making on investing in measures for improving energy efficiency, as well as in future research for developing additional models.

Keywords: energy management, machine learning, support vector machines, artificial neural networks

1 INTRODUCTION

Energy efficiency is an important topic in the context of climate change and has gained much attention in EU directives and national strategic and action plans. Researchers strive to build models that could efficiently predict or explain the main factors that influence energy efficiency. Since buildings are the largest individual energy consumers [20] and building sector itself contains 40% of total primary energy consumption, the efficient models that decision makers could use to allocate resources in reconstructions of buildings are desirable. The methodology for calculating energy efficiency category of a building is determined by professionals. However, there is a lack of analyses that investigate relationships among various buildings' attributes describing construction, geospatial data, climate data, heating data, cooling data and usage, as well as their connection to energy consumption. Since it is a complex issue which includes uncertainty and nonlinearity, it requires advanced methodology. In this paper, the aim is to investigate possibilities of several machine learning methods in extracting important predictors of yearly energy consumption of electricity and natural gas. The methods of artificial neural networks, CART decision trees, conditional inference trees, random forest, and support vector machines are tested on a real dataset of Croatian public buildings. The similarities and differences among the three tested methods in modelling energy consumption are discussed.

2 PREVIOUS RESEARCH

Previous research in the area of modelling energy efficiency shows that the authors have used several approaches: (1) individual statistical methods such as linear regression, time series

analysis, probability density functions, or similar methods, (2) comparison of statistical methods with machine learning methods, and (3) simulation modelling [21]. More advantage could be given to modeling procedures that include integration of statistical, machine learning and simulation methods. Statistical methods such as linear regression, time series analysis, and probability density functions were used in the area of predicting energy consumption by [6], [13], and [14]. Mangold et al. [14] investigated possible improvements of the process of issuing Energy Performance Certificates (EPC) in Sweden according to EU directives. They improved the estimation of energy efficiency used in certificates by using a stepwise regression model. A comparison of statistical methods with machine learning has been conducted in [5] and [8]. Farzana et al. [8] predicted the energy demand in the urban residential buildings of Chongqing in south west China. They collected data by a structured questionnaire survey on household energy consumption data, and preformed prediction of energy demand by using an ANN model, two Grey models, a regression model, a polynomial model and a polynomial regression model. Their results show that the most accurate model based on MRPE (%) and statistical tests was the ANN model. Son et al. [18] suggested a model for predicting government-owned building energy consumption based on an RreliefF variable selection algorithm and support vector machines method of machine learning. Chou and Bui [5] used the largest number of methods in estimating energy performance of buildings, such as support vector regression (SVR), artificial neural network (ANN), classification and regression tree, chi-squared automatic interaction detector, general linear regression, and ensemble inference model, and conducted an extensive research on 768 experimental datasets from the literature with 8 input parameters and 2 output parameters (cooling load (CL) and heating load (HL)). Their results show that the most accurate models were an ensemble of SVR and ANN for cooling energy load and the SVR model for heating energy load. In summary, the previous research shows that statistical as well as machine learning and simulation methods have been used to predict energy consumption of building. However, there is a lack of a comprehensive methodological framework that could be implemented in practice for detection of possible savings in energy consumption and resource allocation. By comparing more machine learning methods and suggesting their implementation, this paper is a step towards providing such framework in the domain of public buildings.

3 MACHINE LEARNING METHODS FOR REGRESSION PROBLEMS

Several machine learning methods are used in this research: artificial neural networks, support vector machines, and three methods based on recursive partitioning, such as CART decision tree, Conditional inference tree (CTREE), and random forest (RF). All the observed methods basically try to fit a function by case-based reasoning, which includes usage of training data for estimating parameters, and test data to validate the predictions. Due to the fact that we deal with estimating a continuous variable as the output, the above methods were used with algorithms adjusted to work with regression type of problem. All the computations were performed in R software package.

3.1 Artificial neural networks

The advantages of ANNs over standard statistical methods in research as well as in real-world applications were emphasized by a number of researchers [16]. The most common type of ANN is the multilayer perceptron (MLP) suggested by Werbos in 1974 and improved by Rumelhart et al. [15]. They differ from standard linear regression in a way that they do not assume a linear connection between inputs and the output, and by using a multi-layered structure, they simulate multiprocessing of biological networks. A typical MLP consists of

three layers (input, hidden, and output), and the basis computation can be described as follows. The input values $x_i \in R$, $i=1,2,...,$ n, as well as the bias are loaded in the input layer and multiplied by weights w_i (initially set randomly). Each unit in the hidden layer receives the weighted sum of all x_i values from the input layer, and produce the output by using a non-linear activation function f which can be sigmoid, tangent hyperbolic, exponential, sine, linear, step or other (see [15]). The output layer then produces the output by using a linear or nonlinear activation function, compares its output with the real output and computes the error ε. This process describes one iteration of ANN learning. The process is repeated such that in each iteration k the error ε is used to adjust the weights of the input vector according to a learning rule, usually the Delta rule [15]:

$$\Delta w_{ji} = \eta \cdot y_{cj} \cdot \varepsilon_i \tag{1}$$

where η is the learning parameter (set to 0.01 in our research), and y_{cj} is the computed output. The aim is to find the weight vectors that minimize the error. The standard minimization algorithm used in ANNs is the backpropagation [15], while in this research we test resilient backpropagation which. The number of hidden units m in ANN is experimentally set by a cross-validation procedure (testing all ANN structures with $m=1,…,20$, and the m which produces the minimum error is saved in ANN model. The training time is determined in an early-stopping procedure which saves the network with the lowest error. In order to obtain equal results in each attempt, the seed of randomization is set to 500 in all ANN models.

3.2 Support vector machine regression

Support vector machine (SVM) suggested by Vapnik and coauthors [4] is aimed for non-linear mapping of the input vectors into a high-dimensional feature space using optimal separation of hyperplanes based on the maximum margin. Its advantages are in high generalization performance, and absence of the local optima problem [2]. Although it is first aimed for classification type of problems, the SVM for regression was presented by Smola and Schoelkopf [17] . In a set of training data $x_i \in R^n$ with the desired output $y_i \in \{+1,-1\}$, the input x is first mapped onto an m-dimensional feature space using a nonlinear kernel function, and then a linear model is constructed in the feature space. In case of a nonlinear problem, the non-negative Lagrange multipliers $\boldsymbol{\alpha}$ can be searched by optimizing the function:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha^*) k(x_i, x) + b \tag{2}$$

Details of computation can be found in [17]. The kernel function can be linear, sigmoid, radial basis or polynomial. In our research, the radial basis function kernel is used.

3.3 Regression tree partitioning – CART, CTREE and Random forest

Tree-structured recursive partitioning aims to build a binary tree by splitting the input vectors at each node according to a function of a single input. The standard algorithm is the classification and regression tree (CART) suggested by Breiman et al. [3] which basically splits each input variable at all possible split points, for each split point splits the parent node into two child nodes by separating the objects with values lower and higher than the split point for the considered input variable, selects the variable and split point with the highest reduction of impurity, performs the split of the parent node into the two child nodes according to the selected split point, and repeats the above steps using each node as a new parent node until the

tree has maximum size. After reaching the maximum tree, the algorithm prunes the tree back using the cross-validation procedure to select the right-sized tree. The evaluation function used in this research for splitting is the Gini index defined as [1]:

$$Gini(t) = 1 - \sum_i p_i^2$$

(3)

where t is a current node and p_i is the probability of class i in t. The CART algorithm considers all possible splits in order to find the best one by Gini index. Prune of missclassification error was used as the stopping rule, with minimum $n=5$. In case of regression trees where the output produce a real number instead of a class probability, the response for any observation is computed by following the path from the root node down to the appropriate terminal node of the tree, where the values for the splitting variables are observed, and the predicted response value is calculated by averaging response in that terminal node [10]. The limitation of CART trees is in their biasness regarding the variable selection bias, since it does not treat fair the variables of different types, categories, or missing values [10]; thus Hothorn, Hornik, and Zeileis [12] proposed the conditional inference trees (CTREE) that use multiplicity-adjusted conditional tests such that the test of the null hypothesis of no association between any of the predictor and the real output is peformed, at first globally for each node, and than separately for each individual variable in non-terminal nodes. The variable that has the smallest p-values becomes a split variable. If there is no further statistically significant split, the tree stops growing. The advantages of this method over standard CART are in absence of pruning, and in robustness regarding different types of variables and missing data. A Random forest (RF) is a tree partitioning method that consists of a large number of trees where each tree is based on a random subset of the observations, and each split within each tree is created based on a random subset of candidate variables. The forest computes its response by averaging responses of the individual trees. In our research a RF-CART algorithm was used as a random forest based on CART, suggested by Breiman et al. [3]. The advantages of RF algorithm are in overcoming instability of single-tree techniques, improving its performance, however the limitation is in its complexity [7], [11]. In our research, the complexity parameter $Cp=0.01$ with ANOVA was used to save computing time in pruning with CART, such that the overall R-squared must increase by cp at each step. Any split which does not improve the fit by cp will likely be pruned off by cross-validation. In addition, the maximum depth parameter in CART was set to 30 which determines the maximum depth of any node of the final tree.

3.4 Evaluation of model accuracy

All the models obtained by ANNs, SVM, CART, CTREE, and RF were trained on the same training sample and validated on the sample hold-out test sample. Data were normalized before training, and denormalized before computing the evaluation measure. The objective function in all algorithms was the mean sqare error (MSE), but to be able to interpret accuracy in percentage, the mean average percentage error was computed on the test sample. The measures were computed according to:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_t - y_c)^2 \quad , \quad SMAPE = 100\frac{1}{n}\sum_{i=1}^{n}\frac{|y_t - y_c|}{|y_t| + |y_c|}$$

(4)

where is the real output value, is the predicted value, and n is the number of cases in the test sample. SMAPE is recommended as a measure of model successfulness by researchers due to the fact that MSE do not treat equally residuals that are larger and smaller from the real output [19].

4   DATA AND MODELING PROCEDURE

Data used to test ML methods in predicting energy consumption included a real dataset of 2048 public buildings in Croatia with their 141 attribute including geospatial, construction, heating, cooling, meteorological and energy data, obtained from The Agency for Legal Trade and Real Estate Brokerage (APN) in Croatia in 2017. The output variable was the total electricity consumption of each building in 2016. The initial insight into the data revealed that some preprocessing is needed. The analysis showed that the output variable (denoted as SUM_kWH) consisted of outliers, i.e. 190 cases were above the upper quartile. After removing outliers, the sample consisted of 1858 cases. The input space consisted of 141 input variables (122 continuous, and 19 categorical) that could be divided in four groups: geospatial, construction, heating, cooling, meteorological and energy coefficients data. Due to a large number of attributes, their descriptive statistics is available on request. The output variable had the minimum in 0, median was 13404, and maximum was 105 739 kWh.

4.1   Variable reduction procedures and sampling

The next stage of the pre-processing included correlation analysis to investigate interdependence among data. The Shapiro-Wilks test showed that the output variable was not normally distributed (W = 0.83974, p-value < 2.2e-16). Therefore, the Spearman correlation was used and it revealed that there is a high correlation (>0.7) between some continuous input variables. The procedure performed on each highly-correlated pair of input variables excluded the one which had higher average correlation with other variables. Such procedure excluded 45 out of 122 continuous variables. Then the chi-square test of independence is applied pairwise to categorical variables, showing that 13 out of 19 categorical variables are mutually dependent and are removed from the model. The final input space after variable reduction consisted of 77 continuous and 3 categorical variables in the sample. In order to obtain systematic training, testing and validation of NNs, the sample is divided into three subsamples, using equal distribution of output variables in the train (60% of cases) and test sample (20% of cases), while the rest of the cases is added to the validation sample (20% of data). The number of hidden units in ANN is optimized by a cross-validation procedure where the architectures from 1 to 20 hidden units were tested and the number of units which produces the minimum MAPE error is retained.

5   RESULTS – LEVERAGING THE STRENGTHS OF MACHINE LEARNING METHODS IN ENERGY CONSUMPTION

After data preprocessing, modelling procedure started with the stepwise multiple regression model to check for the data linearity. The Wilks-Shapiro test has shown that the assumption of normally distributed residuals of the output variable was not satisfied (W = 0.90966, p-value < 2.2e-16),   thus the results of multiple regression cannot be interpreted and the non-linear modeling was conducted by machine learning methods: ANN, SVM, CART, CTREE, and the RF model.

The final ANN model with 3 hidden units and logistic activation function is presented in Figure 1.
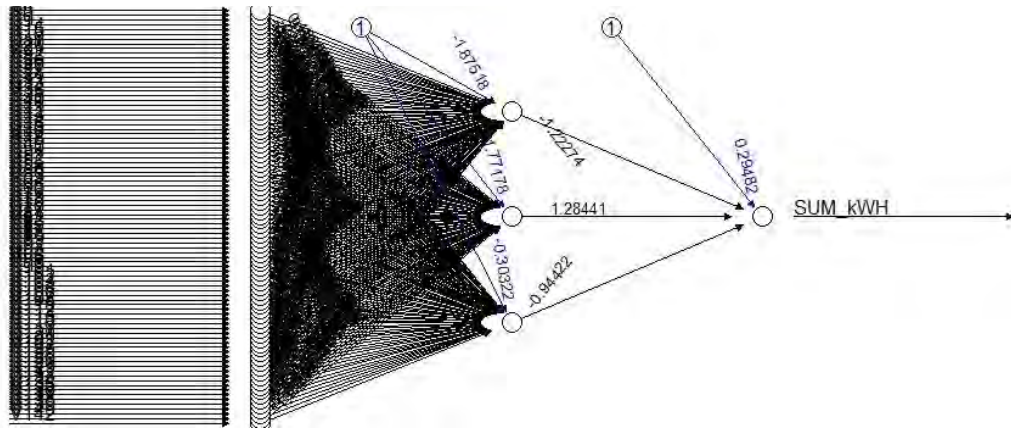
Figure 1: The structure of the ANN model for predicting electricity consumption

The fitted model was used to generate predictions on the test sample, and the calculated MAPE was 30.5516. The varImp function in R software was used to obtain variable importance. In case of ANN, it is based on Gevrey et al [9], which uses combinations of the absolute values of the weights. The procedure extracted 20 variables as important predictors that are listed in Table 1. It can be seen that the ANN model considers *V139 (object geo type)* as the most important predictor, followed by *V9 (heated surface of the building)*, and *V60 (total number of compact fluorescent luminaries)*. Also, variables representing cooling power and heating capacity are extracted, then construction variables, as well as the number of employees representing a usage of buildings.

The CART decision tree model has produced the MAPE of 33.7324%, with the error conversion for different values of *cp* parameter as shown in Figure 2. The plot of the pruned CART tree is shown in Figure 3. It can be seen that only five predictors were left out in the pruned tree, namely *V9 (heated surface of the building)* in the first node of the tree, followed by the variable *V40 (total building power of cooling in kW)* in the left branch, and variables *V17 (number of working hours per workday)* and *V5 (number of users)* in the right branch.
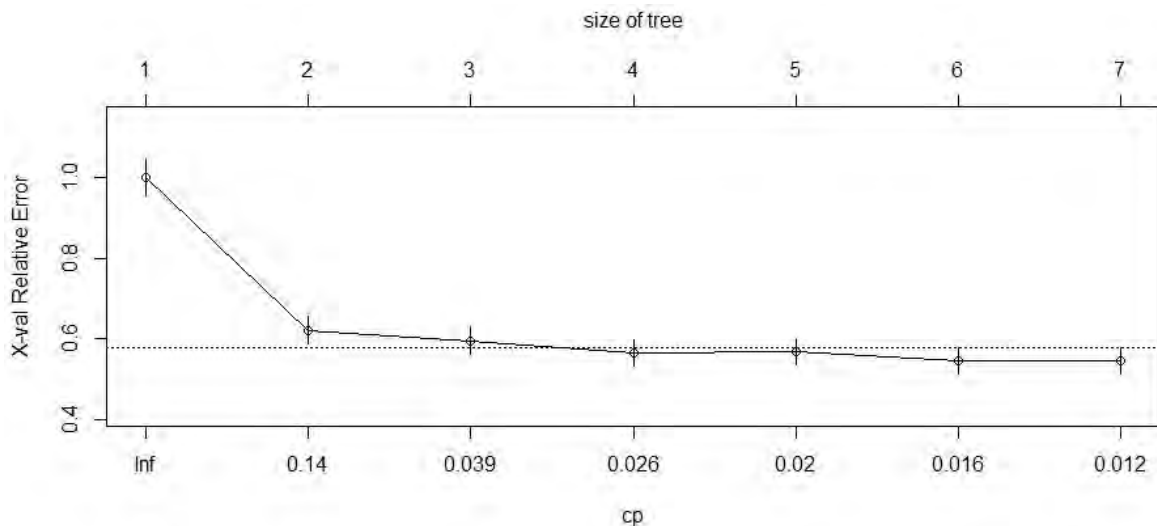
Figure 2: Error conversion of CART decision tree for predicting electricity consumption

12
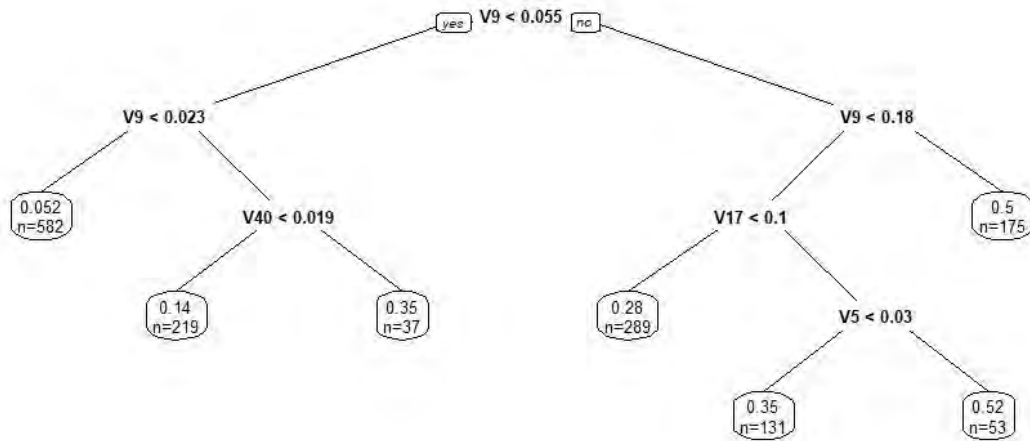
Figure 3. Pruned CART decision tree for predicting energy consumption

When conditional inference tree method (CTREE) was used to build the prediction model, it retained a larger number of variables than the CART method. Graphical presentation of CTREE in Figure 6 shows that the final tree contains 15 split variables in 20 nodes, also extracting *V9 (heated surface of the building)* as the most important variable for splitting the tree. Besides that, the CTREE nodes contain all the other variables also used in CART split and additional 10 extracted predictors listed in Table 1. The CTREE model produces the MAPE of 31.2056 %, which is lower than the MAPE of ANN and CART models. The random forest (RF) method produced the highest accuracy out of five tested methods. The RF MAPE was 18.8015%, it was obtained by using 500 trees in the forest, and the graph of error convergence in generated trees is shown in Figure 4.
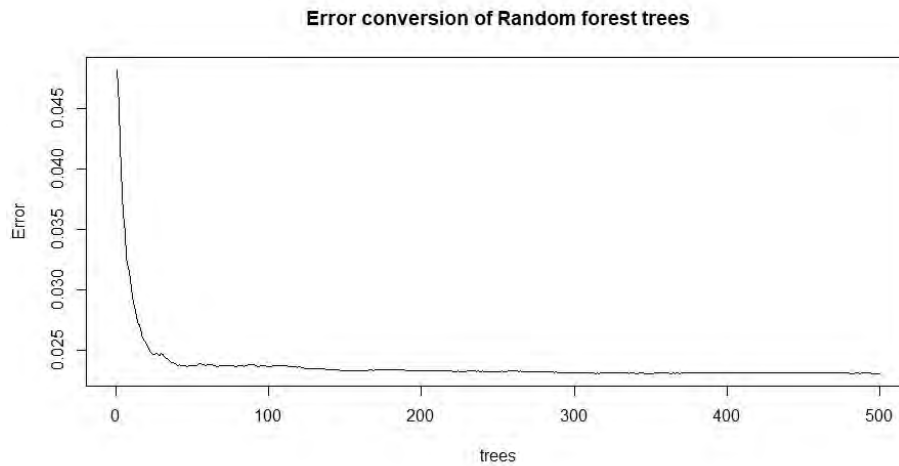
Figure 4: Error conversion of Random forest modeling

To compute variable importance, the RF uses a measure called the total decrease in node impurities measured by the Gini index attributed to each variable at each split, averaged over all trees. With the cut-off of 1, and the model extracted 15 predictors. The list of predictors ordered by importance is given in Table 1, and it can be seen that the variable with the highest impact is *V9 (heated surface of the building),* matching the selection made by CART and CTREE. The next two extracted variables are related to the usage of building *V4 (number of employees)* and *V5 (number of users),* followed by six variables that were not extracted by

previous models: *V29 (total installed thermal power of heaters in kW), V70 (number of interior light luminaries), V35 (total cooling power of coolers), V33 (total heating power), V85 (the maximum allowed coefficient of transmission loss of heat), V21 (building shape factor), and V15 (number of floors)*. The last five predictors extracted by the RF model were also extracted by CTREE (see Table 1). It can be seen that RF and CTREE match in 10 extracted variables, while RF and CART match in 4 extracted variables.

The last tested method was the SVM, which produced the model with MAPE of 28.5691% and extracted six predictors. Like in all previous models except in ANN, a construction variable *V9 (heated surface of the building)* was the most important. It is interesting that the SVM model extracted three attributes that describe the usage of a building: *V4 (number of employees), V5 (number of users),* and *V6 (number of working days per week)*, as well as another construction variable: *V2 (share of use of total building area in m^2)*. The MAPE errors and extracted variables of all five models are presented in Table 1.

Table 1: Results of machine learning models for electricity prediction obtained on the validation sample

Model	*SMAPE (%)*	*Selected features (ordered by importance)*
ANN	30.5516	*V139 (object geo type)* *V9 (heated surface of the building)* *V60 (total number of compact fluorescent luminaries)* *V33 (total heating power in kW)* *V34 (total body coolers)* *V22 (total heat capacity of heat pump)* *V16 (average yearly temperature)* *V17 (number of working hours per workday)* *V87 (annual thermal energy needed for heat)* *V14 (cooled volume area of the building in m^2)* *V32 (installed heat power of split system for heat)* *V94 (object construction roof surface)* *V6 (number of working days per week)* *V64 (total number of halogen luminaries)* *V44 (total heating capacity in kW)* *V54 (total number of luminaries with incandescent)* *V15 (number of floors)* *V103 (object construction thickness of windows)* *V89 (object cooling area)* *V4 (number of employees)*
CART	33.7324	*V9 (heated surface of the building)* *V40 (total building power of cooling in kW)* *V17 (number of working hours per workday)* *V5 (number of users)*
CTREE	31.2056	*V9 (heated surface of the building)* *V31 (installed electric power of split system for heating in kW)* *V40 (total building power of cooling in kW)* *V14 (cooled volume area of the building in m^2)* *V132 (cool energy generating product code)* *V81 (total installed power of office equipment in kW)* *V16 (average yearly temperature)* *V17 (number of working hours per workday)* *V125 (climate data – region)* *V135 (domestic hot water fuel-1 code)* *V50 (total installed electric power of domestic hot water in kW)* *V51 (tank capacity central in liters)* *V6 (number of working days per week)* *V4 (number of employees)*

		V5 (number of users)
		V82 (total installed power of kitchen equipment in kW)
Random forest	18.8015	V9 (heated surface of the building), V4 (number of employees) V5 (number of users) V29 (total installed thermal power of heaters in kW, V70 (number of interior light luminaries) V35 (total cooling power of coolers) V33 (total heating power) V85 (the maximum allowed coefficient of transmission loss of heat) V21 (building shape factor) V15 (number of floors) V81 (total installed power of office equipment in kW) V14 (cooled volume area of the building in m^2) V40 (total building power of cooling in kW) V16 (average yearly temperature) V6 (number of working days per week)
SVM	28.5691	V9 (heated surface of the building), V4 (number of employees) V5 (number of users) V14 (cooled volume area of the building in m^2) V6 (number of working days per week) V2 (share of use of total building area in m^2)

It can be seen from Table 1 that the most accurate model was produced by the RF method of recursive partitioning. The t-test of difference in proportions shows that the RF model is significantly different (p=0.0001) from the closest one in accuracy (SVM model), therefore the RF MAPE is significantly better than the MAPEs of other models.

6 DISCUSSION AND CONCLUSION

The experiments conducted in this research leverage the performance of five machine learning methods in predicting energy consumption. The procedure included data preprocessing stage and the modeling stage. In case of using real datasets in domain of energy consumption, a modeller needs to deal with the problems of high-dimensionality, missing data, outliers, and inter-correlations, and use appropriate methods regarding the nature of data. Due to a high dimension of input space and presence of continuous as well as categorical data, the variable reduction procedure in our research included selection based on inter-correlations and chi-square test of independence. The reduced input was used by five machine learning methods to model the prediction of yearly electricity consumption of buildings based on their construction data, geo-spatial data, heating, cooling, and usage data. All the methods used the same normalized training sample to fit the model, and the same test sample to validate their result. The significantly lowest MAPE error was produced by the random forest method with 500 trees, extracting 15 variables as important predictors. It can be noticed that there was a certain overlap in the set of extracted features by different methods. The four out of five methods recognize the *heated surface of the building* as the most important variable. Also, both RF and SVM put the *number of employees* to the second place of importance. In accordance with previous research, the most successful model has shown that construction characteristics, geo-spatial data, energy data, as well as usage are important in predicting yearly energy consumption. However, unlike in other authors' research, the ANN model was not the most successful one. The significant over performance of random forest algorithm could be caused by high dimensionality of data, which requires more complex models. In this research, only the electricity consumption was observed in the output since this form of energy is used in all buildings in the dataset. In future research, other types of energy consumption, such as natural

gas and hot water will be included to create a more general model of energy consumption. Besides predicting the total yearly energy consumption, monthly consumptions and their dynamic behavior in relation to reconstruction measures should be investigated in the future in order to create a methodological framework for decision makers that could assist in allocating state resources to obtain savings in energy consumptions and healthier environment.

References

[1] Apte, C. and S. Weiss, 1997. Data Mining with Decision Trees and Decision Rules. Future Generation Computer Systems, 13: 197-210.

[2] Behzad, M., Asghar, K., Eazi, M. and Palhang, M., 2009. Generalization performance of support vector machines and neural networks in runoff modeling. Expert Systems with Applications, 36: 7624–7629.

[3] Breiman, L., Friedman, J., Olshen, R., and Stone, C. 1984. Classification and Regression Trees, Belmont, CA: Wadsworth International Group.

[4] Chapelle O., Vapnik, V. 1999. Model selection for support vector machines. In Advances in Neural Information Processing Systems, 12 (Solla A., Leen T. K. and Mueller, K., eds.), 230-236.

[5] Chou, J.S., Bui, D.K. 2014. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy and Buildings, 82: 437-446,

[6] Chung, M., Park, H.C. 2015. Comparison of building energy demand for hotels, hospitals, and offices in Korea. Energy, 92, pp. 383-393, DOI: 10.1016/j.energy.2015.04.016

[7] Fan, G. Gray, J. B. 2005. Regression Tree Analysis Using TARGET. Journal of Computational and Graphical Statistics, 14 (1): 206-218.

[8] Farzana, S., Liu, M., Baldwin, A., Hossain, M.U. 2014. Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China. Energy and Buildings, 81: 161-169.

[9] Gevrey, M., Dimopoulos, I., Lek, S. 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecological Modelling, 160(3): 249-264.

[10] Grömping, U. 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. The American Statistician, 63(4): 308-319.

[11] Hartshorn, S. 2016. Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners. Amazon Digital Services LLC: Seattle, Washington, USA.

[12] Hothorn, T., Hornik, K., Zeileis, A. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics, 15(3): 651-674.

[13] Liang, X., Hong, T.Z., Shen, G.Q. 2016. Improving the accuracy of energy baseline models for commercial buildings with occupancy data. Applied Energy, 179: 247-260, doi: 10.1016/j.apenergy.2016.06.141.

[14] Mangold, M., Osterbring, M., Wallbaum, H. 2015. Handling data uncertainties when using Swedish energy performance certificate data to describe energy usage in the building stock. Energy and Buildings, 102: 328-336, doi: 10.1016/j.enbuild.2015.05.045.

[15] Masters, T. Advanced Algorithms for Neural Networks. New York: John Wiley & Sons, 1995.

[16] Paliwal, M. and Kumar U.A., 2009. Neural networks and statistical techniques: A review of applications. Expert Systems with Applications, 36: 2–17.

[17] Smola A., Schölkopf, B. 2004. A tutorial on support vector regression. Statistics and Computing, 14: 199–222.

[18] Son, H., Kim, C., Kim, C., Kang, Y. 2015. Prediction of government-owned building energy consumption based on an RReliefF and support vector machine model, Journal Of Civil Engineering And Management, 21(6): 748-760.

[19] Tofallis, C. 2015. A better measure of relative prediction accuracy for model selection and model estimation. Journal of the Operational Research Society, 66(8): 1352-1362.

[20] Tommerup, H., Rose, J., Svendsen, S. 2007. Energy-efficient houses built according to the energy performance requirements introduced in Denmark in 2006. Energy and Buildings, Volume 39(10): 1123-1130.

[21] Zekić-Sušac, M. 2017.Overview of prediction models for buildings energy efficiency. Proceedings of the 6th International Scientific Symposium Economy Of Eastern Croatia – Vision and Growth. Mašek Tonković Anka (ed.) Osijek: Faculty of Economics in Osijek, 25.05. - 27.05.2017. 697 -706.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 1:
Advances in Modelling and Statistical Research of the Western Balkan Countries in the Times of Economic Crisis

A NOTE ON HOUSING WEALTH EFFECT IN SELECTED WESTERN BALKAN COUNTRIES

Anita Čeh Časni
University of Zagreb, Faculty of Economics and Business, Department of Statistics
10 000 Zagreb, Croatia
E-mail: aceh@efzg.hr

Abstract: The aim of this study was to explore empirically whether there is a direct housing wealth effect on personal consumption in the long and the short run in selected countries. For that purpose the estimator for nonstationary heterogeneous dynamic panels was employed. Namely, pooled mean group estimator that allows for heterogeneous short-run dynamics and common long-run income and housing wealth is used. The results of the empirical analysis reveal that for the selected group of countries there is statistically significant long-run relationship between consumption, income and housing wealth which is in line with life-cycle theory.

Keywords: direct housing wealth effect, dynamic panels, pooled mean group estimator, Western Balkan countries

1 INTRODUCTION

The main objective of this paper is to explore the influence of permanent changes in housing wealth on consumption in 3 Western Balkans countries (Croatia, FYR Macedonia and Turkey) and 10 Central and Eastern European Countries (Bulgaria, Czech Republic, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia and Slovenia[1]) in the short and long run. Generally, empirical literature studying the impact of housing and stock market wealth on consumption can be broadly divided in two categories. One group of papers models direct wealth effect using aggregate macroeconomic data, while the other group of papers assesses indirect wealth effect using disaggregated (usually household level) data.

In this study I follow the former approach which assumes that rising asset prices (house prices) increase household wealth, which in turn increases consumption via the budget constraint. Direct wealth effect is most often modeled using cointegration and error correction models which allow one to distinguish between the short run and the long run relationship between consumption, income and wealth. Moreover, this approach identifies the variables that adjust after the shock in order to restore the long-run equilibrium.

Besides aforementioned two groups of papers, there is another part of literature that deals exclusively with housing wealth effect on consumption in industrialized countries. However, studies like [4, 6, 8, 11, 16] differ greatly on the exact effect of housing wealth on consumption.

As far as studies on housing wealth effect in post-transition countries are concerned, to the best of my knowledge only few papers exist, mostly due to data availability that prevent complete and effective empirical analysis. Very few recent studies provide evidence of significant housing wealth effect in European post-transition countries, [1, 9, 10, 18] to name a few.

In accordance with the objective and presented overview, this study proposes to: model the impact of permanent changes in housing wealth on consumption in Western Balkan, Central and Eastern European Countries and differentiate between short run versus long run impact of changes in wealth by applying panel cointegration and error correction models introduced by [17]. This study makes use of the previous study by [9] but presents the new

---

[1] Bulgaria, Romania and Slovenia are sometimes grouped into the Balkans.

estimates of housing wealth effect on personal consumption by enlarging the sample of countries focusing on the Balkans and by using annual data series from altered data sources.

The structure of this paper is as follows. After the brief introduction, the methodology is presented in section two. In section three the data is presented. Section four gives the results of the empirical analysis and section five concludes.

2 METHODOLOGY

The main purpose of this study is to assess the relative importance of housing wealth, across post-transition countries of Western Balkan and CEE countries. Wealth is important for consumption behavior because in applied macroeconomics literature, consumption function usually involves household income and wealth. This rather simple consumption function model is motivated by several theories, comprising the permanent income theory proposed by [12] and the life cycle theory by [2]. Cointegration and error correction methods are typically the methods of choice when modelling wealth effect on consumption since the basic prediction of the life-cycle model of household spending suggests that expected changes in asset prices should not lead to changes in planned consumption, while unexpected changes should generate a response.

In this study, method for cointegrated panels developed by [17] will be applied to the analysis of the relationship between consumption, income and housing wealth. Before testing for panel cointegration, test for panel unit roots using one or more tests suggested by [5] will be employed. Pooled mean group (PMG) estimator proposed by [17] is particularly attractive since it pools long run relationships between countries while short run responses are flexible and unrestricted across countries. In other words, PMG estimator allows imposing a cross-country restriction on the long-run marginal propensity to consume out of wealth (mpcw), while taking into account possible cross-country differences in the adjustment. Hence, PMG estimate of the mpcw provides a theory-consistent guide to the wealth effects on consumption superior to estimates obtained from traditional panel methods [14].

The benchmark specification will pool all countries and all time periods. Thus, the empirical analysis starts with the following panel long run equilibrium relationship[2]:

$$C_{t,i} = \delta_{0,i} + \gamma_i C_{t-1,i} + \beta_{10,i} W^h{}_{t,i} + \beta_{20,i} Y_{t,i} + \varepsilon_{t,i} \tag{1}$$

where $C_{t,i}$ is private consumption, $W^h{}_{t,i}$ refers to housing wealth, $Y_{t,i}$ is income and subscripts i and t denote country and time, respectively. The error term entailing the effects of unanticipated shocks to consumption is denoted by $\varepsilon_{t,i}$.

Rewriting equation (1) yields the panel error correction specification:

$$\Delta C_{t,i} = \alpha_{0i} + \phi_i\left(C_{t-1,i} - \alpha_{1i} W^h{}_{t,i} - \alpha_{2i} Y_{t,i}\right) + \beta_{10i}\Delta W^h{}_{t,i} + \beta_{20i}\Delta Y_{t,i} + \varepsilon_{t,i} \tag{2}$$

where long run parameters α_{0i}, α_{1i}, α_{2i} and the adjustment coefficient (ϕ_i) are calculated from equation (1) parameters.

3 DATA

Although this analysis entails only three variables, there are several limitations related to the data availability. In order to conduct the analysis annual data series on consumption, income

---

[2] The choice of a country specific lag order of panel ARDL model for the sample of analysed countries, concidering the annual data and according to SBC information criterion is (1,0,0). Thus, personal consumption is lagged once and disposable income and house prices are given in levels.

and housing wealth for the panel of 13 countries are required. Out of these 13 countries, 10 countries are post-transition European countries (Bulgaria, Czech Republic, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia and Slovenia) and 3 countries are Western Balkans countries (Croatia, FYR Macedonia and Turkey). The data set consists of annual data spanning, when available, from 1990 to 2016. The private final consumption expenditure, gross wages and gross national disposable income are expressed in milliards of Euros and are taken from AMECO and Vienna institute for Economics (WIIW) database. Apart from that, real estate price indices are taken from BIS database and are all recalculated to the same base (2015=100). Furthermore, all data is recalculated into logarithms, so the parameters can be interpreted as elasticities of dependent variable (consumption) to changes in independent variables (income and housing wealth).

House price series is used as a proxy for housing wealth since housing wealth series (or housing stock series from which housing wealth series are derived) are not available for most post-transition countries in a panel. House prices were also used as a proxy for housing wealth in studies on wealth effect on consumption in [3, 13, 14]. Given the fact that the span of house price series varies across countries, the panel will be unbalanced.

Data on total aggregate consumption are used, although in the empirical literature non-durable consumption series are also used. Namely, several authors including [7] and [15] advocate using total consumption series when testing for wealth effect because the stock market crashes usually only affect durable consumption.

Finally, two data series for income will be used. Consumption theories suggest using only labor income, but given the lack of comparable labor income series for the countries in this data sample, I intend to use disposable income series published by AMECO database and wage series published by WIIW database. This approach will enable performing additional robustness checks.

4 THE RESULTS OF THE EMPIRICAL ANALYSIS

Following PMG modelling approach, firstly, the panel unit root tests for the variables of interest were employed. Namely, tests that assume individual unit root processes: Im, Pesaren and Shin (IPS) test, Augmented Dickey Fuller (ADF) test and Philips-Perron (PP) test were conducted and resulted in non-rejecting the null hypotheses of a unit root for all the variables of interest. Thus, consumption, income (wage) and housing wealth were integrated of order one or difference stationary. After performing the panel unit root tests, panel cointegration tests were employed. Namely, Johansen-Fisher Panel cointegration test, Kao Residual cointegration test and Augmented Dickey-Fuller test showed that consumption, income and housing wealth were cointegrated. Additionally, Westerlund panel cointegration tests based on structural dynamic were conducted and showed that personal consumption and housing wealth are cointegrated in the long-run, since the null hypothesis of no cointegration is strongly rejected in all cases with statistical significance of 1% and 5%[3].

The estimation result of consumption model given by expression (2) is shown in Table 1. Accordingly, the adjustment coefficient for analysed panel of 13 countries has the expected negative sign and is statistically significant on 10% significance level. Thus, the error-correction mechanism is in place and the long-run equilibrium is reached in about four years. Also, one can notice that housing wealth effect on personal consumption does exist in the long-run as well as in the short-run, with both coefficients being statistically significant on 1% significance level and having the correct positive sign. Namely, in the long-run the

---

[3] In order to save space, panel unit root tests and panel cointegration tests are not presented here, but are available from the author upon the request.

elasticity of consumption to changes in housing wealth is 0.111 and in the short-run the consumption is more responsive to changes in housing wealth, with the elasticity coefficient of 0.279. Furthermore, the consumption is also responsive to the changes in income, which is again more pronounced in the short-run (0.862) compared to the long-run (0.467) according to the estimated elasticity coefficient.

Table 1: PMG Estimation results

	PMG	
	Long-run coefficients	*Short-run coefficients*
Speed of adjustment, ϕ_i	-	-0.257*
Housing wealth	0.111***	0.279***
	(0.018)	(0.066)
Income	0.467***	0.862***
	(0.062)	(0.178)
constant	-	0.052
		(0.059)
Number of observations	128	
Number of countries	13	
Hausman test of poolability of countries	0.21	

Notes: estimations are performed using PMG estimator of Pesaran et al. (1999); all equations include a constant term, standard errors are in brackets, ***,**,* denote significance at 1%, 5% and 10% significance level, respectively
Source: Author`s calculations.

Since the PMG estimation procedure allows for short-run heterogeneity, it is possible to estimate short-run country-specific error correction models. The result of this econometric exercise is given in table 2. Accordingly, there is a housing wealth effect on personal consumption in Bulgaria, Latvia, Lithuania, Hungary and Macedonia. In Estonia, the short-run housing wealth effect has the negative sign and it is very small (around zero), so one can conclude that there is no housing wealth effect on personal consumption in the short-run in Estonia. Also, according to the estimated model, the most pronounced short-run housing wealth effect is recorded for Hungary (0.745), Latvia (0.435) and for Western Balkan country- FYR Macedonia, with the elasticity coefficient of 0.373. For the other analysed countries, there is no statistically significant housing wealth effect in the short-run. This can be due to higher transaction costs that prevent conversion of housing wealth into the money that can be used for consumption. Additionally, consumption is responsive to short-run changes in income, with properly signed and statistically significant coefficients recorded for: Czech Republic, Estonia, Croatia, Lithuania, Hungary and Slovakia, with the highest coefficient recorded for Estonia (1.976).

Table 2: Short-run country-specific estimates of personal consumption model

Country	Full PMG			
	ϕ_i	Δw^h	ΔY	constant
Bulgaria	-0.866***	**0.293***	-0.051	0.249**
	(0.258)	(0.0.075)	(0.028)	(0.104)
Czech Republic	-0.480***	0.004	1.203***	0.265***
	(0.158)	(0.144)	(0.152)	(0.100)
Estonia	-1.152***	**-0.074***	1.976***	0.039
	(0.203)	(0.028)	(0.403)	(0.074)
Croatia	-0.411***	0.301	1.588***	0.137**
	(0.116)	(0.220)	(0.504)	(0.055)
Latvia	-0.305***	**0.435***	0.025	0.029
	(0.083)	(0.052)	(0.125)	(0.025)
Lithuania	-1.193***	**0.235***	0.568***	0.064***
	(0.045)	(0.036)	(0.168)	(0.017)
Hungary	-0.272***	**0.745***	1.195***	0.065*
	(0.064)	(0.123)	(0.165)	(0.034)
Poland	0.384**	0.129**	1.408***	-0.340**
	(0.171)	(0.062)	(0.155)	(0.159)
Romania	0.924***	0.044***	1.069*	0.453**
	(0.317)	(0.015)	(0.573)	(0.199)
Slovenia	-0.235	0.202	0.618	0.038
	(0.210)	(0.127)	(0.423)	(0.041)
Slovakia	-0.017	0.278***	0.906***	0.011
	(0.068)	(0.064)	(0.185)	(0.025)
FYR Macedonia	-0.869***	**0.373***	-0.010	-0.301***
	(0.223)	(0.117)	(0.417)	(0.098)
Turkey	0.148	0.664	0.796***	-0.030
	(0.160)	(0.695)	(0.256)	(0.026)

Notes: ***,**,* denote significance at 1%, 5% and 10% significance level, respectively; numbers in brackets are standard errors for full PMG
Source: Author`s calculations.

With the aim of the robustness check[4] of the estimated model presented in tables 1 and 2, another variable for income (wages) was used. Accordingly, the results were somewhat the same, confirming the importance of housing wealth effect in short-run and in the long-run for the analysed countries.

5 CONCLUSION

This empirical paper generated several conclusions. Namely, the results of the estimated panel error correction model suggest that personal consumption, housing wealth and income form a long-run equilibrium relationship in the selected Western Balkans and CEE countries. Furthermore, housing wealth effect on personal consumption does exist in the long-run as well as in the short-run, with both coefficients being statistically significant on 1% significance level and having the correct positive sign. Also, the short-run housing wealth effect has shown to be more pronounced than the long-run housing wealth effect. Finally, the long-run housing wealth effect in the baseline model as well as in alternative model specification suggests that the Western Balkan and CEE countries are sensitive to the developments in housing sector and the policy makers should take this into account.

---

[4] The robustness check results are not presented here, but are available from the author upon the request.

References

[1] Ahec Šonje A., Čeh Časni A., Vizek M. 2012. Does housing wealth affect private consumption in European post transition countries? Evidence from linear and threshold models. Post-communist Economies, Vol. 24, No. 1, pp. 73-85.

[2] Ando, A., Modigliani, F. 1963. The "Life Cycle "Hypothesis of Saving: Aggregate Implications and Tests, The American Economic Review, 53 (1), 55-84.

[3] Aoki, K., Proudman, J., Vlieghe, G. 2003. House prices, consumption, and monetary policy: a financial accelerator approach. Bank of England Working Paper Series, No. 169, pp. 1-38.

[4] Attanasio, O., Blow, L., Hamilton, R., Leicester, A. 2009. Booms and busts: consumption, house prices and expectations. Economica, Vol. 76, No. 301, pp. 20-50.

[5] Baltagi, B., Kao, C. 2000. Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey, CPR Working Papers, No. 16.

[6] Bover, O. 2005. Wealth effects on consumption: microeconometric estimates from the Spanish survey of household finances. Documentos de Trabajo No. 0522, Banco de Espana.

[7] Brady, P. J., Canner, G. B., Maki, D. M. 2000. The effects of recent morgage refinancing, Federal-Reserve-Bulletin, 86(7), 441-450.

[8] Campbell, J., Cocco, J. 2007. How do house prices affect consumption? Evidence from micro data, Journal of Monetary Economics, 54, 591-621.

[9] Čeh Časni, A. 2014. Housing Wealth Effect on Personal Consumption: Empirical Evidence from European Post-Transition Economies. Czech Journal of Economics and Finance, Vol. 64, No. 5, pp. 392-406.

[10] Čeh Časni, A. 2016. Is there a housing wealth effect in European countries? Croatian Review of Economic, Business and Social Statistics, 2(2).

[11] Disney, R., Bridges, S., Gathergood, J. (2006). Housing Wealth and Household Indebtedness: Is there a Household 'Financial Accelerator?, CNB Working Paper Series, No. 12.

[12] Friedman, M. 1957. A theory of the Consumption function. Princeton University Press, Princeton.

[13] Girouard, N., Blöndal, S. 2001. House prices and economic activity, OECD Economics Department Working Papers No. 279.

[14] Ludwig, A., Sløk, T. 2004. The Relationship between Stock Prices, House Prices and Consumption in OECD Countries.

[15] Mehra, Y. P. 2001. The wealth effect in empirical life-cycle aggregate consumption equations. Federal Reserve Bank of Richmond Economic Quarterly, Vol. 87, No. 2, pp. 45- 68

[16] Morris, E. 2007. Examining the wealth effects from home price appreciation. Job-Market Paper, University of Michigan.

[17] Pesaran, H., Shin, Y., Smith, R. P. 1999. Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. Journal of the American Statistical Association, Vol. 94, No. 446, pp. 621-634.

[18] Seč, R., Zemčík, P. 2007. The Impact of Mortgages, House Prices and Rents on Household Consumption in the Czech Republic, CERGE-EI Discussion Paper, 2007-185.

INTER-INDUSTRY DIFFERENCES IN CAPITAL STRUCTURE: THE EVIDENCE FROM BOSNIA AND HERZEGOVINA

Ksenija Dumičić
University of Zagreb, Faculty of Economics and Business – Zagreb
Trg J. F. Kennedyja 6, HR-10000 Zagreb, Croatia
E-mail: <kdumicic@efzg.hr>

Emina Resić
University of Sarajevo, School of Economics and Business Sarajevo
Trg oslobođenja - Alija Izetbegović 1, 71000 Sarajevo, Bosnia and Herzegovina
E-mail: <emina.resic@efsa.unsa.ba>

Jasmina Mangafić
University of Sarajevo, School of Economics and Business Sarajevo
Trg oslobođenja - Alija Izetbegović 1, 71000 Sarajevo, Bosnia and Herzegovina
E-mail: <jasmina.mangafic@efsa.unsa.ba>

Abstract: This paper deals with the empirical investigation of the existence of inter-industry differences in the capital structure of entrepreneurial, non-financial firms in a transitional economy. Such research contributes towards a better understanding of their financial behavior. The technique used for this cross-sectional analysis is nonparametric analysis of variance. The aim was to test the hypothesis that the firms in the same industry have similar capital structure, since they are facing relatively similar economic conditions. The paper demonstrates significant differences among capital structure depending on the industry where the company operates.

Keywords: Capital Structure, Industry Effect, Non-parametric analysis of variances, Financial Leverage, Balance-Sheet Leverage

1 INTRODUCTION

Apparently and at the same time, unfortunately, there is no magic proportion of debt that a company can take on. The debt-equity relationship varies according to industries involved, a company's line of business and its stage of development. Since a number of factors influence the capital structure decision of a company, the judgment of the person making the capital structure decision plays a crucial part. Two similar companies will have different capital structures, should the decision makers differently judge the significance of various factors. That is why, a theoretical model alone perhaps cannot adequately handle all those factors, which affect the capital structure decision. These factors are highly psychological, complex and qualitative and do not always follow the accepted theories, since capital markets are not perfect and the decision has to be make in perfect knowledge and risk.

The relating vast amount of literature is primarily written by Anglo-Saxon authors and the capital structure choice was empirically mostly examined on the first listed in developed countries that share many institutional similarities [2]. Comparatively, little attention is devoted to developing economy context, especially in transitional countries. Bosnia and Herzegovina is no exception.

Considering the market imperfections typical of transitional economies, the aim of this study is to contribute to the existing empirical financial literature on the inter-industry differences in capital structure in the context of entire population of privately owned non-financial companies in Bosnia and Herzegovina (BiH), for the period of ten years (2003-2012). Since the assumptions required for a parametric analysis of variances are not satisfied, a nonparametric method is applied to test for cross-industry differences in the capital structure. The results show that leverage varies significantly across industries.

2 DATA AND METHODS

The Study addressed the following specific research questions:

(i) Are there any differences across the industries when it comes to the capital structure of FBiH firms? and

(ii) Does leverage vary across different industries?

The study was conducted on a sample of non-financial BiH corporations coming from different industries. The Company Law (the Law on Companies/Law on Enterprises) at both BiH entities[1] classifies corporations in the same manner, as joint stock companies (*JSC*) and limited liability companies (*Ltd*) [4]. The data set used contains both listed and non-listed companies and therefore we use the book values from the companies' financial statements. Consequently, after the initial screening, our dataset contains a total of 179.330 firm-year observations over a ten-year period from 2003 to 2012. This is the final sample on which further analysis is based. The panel is unbalanced and not all firms are present in all observation years. The sample reveals, as the population itself, the dominance of companies in the wholesale and retail trade, manufacturing and real estate service industries, and Ltd as a domineering form of organization comprises at least 96% of the sample in each observed year.

Similar to the competing theories, there is no universally accepted definition of capital structure in the literature. Many different forms of debt, equity, and different blends of the two exist. Therefore, the appropriate definition is not obvious as to which kind of debt-to-equity ratio should be used in the empirical research. Researchers are in an agreement that the measures of capital structure should be different, depending on the analysis purpose [1, 3]. Before bringing the final decision on the measure used in this study we want it to check whether the non-financial liabilities are small enough that they can be ignored. Therefore we wanted to take a look at the size and structure of debt in the FBiH companies. In other words we took at their financial structure.

When looking at the financial vs. non-financial component of total liabilities, the importance of the latter is obvious, comprising up to 67%. That resulted in the following dilemma. If we want to include the non-financial liabilities meaning to treat financial and non-financial liabilities alike, then we should use the ***liabilities-to-asset ratio*** (balance sheet leverage). In case of liquidation, this measure can be seen as a proxy of the value that remains for the equity holders. Still, this measure does not serve as a good indicator of whether the company faces an impending a risk of default. Due to the fact that total liabilities also include the items such as accounts payable, utilized to make transactions rather than for financing, it is likely that the leverage level is overstated. Even more so, this leverage proxy might be affected by the provisioning and reserves, such as the obligations for the retirement funds. If we still want to focus only on financial leverage, then we should use the ***financial debt-to-capital ratio*** (financial leverage) where the denominator is financial debt plus equity. This measure of leverage observes the "employed" and hence it best demonstrates the impact of the financing decisions. Also, the measure is directly linked to the debt associated agency problems. Mindful of all these, we decided to include in our analysis the ratio that treats both type of debts equally to see are there variables for which the inference depends on the measure of leverage. We used this alternative definition of leverage in order to examine the robustness of the effects of determinants on the leverage decisions of firms.

---

[1] BiH is made of two entities (the Federation of BiH and the Republika Srpska) and the Brčko District.

3 RESEARCH METHODOLOGY AND EMPIRICAL RESULTS

The next step is to make a direct probe into the variations across 16 industries in the basic capital structure in the sampled companies (the industry classification is given in the Appendix 1). No prior empirical research within the BiH setting has examined the heterogeneity in the basic capital structure across various industries. The main purpose is to test the hypothesis that firms in the same industry, as they face similar economic conditions, have similar capital structures. The analysis is based on two alternative ratios as measures for the capital structure: the balance-sheet leverage and financial leverage. Since the normality assumption required for a parametric analysis of variance are not satisfied by the data (Table 1), a non- parametric test called the Kruskal-Wallis one-way analysis of variances is applied. It tests the null hypothesis that the samples are from identical populations with the same median against the hypothesis that at least one pair of groups has different medians. The Table 1 reports the sample statistics for the ten years average of **financial leverage** measured by the ratio of financial debt to capital for the individual industries and the entire sample.

Table 1: Kolmogorov-Smirnov Test Results (Financial Leverage) for industries following the classification given in Appendix 1

Financial leverage	Total	A	B	C	D	E	F	G	H
Mean	.301	.302	.334	.278	.302	.235	.259	.338	.320
Median	.118	.136	.099	.132	.164	.034	.085	.191	.052
Std. Deviation	.358	.353	.381	.336	.341	.354	.331	.369	.400
Skewness	.831	.852	.593	1.093	.861	.339	1.083	.627	.765
Kurtosis	-.821	-.748	-1.310	-.181	-.652	.150	-.190	-1.151	-1.147
KS value	.342	.197	.248	.204	.188	.280	.217	.198	.252
p value	.000	.000	.000	.000	.000	.000	.000	.000	.000

Financial leverage	Total	I	J	K	L	M	N	O	P
Mean	.301	.310	.280	.223	.243	.231	.124	.260	.075
Median	.118	.150	.000	.000	.000	.000	.000	.001	.000
Std. Deviation	.358	.354	.390	.341	.487	.366	.270	.367	.219
Skewness	.831	.745	.948	1.283	2.000	1.223	2.288	1.090	2.996
Kurtosis	-.821	-.931	-.848	.117	4.000	-0.205	3.939	-.441	8.982
KS value	.342	.200	.308	.304	.441	.352	.360	.273	.493
p value	.000	.000	.000	.000	.000	.000	.000	.000	.000

The results of the Kolmogorov-Smirnov test are also reported in the Table 2. Since the assumptions of normal distribution are rejected and all 16 samples have very skewed distributions, the nonparametric method is more appropriate. We apply the Kruskal-Wallis non parametric analysis of variances to test the null hypothesis of no variation in leverage across industries during 2003-2012.

The Kruskal-Wallis non-parametric analysis of variances yields the p-value equal to zero which reject the hypothesis that all the samples from the same populations. This means that at least one pair of group has different medians. In other words, there is a difference in leverage across industries during 2003-2012. We found that capital structure positions among industries (interindustry) have significant differences according to statistical evidence using a non-parametric analysis of variance (Table 2).

Table 2: Kruskal-Wallis Test Results (Financial Leverage)

Year	Financial leverage		Year	Financial leverage	
2003	Chi-Square	271.522	2008	Chi-Square	282.503
	p value	.000		p value	.000
2004	Chi-Square	338.551	2009	Chi-Square	421.269
	p value	.000		p value	.000
2005	Chi-Square	429.529	2010	Chi-Square	313.005
	p value	.000		p value	.000
2006	Chi-Square	322.125	2011	Chi-Square	338.608
	p value	.000		p value	.000
2007	Chi-Square	340.443	2012	Chi-Square	294.476
	p value	.000		p value	.000

The Table 3 reports the sample statistics for the ten years average of **balance-sheet leverage** measured by the ratio of financial debt to capital for the individual industries and the entire sample.

Table 3: Kolmogorov-Smirnov Test Results (Balance-Sheet Leverage)

Balance-sheet leverage	Total	A	B	C	D	E	F	G	H
Mean	.703	.604	.609	.550	.603	.522	.628	.707	.671
Median	.711	.646	.696	.559	.625	.501	.668	.791	.797
Std. Deviation	16.268	.326	.354	.329	.364	.395	.495	.310	.357
Skewness	300.480	-.289	-.429	-.097	19.601	2.295	45.393	.901	.316
Kurtosis	90812.547	-.959	-1.337	-1.266	1547.769	30.079	3405.032	250.820	7.699
KS value	.138	.111	.147	.086	.137	.110	.225	.170	.174
p value	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**

Balance-sheet leverage	Total	I	J	K	L	M	N	O	P
Mean	.703	.647	.533	1.007	.540	.633	.471	.620	.654
Median	.711	.708	.540	.594	.572	.718	.409	.680	.668
Std. Deviation	16.268	.480	.392	45.510	.501	.340	.344	.422	.036
Skewness	300.480	50.793	.210	107.451	-.025	-.372	1.248	13.247	-1.564
Kurtosis	90812.547	4555.958	-.212	11606.610	-3.265	-1.009	9.117	516.369	2.358
KS value	.138	.229	.130	.499	.314	.139	.085	.181	.283
p value	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**

The results of the Kolmogorov-Smirnov test are also reported in the table. Since the assumptions of normal distribution are rejected and all 16 samples have very skewed distributions, the nonparametric method is more appropriate. We apply the Kruskal-Wallis non parametric analysis of variances to test the null hypothesis of no variation in leverage across industries during 2003-2012.

Table 4: Kruskal-Wallis Test Results (Balance-Sheet Leverage)

	Year	Balance-sheet leverage		Year	Balance-sheet leverage
2003	Chi-Square	650.593	**2008**	Chi-Square	462.131
	p value	.000		p value	.000
2004	Chi-Square	582.597	**2009**	Chi-Square	526.319
	p value	.000		p value	.000
2005	Chi-Square	621.606	**2010**	Chi-Square	509.282
	p value	.000		p value	.000
2006	Chi-Square	631.328	**2011**	Chi-Square	444.789
	p value	.000		p value	.000
2007	Chi-Square	594.264	**2012**	Chi-Square	417.158
	p value	.000		p value	.000

Again we found that capital structure positions among industries have significant differences according to statistical evidence using a non-parametric analysis of variance.

4 CONCLUSION

Using Kruskal-Wallis one-way analysis of variance, we present evidence that capital structures of firms vary across a wide array of industries. The results are obviously robust to alternative measure of the firm capital structure. The literature suggests that inter-industry differences in capital structure are related to differences in operating characteristics (including the nature of assets and technologies used) and the industry specific regulations. Our evidence signifies the important role that the industry specific technologies and regulations play in the firm's capital structure decisions. Since an industry dummy variable can be used to model the effect of differences across industries on the capital structure, the industry dummies, as explanatory variables, should be part of the regression models.

Taken altogether, the companies operating in the same industry are faced with the same economic conditions but the economic conditions may vary among industries. Consequently, industry classification can be used as a proxy for the business risk and the business risk will vary from industry to industry. As a consequence, the capital-structure norms will vary from industry to industry. A systematic variation in leverage among industries may be seen as evidence to reject the irrelevancy of capital structure. Therefore, it is worthwhile for the governments, policymakers and other stakeholders to craft the policy interventions with great prudence, meaning that a "one shoe fits all" type of policy intervention may not be effective for all the industries in a country.

The old question remains as to what extent does the "adherence" on the past information influence current capital structure decision making? Value indicators reflect the results of the past decisions and may not be predictive for future results. For many of the indicators in the Report, there are no ranges of value that are generally accepted to constitute a "good performance" or a "bad performance". There are no official publications or statistics, neither on the industry comparative ratios, nor on the benchmark companies.

Although there is no prior Study which attempted to investigate inter-industry variations in financing decisions of a firm within the context of BiH firms, we are mindful of the fact that this approach does not tell us *how* the industry-specific factors determine the firm financial structure, *nor why* the financial structures vary so widely across firms within a given industry. Future research should consider investigating how the industry factors such as the industry competition and concentration, technology and risk influence financial decisions of a firm. Such research would also proffer crucial information to governments and policymakers in their effort to craft industry-specific policy interventions.

References

[1] Bevan, A.A. & Danbolt, J. (2002). Capital structure and its determinants in the United Kingdom – a decompositional analysis, *Applied Financial Economics*, *12*(3), 159-170.

[2] Booth, A., Demirgüç-Kunt, A., & Maksimovic, V. (2001). Capital structures in developing countries. *Journal of Finance*, *56*, 87-129.

[3] Rajan, R., & Zingales, L. (1995). What do we know about capital structure? *Journal of Finance*, *50*(5), 1421-1460.

[4] Trivun, V., Silajdžić, V., Mahmutćehajić, F., Mrgud, M. (2009). Applied business law. School of Economics and Business, Sarajevo.

Appendix 1. Industry Classification

Table 5: Industry Classification

	Abbreviation	Industry
PRODUCTION	A	Agriculture, Hunting and Forestry
	B	Fishing
	C	Mining and Quarrying
	D	Manufacturing
	E	Production and Supply of Electricity, Gas and Water
CONSTRUCTION	F	Construction
SERVICES	G	Wholsesale and Retail Trade; Repair of Motor Vechicles and Personal and Households Goods
	H	Accomodation and Food Services Activities
	I	Transport, Storage and Supportc Activities for Transportations
	J	Financial Services
	K	Real Estate Services
	L	Public Administration and Defence; Compulsory Social Security
	M	Education
	N	Human Health and Social Work Activities
	O	Other Public Services
	P	Activities of Households as Employers of Domestic Personnel

Source: *2010 Classification of Activities in BiH,* Table T.2

SOCIAL ACCOUNTING MATRIX - METHODOLOGICAL BASIS FOR SUSTAINABLE DEVELOPMENT GOALS ANALYSIS IN THE WESTERN BALKANS COUNTRIES

Sasho Kjosev, PhD, Full Professor
University "Ss. Cyril and Methodius", Faculty of Economics
Blvd. Goce Delchev 9V, 1000 Skopje, Republic of Macedonia
E-mail: skosev@eccf.ukim.edu.mk

Blagica Novkovska, PhD, Assistant Professor
University of Tourism and Management, Faculty of Economics
Blvd. Partizanski Odredi No. 99, 1000 Skopje, Republic of Macedonia
E-mail: blagica@novkovski.com

Abstract: Social accounting matrix (SAM) technique has been shown to be exceptionally useful in providing the basis in-depth analyses of all the economic flows and connections between them. Highly developed methods have already applied in developed countries, where they have shown their potential for exhaustive analyses and monitoring of the realization of sustainable development goals. Established SAM in institutions producing official statistical data are of particular importance for the production of SAM organized data. In the case of Western Balkan countries such institutional capacities for production of SAM and their use have not been yet established. In this work we propose creating of a Social Accounting Matrix network of the State Statistical Offices in the Western Balkans region, for developing SAM, SESAME and NAMEA methodology adjusted to the Western Balkan countries specifics, needs and requirements. Expected benefits from this network are specified.

Keywords: Social Accounting Matrix, SESAME, NAMEA, Western Balkans, Sustainable development, Regional development

1 INTRODUCTION

Sustainable development goals analysis needs adequate quantitative analysis of micro and macro-economic policy, based on reliable data and relevant analytical tools. Social accounting matrices (SAMs) provide a coherent, detailed data base on major macroeconomic aggregates in the economy, which support macroeconomic and development policy making. SAM goes further than the improvement of statistics. Rather, it is a common ground of economic planners and development economists, on the one hand, and statisticians, on the other. In addition, the SAM approach has proved to be a practical quantitative tool of significant importance for achieving the best use of available data and in providing a quantitative basis for analysis.

2 LITERATURE REVIEW

The genesis of the Social Accounting Matrix (SAM) goes back to Richard Stone's pioneering work on social accounts. Subsequently, the genuine theoretical foundations for the use of Social Accounting Matrices has been laid down with the highly significant World Bank publication [13], followed by the joint work of EUROSTAT, IMF, OECD, UN and World Bank [3], the European Commission [4], as well as Erik Thorbecke (13). Applied economic analyses of social dimension of sustainable development goals are often seriously troubled by the lack of a complete data framework. The so-called SESAME is presented as a possible solution to problems in connection to SD indicators, where great support for developing and implementation of SESAME was given by Steven Keuning and his team at Statistic Netherlands [5] and [6]. On the other hand, the NAMEA has been developed to

systematically supplement the national accounts with environmental statistics. De Haan's PhD thesis [2] are used to present, very briefly, its basic characteristics. The authors, Kjosev and Novkovska [7], [8], [10] and [11], tried to incorporate the main theoretical findings in a "Western Balkan" context.

3 THEORETICAL CONSIDERATIONS

In order to better understand the concept of sustainable development goals analysis and monitoring, one has to develop an appropriate methodological instrument. In parallel with the introduction of the sustainable development concept, the SNA 1993 introduced (and SNA 2008 further developed) the Social Accounting Matrix (where later its extensions, the SESAME and NAMEA approaches, have been developed) as a methodological basis for the sustainable development goals analysis. Below, we briefly discuss the following elements of this methodology:
   - Social Accounting Matrix (SAM);
   - SAM extended with social indicators (SESAME approach); and
   - SAM extended with environmental indicators (NAMEA approach).

3.1 Social Accounting Matrix

Social accounting matrix (SAM) is a technique related to national income accounting, providing a conceptual basis for examining both growth and distributional issues within a single analytical framework in an economy. In SAM the flows of all economic transactions occurring in a regional or national economy are included. It can be seen as means of presenting in a single matrix the interaction between production, income, consumption and capital accumulation [2]. Therefore, SAM can be regarded as a square matrix (the incomings and outgoings for each account are shown as a corresponding row and column of the matrix) representation of National accounts.

The SAM is a comprehensive (it portrays all the economic activities of the system - consumption, production, accumulation and distribution), flexible, and disaggregated framework (there is a large measure of flexibility both in the degree of disaggregation and in the emphasis placed on different parts of the economic system) which elaborates and articulates the generation of income by activities of production and the distribution and redistribution of income between social and institutional groups. A principle objective of compiling a SAM is, therefore, to reflect various interdependencies in the socio-economic system as a whole by recording, as comprehensively as is practicable, the actual and imputed transactions and transfers between various agents in the system. Hence, the SAM can be visualized as a one-time snapshot of an economy incorporating the interdependence that exists within a socioeconomic system through a consistently organized complete data system.

The SAM has, mainly, two basic tasks:
   - to enable presentation of information about the economic and social structure of the regional and national economy; and
   - to provide analytical and accounting framework as a basis for construction of macroeconomic models for analyzing the regional and national economy and the effects from the implementation of the macroeconomic and development policy measures [3].

SAM can also be used for cross-country analyses and creation of international economic policies [1]. Namely, a social accounting matrix can also be constructed on a regional level, or in other words, the economy can be divided into separate sub-national economies or regions. In this model, blocks within a region in the economy may receive transfers from

various blocks in other regions, albeit still in the same country. Likewise, blocks within a region may also send payments to other blocks outside the region. Needles to say, the development of a regional social accounting matrix system may be approached from two standpoints. One approach is to disaggregate a SAM for the economy, taken as a whole, into its constituent regional components. The alternative approach is to combine SAMs for two (or more) regions into an integrated system.

According to Thorbecke [12], "…distinguishing regions within a country SAM can enhance both its realism and its usefulness. If the economy displays significant regional differences in the types of goods produced, structure of production and technology, these differences could affect the standards of living of different household groups". Thorbecke also mentions the fact that a large number of policy means (e.g. investment projects, current government expenditures on health or education) are location-specific as another advantage of having a regional dimension in SAM.

3.2 SESAME approach

SESAME (System of Economic and Social Accounting matrices and Extensions) is a statistical information system in matrix format, from which a set of core economic, environmental and social macro-indicators is derived. The system is driven, to a large extent, by the kind of information required for monitoring and policy-making at the macro level. Although it is impossible to capture socio-economic development in a single indicator, it is equally clear that a prime task of national statistical offices is to comprise the countless numbers they collect to a manageable, "executive" summary. Such a summary typically describes trends in main indicators. At the same time, for analytical purposes a more detailed data framework is required. Obviously, the communication between policy-makers and analysts is optimally served if the core macro-indicators are all derived from an integrated information system such as SESAME [5].

Keuning [6] points out to the following advantages of the SESAME approach:
- Just like conventional national accounts, *SESAME provides both core macro-indicators and an underlying information* system;
- SESAME promotes the use of uniform units, classifications, concepts, etc. throughout a statistical system; that is, not only in economic statistics, but also in social statistics. Among the advantages of such a harmonization is a much easier matching of results from different surveys;
- SESAME is an inherently *flexible* framework. It can readily be adapted to the specific characteristics, needs and capabilities of every country or region; and
- Finally, it should be mentioned that *SESAME essentially aims at a better use, through integration, of existing statistics.*

3.3 NAMEA approach

The NAMEA has been developed to systematically supplement the national accounts with environmental statistics. Its hybrid accounting structure, i.e. the combined presentation of physical and monetary accounts, indicates that in the NAMEA environmental imputations in the core national accounts framework are avoided.

Therefore, NAMEA has been developed to link environmental and economic statistics. An important characteristic of environmental accounting is that the data are consistent with the National Accounts which mean that the environmental data can be directly compared to well known macro-economic indicators such as GDP, inflation and investment rates, developed in the System of National Accounts (SNA) [7].

De Haan [2], points out to the following main characteristics of the NAMEA:
- The NAMEA maintains a strict borderline between the economic sphere and the natural environment, established by monetary accounts on the one hand and accounts denominated in the most relevant physical units on the other;
- The NAMEA maintains a clear distinction between physical inputs (extraction of resources) on the one hand and outputs (emission of pollutants) on the other, and
- Finally, the NAMEA provides an institutional representation of the economy and its relationship with the environment.

4 WHAT SHOULD BE DONE IN THE WESTERN BALKANS?

The methodology for sustainable development goals analysis is of highest importance for the unity, complexity and consistency of the sustainable development planning system. It should enable methodological consistency in the process of evaluation of the development conditions, problems and perspectives, perception of interests, objectives and tasks of the relevant stakeholders and their harmonization, the simultaneity of the planning process, as well as the mandatory preparation and execution of plans.

Particular attention in the establishing of SAM is to be devoted to the differences between the regions in the countries [10], since there are multiple important issues connected with regional disparities to be studied based on the data from SAM [11].

Having in mind the wide variety of socio-economic problems and future challenges in the Western Balkans, we are free to suggest the following activities in order to implement this methodology in our region:
- Creating a Social Accounting Matrix network of the State Statistical Offices in the Western Balkans region, for developing SAM, SESAME and NAMEA methodology adjusted to the Western Balkan countries specifics, needs and requirements;
- Creating a coordination institution and mechanism;
- Activities for institutional capacity building required for building and implementing the SAM methodology
- Organize and implement adequate professional trainings in order to overcome the lack of skilled specialists, capable of working on/with Social Accounting Matrices
- Creating highly professional working groups in each of the Western Balkan countries
- As a result of all the above mentioned, building a Western Balkan countries SAM that will be able to analyze the regional impact of the national macroeconomic and development policies across a wide range of economic, social and environmental indicators.

5 WHY DO WE NEED SOCIAL ACCOOUNTING MATRIX METHODOLOGY IN THE WESTERN BALKANS?

First, the practice of developed countries shows multiple benefits from established SAM systems. Thus, sophisticated analysis comparing theory and empirical data using official NAMEA for Italy, produced by national statistical institution (ISTAT), has been used in [9] to study the link between environmental efficiency and labour productivity.

Second, the statistical systems of the Western Balkan countries are compatible to a large extent; they are using mainly the same methodologies and already have established significant links for cooperation. This is an asset promising successful implementation of the system of SAMs that will be mutually comparable and susceptible to cross-linked analyses.

Compilation, analysis and use of SAMs in setting up projection tools allows to penetrate into all sophisticated socio-economic processes, while regulating such processes from the perspective of principles of economics. Such a great analytical capacity can play a great role not only in terms of governing problems of economic regulation within one country, but also managing inter-country economic cooperation issues at a level of a number of countries. In this context, it is important for the Western Balkan countries to compile SAM and use it for creation and implementation of the respective countries macroeconomic and strategic policy documents and analysis. Using such tools in the Western Balkan countries might play a great role in defining areas of cooperation and raising efficiency of mutual benefits.

Hence, establishing and implementing sustainable development goals planning methodology is of highest importance for the organization of efficient development planning systems in the Western Balkan countries. It should support methodological consistency in the evaluation process of the development conditions, problems and perspectives, the simultaneity of the planning process, as well as the mandatory preparation and execution of strategic planning documents [8].

The institutions in the respective Western Balkan countries can use this Social Accounting Matrix methodology for:
- preparation of macroeconomic analyses for the economic, social and regional development;
- preparation of medium and long-term strategic sustainable development planning documents;
- preparation of studies and strategies for the national and regional/local economic development in the region;
- preparation of analyses for the situation and problems of the socio-economic and environmental development in the region;
- planning and implementation of the Governments macroeconomic and development policies and strategies; and
- development and implementation of an integrated approach to policymaking on implementing sustainable development goals in the Western Balkans.

In addition, implementing the Social Accounting Matrix will create number of additional benefits for the Western Balkans national economies:
- build dynamic and innovative economies that provide prosperity for their citizens;
- construction of tolerant, inclusive and stable societies that provide improved quality of life for their citizens;
- provide an appropriate balance between responsible use and preservation of natural resources in support of a better quality of life and quality of environment; and
- construction and application of a mechanism for macroeconomic and development policy that support the application of the sustainable development concept in the Western Balkans national economies.

6 CONCLUSIONS

Governments of the Western Balkans countries are responsible for the preparation and realization of efficient macroeconomic and development policies. Such policies are the basis for the implementation of the sustainable development planning system in the national respective national economies. Moreover, the macroeconomic policy implementation is looking for utilization of modern planning and forecasting techniques, as well as a developed information system as a basis for an efficient macroeconomic and development policy. This will lead to a continuous improvement of the economic policy instruments, as well as the other types of planning and programming of the respective national economies development.

Having this in mind, one can support the fact that developing and implementing Social Accounting Matrices for the Western Balkan countries will provide a solid basis for implementation of an integrated, economy-wide planning and modelling methodology that captures all segments and sectors of the national economies, within one integrated comprehensive framework.

The Social Accounting Matrix for the separate Western Balkan countries will be used for elaborating a more realistic and empirically based development and macroeconomic policy and strategic vision for the respective economies.

Acknowledgement

References
[1] Ashimov, A., Borovskiy, Y., Borovskiy, N., Adilov, Z., Alshanov, R. and Sultanov, B. 2014. Evaluation of optimal international economic policy based on both the parametric control theory and global computable general equilibrium model. Procedia Computer Science, 31: 701–710.

[2] de Haan, Mark. 2004. Accounting for Goods and for Bads: Measuring Environmental Pressure in a National Accounts Framework (PhD thesis). Voorburg, the Netherlands. Statistics Netherlands

[3] EUROSTAT, IMF, OECD, UN and World Bank. 1993. System of National Accounts 1993. Brussels/Luxembourg, New York, Paris & Washington D.C.

[4] Leadership Group SAM. 2003. Handbook on Social Accounting Matrices and Labour Accounts. Population and social conditions 3/2003/E/N°23. Luxembourg: European Commission

[5] Keuning, Steven. 1998. Interaction between national accounts and socio-economic policy. Review of Income and Wealth, 44(3): 345–359.

[6] Keuning, Steven. 2000. Accounting for welfare in SESAME. In: Household Accounting – Experience in Concepts and Compilation, Handbook of National Accounting,Series F,75(2): 273–307. New York: United Nations.

[7] Kjosev, Sasho. 2012. Social Accounting Matrix – Methodological Basis for Sustainable Development Analysis. In Chaouki Ghenai (Ed.). Sustainable Development - Policy and Urban Development - Tourism, Life Science, Management and Environment (pp. 269–284). Rijeka, Croatia: InTech.

[8] Kjosev, Sasho, Gockov, Gjorgji and Ljupcho Eftimov. 2014. Why is the Social Accounting Matrix Important for the Republic of Macedonia. The Young Economist Journal, 23: 45-50.

[9] Mazzanti, M. and Zoboli, R. 2009. Environmental efficiency and labour productivity: Trade-off or joint dynamics? A theoretical investigation and empirical evidence from Italy using NAMEA. Ecological Economics, 68(4): 1182–1194.

[10] Novkovska, Blagica. 2016. Role of the Clusters in Reduction of Regional Development Disparities in Macedonia. VII Balkan & Black Sea Conference DAYS OF CLUSTERS 2016, Ohrid, Macedonia.

[11] Novkovska, Blagica. 2017. Regional development disparities and their connection with hidden economy. UTMS Journal of Economics *in press*.

[12] Thorbecke, Erik. 2000. The Use of Social Accounting Matrices in Modelling. Paper Prepared for the 26th General Conference of The International Association for Research in Income and Wealth. Cracow, Poland (http://www.iariw.org/papers/2000/thorbecke.pdf)

[13] World Bank. 1985. Social Accounting Matrices: A Basis for Planning (eds. Graham Pyatt and Jeffery I. Round). ISBN 0-8213-0550-6. Washington, DC

WHAT AFFECTS THE EXPORT PERFORMANCE OF CROATIA IN EASTERN EUROPE?

Helena Nikolic
University of Zagreb, Faculty of Economics & Business, Department of Trade
J.F. Kennedy 6, 10 000 Zagreb, Croatia
E-mail: hmiloloza@efzg.hr

Abstract: The paper examines and measures the direction and strength of impacts of obstacles to the export activities of Croatia to the countries of Eastern Europe, especially the insufficiently explored determinants - Market orientation, innovation of products and services and organizational capacity for export activities. Research findings have shown that cultural and geographical differences between countries have no impact on export performance, that market orientation has a positive impact on the organizational capacity of enterprises for export activities and on the innovation of products and services intended for export to the selected market, that organizational capacities, in some segments, have a positive impact on the export performance of Croatian companies in the Eastern European market, while innovations in export products and services have a negative impact.

Keywords: export, determinants of export activities, Eastern Europe

1 INTRODUCTION

Export provides an economical way of rapid sales to new markets [13, 15]. The company's technological readiness, high quality management and high value-added products at competitive prices can ensure long-term growth on single market. The country's export orientation also increases the overall competitiveness of the country. Dominant and growing exports to Eastern European countries have significantly contributed to economic growth of Croatia. Namely, since 2000, when the presence of Eastern European countries in Croatian exports was 26%, there was a constant significant increase in the share of exports to the Eastern European market. According to the latest data, in 2015, this share has reached an enviable 42% [6] and the upward trend is expected in the future.

The complexity of the determinants of export activity is reflected in the fact that they all include factors affecting export activity; from the basic characteristics of the enterprise itself to the characteristics of supreme management that manages the enterprise. Three determinants of export activities - market orientation, organizational capacity for export activities and innovation of products and services - are set at the centre of this paper research.

The philosophical basis of market orientation is the concept of marketing based on its basic settings and foundations. It is consisted of a series of activities related to the application of the marketing concept in business [11]. Numerous factors influence design of market orientation. Some of them are environment, organizational system and employees, the interurban dynamics and top management that defines the company's performance on the market, which is ultimately reflected on export result [9, 16].

Organizational capacities are a key resource of a knowledge-based economy. Employees possess knowledge, skills, competences, experience, creativity and creative ability and are essential for growth and business development [19]. Through their training and raising their knowledge, skills and competencies, organizational capacity is created to create a new value that is confirmed by the market and which ensures long-term competitiveness [1, 2, 18]. What constitutes a problem with organizational capacities, when export is concerned, is the fact that in a company, especially the smaller one, there is only one person in charge of a wide range of business decisions. Management is the one that plays a key role in the choice of export activity plan and program, but in the absence of time, due to the burden on domestic market, export is put in another plan [14]. For this reason it is necessary to provide

qualified organizational capacities to exploit the opportunities and benefits that foreign markets provide.

Innovation represents the implemented new technological product and manufacturing process and significant technological improvement of the existing product and process [17]. In modern economic trends, the ability to innovate continually is one of the basic distinct criteria between successful and unsuccessful companies. Namely, creating new products, services and processes is a complex work with lot uncertainty and risk, but competitiveness in the international market is impossible without developing innovation.

Research into the impact of market orientation on organizational capacities is virtually unavailable in scientific literature. Several studies have been carried out on the impact of market orientation on organizational commitment and organizational learning, but the obtained results indicated the existence of empirical inconsistencies. A number of researchers have come to the conclusion that organizational orientation is a consequence of market orientation [5, 9], while others argue that organizational orientation prevails over market orientation [20]. Adopting market orientation contributes to the improvement of morale, job satisfaction and commitment because all organizational units are geared toward a common goal - external customer satisfaction [11]. Given the greater customer satisfaction, the business results are better and the company has a greater need to expand its business, resulting in the need for recruiting new workforce, ie the need for organizational capacities to be solely responsible for export activities.

Market orientation certainly has an impact on innovation processes in the enterprise. However, the results of recent research into correlation between market orientation and innovation have shown that the effect of market orientation can be both positive and negative.

The aims of this paper are: (i) empirically and statistically determine the impact of export activity obstacles to Eastern European countries on the export performance of Croatian companies in the mentioned market, (ii) empirically and statistically determine the level of market orientation, organizational capacity for export activities and innovation of products and services intended for export to the export performance of Croatian companies in the market of Eastern European countries.

The paper is organised as follows. After the brief introduction, the second chapter explains variables that were used in the research. Third chapter is consisted of methodology. In the fourth chapter results have been shown and the fifth chapter concludes the paper and provides recommendations for further research.

2 DATA DEFINITON AND DESCRIPTION

The CAGE framework encompasses determinants (cultural, administrative, geographical and economic diversity) that influence the decision of an enterprise to which export markets will focus their business and how to shape their own export strategy [7]. The initial assumption was that barriers to export activities, which are set within the CAGE framework, are less in the Eastern European countries than in other markets. In other words, the countries of Eastern Europe are most similar to the Croatian market according to all four criteria. For this reason, Croatian companies are most oriented toward the mentioned market and realize more remarkable export results. The hypothesis set was: H1: *Less export barriers, measured by the CAGE framework, in Eastern Europe compared to other markets, have a positive impact on the export performance of Croatian companies in the Eastern European market.*

Furthermore, the assumption was that market-oriented companies, on one hand, have more developed organizational capacities for export and, and on the other, make more efforts to innovate products and services intended for export markets. Two hypotheses have been

defined from the above assumptions: H2: *The market orientation of Croatian companies positively affects the organizational capacity of enterprises for export activities*. H3: *The market orientation of Croatian companies has a positive impact on the innovation of products and services intended for export to the said market.*

Finally, the aim was to examine the stability of the positive relationship between the development of organizational capacity for export activities and the persistence of the positive relationship between product innovation and export-oriented services and the export performance of companies in the Eastern European market. Hence, the hypotheses were: H4: *The organizational capacity of the company for export activities has a positive impact on the export performance of Croatian companies in the market of Eastern European countries*. H5: *Innovation of exported products and services has a positive impact on the export performance of Croatian companies in the Eastern European market.*

A proposed research model of structural equations is shown below (see Fig. 1)
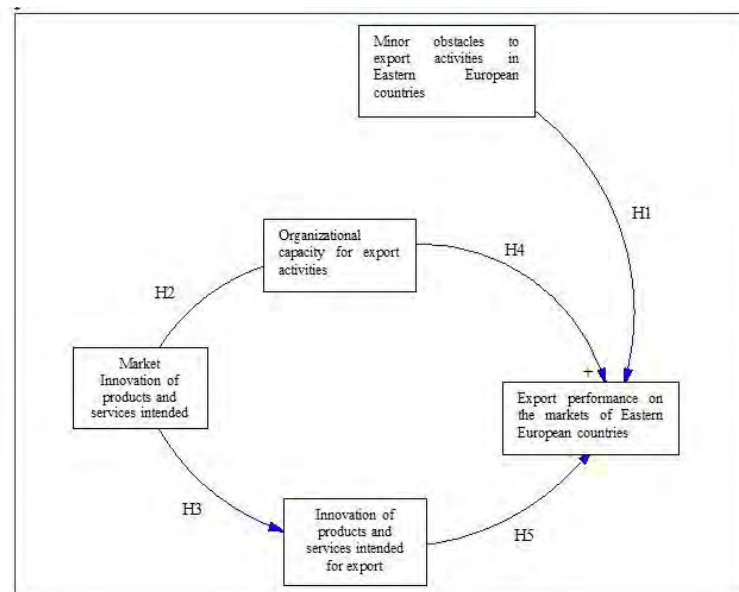
Figure 1: A proposed research model of structural equations

Dependent variable, export performance, was observed on the market of Eastern Europe which is comprised of 20 countries: Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Montenegro, Czech Republic, Croatia, Georgia, Kosovo, Macedonia, Hungary, Moldova, Poland and Romania, European part of Russia, Serbia, Slovenia, Slovakia and Ukraine. The stated classification is according to the European Union Glossary. The main variable represented the share of exports to the market of Eastern European countries in total revenue (%). Structural model variables – Obstacles to export activities in Eastern European countries, Market orientation of enterprises, Organizational capacity for export activities and Innovation of products and services intended for export – were measured by using known metering methods that are available and used in scientific literature (see Tab. 1)

Table 1: Operationalization of model variables

Construct	The definition of the construct
Obstacles to export activities in Eastern European countries	The CAGE framework covers the perceptions of respondents with regard to cultural, administrative, geographical and economic differences between Croatia and the countries of Eastern Europe. The scale consists of a total of 17 statements and is measured by matching the respondents with statements using the Likert scale (1-7) [4].

Market orientation of enterprises	Market orientation is a one-dimensional construct that consists of three behavioural components: consumer orientation, competition orientation, and inter-functional coordination. The MKTOR scale consists of a total of 15 statements. Compatibility of respondents with statements is measured using Likert scale (1-7) [11,3].
Organizational capacity for export activities	The scale consists of five subdivisions: management support, autonomy in decision making, emphasis on the initiative, timeliness and organizational arrangement. The scale consists of a total of 15 selected arguments. Compatibility of respondents with statements is measured using Likert scale (1-7) [12].
Innovation of products and services intended for export	The scale measures enterprise innovations, as well as actively supporting innovation in the enterprise, whereby products and services are viewed. Innovation of products and services must be new to the enterprise, but not necessarily for the market. The scale consists of 10 statements and the consistency of the respondents with statements using the Likert scale (1-7) is measured [8,10].

3 METHODOLOGY

By adopting a positivist approach, based on the study of previous research from relevant literature sources, five hypotheses were tested by using the path of regression models. The research was conducted in two phases. In first phase a pilot survey was carried out, whereby a statistical questionnaire was tested. Based on the observed shortcomings, an adaptation of the questionnaire was adjusted, which was ultimately distributed to a larger number of respondents. A reliability analysis of the scale was performed by using the Cronbach alpha coefficient. A survey of collected data was then carried out; including tests aimed at testing non-hypothetical values in the data as well as testing assumptions about the normal distribution of manifest variables. Finally, data analysis was performed by modelling structural equations.

The target population of the survey included Croatian export companies (regardless of the size of the company). These are companies that, according to the Customs Administration of the Republic of Croatia, realized the export of goods and services. Given the uneven distribution of the company when the size was concerned, a stratified random systematic pattern was applied in the research. The list of companies was created using the Croatian Exporters Register of the Croatian Chamber of Economy. A total of 200 small, medium and 200 large companies were selected in the sample. Of these, the study involved 30 small companies (response rate 15%), 38 medium (response rate 19%) and 62 large Croatian companies (response rate 31%).

Companies were selected in a random systematic pattern using the method of step [k = N (population)/n (sample)], with a randomly selected beginning determined by a number between 1 and k that is determined by random number table. As a result, companies were ranked according to the alphabet in order to avoid periodic repetitions in the sample related to the activity, region or some other characteristic that could affect the sample representativeness.

The researcher was the director of a company, a board member or a person in charge of international business operations. The survey was conducted using an online questionnaire. Afterwards, control of the responses collected in the direction of the structure of the organizations that participated in the research as well as in the direction of completing the survey research instrument was carried out. The information that legitimizes the surveys, such as the Timestamp and the ID number of each questionnaire, were also checked.

4 RESULTS

The survey was attended by the largest companies in the manufacturing industry. The largest number of companies is predominantly privately owned (87), while the smallest number of enterprises is predominantly state-owned (4). They are mostly independent companies (111), while there are fewer subsidiaries of some multinational companies (19). Respondents match the profile of managers in the Republic of Croatia. Most of the respondents are male, aged between 31 and 40 years, with related degrees.

The share of exports to Eastern European countries in the total enterprise income is 15.05% (see Tab. 2). Large companies are more export oriented and the share of exports in total revenue is higher. However, 68 small and medium-sized enterprises (52.31%) participated in the survey. To them, export represent a subsidiary activity and they carefully perform in foreign markets; most often, in the role of passive or reactive participants. Their income from abroad is lower, as well as their share in total earnings.

Table 2: Descriptive statistics the share of exports to the countries of Eastern Europe in total revenue companies

	N	Min	Max	Average	Std. Dev.
The share of exports to Eastern European countries in the total revenue of the enterprise	130	0,2	99	15,05	24,47

Source: Copyright Research, February-May, 2015.

Respondents as the biggest problem in export activities see the difference in legal systems (corruption, weak regulation, etc.) of Eastern European countries. On the other hand, differences in language are least expressed among Croatian companies and Eastern European countries, and according to them, they are least avoided when exporting to the Eastern European market.

Within the dimension of Cultural Differences, Different Value Systems have the highest average rating, and within the dimension of Administrative-Political Differences it is the Legal System. By observing the dimension of Geographic Differences, the variables of the geographic distance have been shown as the dimension with the highest average rating, and within the dimension of the Economic Differences is the Restricted Infrastructure.

Respondents evaluate all the dimensions of market orientation almost equally. There is no big difference between consumer orientation, competition orientation, and inter-functional coordination. Companies pay the most attention to the market by understanding and meeting the needs of their target consumers to create value for their business through their business. However, the biggest problem is in sales-staff who does not regularly exchange information about competitors.

Superiors are aware of the importance of employees who are ready to accept new challenges. Therefore, employees, in line with the results achieved, enjoy greater autonomy and are more responsible for their responsibilities. However, their activity is still, mostly, under the supervision and control of their superiors. Also, deprivation of freedom of decision was noted. Free decision making is not the practice of doing business in Croatian companies, and the problem lies in the lack of self-initiative of employees who have no interest in generating new ideas and functions in the enterprise. Employees often act as passive observers who routinely and reluctantly execute set tasks, which create mistrust of their superiors.

When innovations are concerned, small firms in the international market can hardly be in the role of market leader. There is a lack of human resources that will devote their time to innovations and are faced with a lack of financial and material resources. Also, they are limited to innovation in the business they are dealing with. Small businesses, however,

endeavour to implement innovations, at least in the segment of services that do not require extensive and cost-effective examinations. Medium-sized companies are not prone to innovation, and when trying out new ideas, they are very cautious, but they are trying to promote at least innovative high-risk services. For large and state-owned enterprises, there has been a marked increase in product and service innovation over the past five years. Big companies are trying to become market leaders, innovators in their activities, and to gain new markets thanks to their innovations. Private companies constantly introduce and test new services but have no primacy in launching new products and / or services on the market.

By testing hypothesis H1, it has been shown that the share of exports to Eastern European countries in the total income of enterprises has statistically significant negative impacts on political-administrative differences with probability of 1% and statistically significant positive influence on economic diversity with probability of 5%. It has been shown that the share of exports to Eastern European countries in the total income of the company has a statistically significant negative impact on political and social conflicts with probability of 1% and statistically significant positive influence on the differences in the economic power of the state with probability of 1%.

By testing hypothesis H2, it has been shown that the organizational capacity of the company for export activities has a statistically significant positive impact on the orientation towards competition with a probability of 1%. It has been shown that the organizational capacity of the company for export activities has a statistically significant positive effect on the exchange of information on competitors between sales staff, the focus on those consumers in which companies can gain competitive advantage and the development and implementation of post-sales services with a probability of 1%.

By testing hypothesis H3, It has been shown that the innovativeness of the company has a statistically significant positive impact on the orientation of the competition with a probability of 1%. It has been shown that company innovations have a statistically significant positive impact on the development and implementation of post-sales services, focus on those consumers in which companies can gain competitive edge and management awareness of the fact that any business activity can contribute to creating value for consumers.

By testing hypothesis H4, it has been shown that the share of exports to Eastern European countries in the company's total revenue has a statistically significant negative effect on management support in making risky decisions with a probability of 5% and statistically significantly positively influenced organizational organization of the company with probability of 1%. Employee's risk-taking as well as self-empowering employees in higher positions on new jobs and projects has a negative impact on the export performance of companies in the Eastern European market. On the other hand, it has been shown that the share of exports to Eastern European countries in the total income of the company has a statistically significant positive effect on providing employees with opportunities to be creative and tempt their own methods of doing business tasks, as well as the existence of standard procedures and compliance with customary business practices.

By testing hypothesis H5, there was no statistically significant influence. It has been shown that the share of exports to Eastern European countries in the company's total revenue has a statistically significant negative impact on the promotion of new services with a probability of 5%. However, such a result can be explained by the fact that the markets of the Eastern European countries are economically less developed (in line with the CAGE framework) and are therefore oriented towards products that do not represent innovations for our companies but are relatively new in the markets of Eastern European countries interesting.

5 CONCLUSION

This paper has emerged from the observed problem of the importance of integrating small markets into world economic trends, the ubiquitous trends of globalization and internationalization, the important role of exporters as the backbone of the national economy, and the constantly present negative trade balance, questionable competitiveness and sustainability of the Croatian economy. Croatian companies are generally, irrespective of size, at a lower level of internationalization and are still failing to achieve remarkable results in international marketplaces dominated by better competitors.

Path analysis has demonstrated the positive impact of the company's market orientation on the organizational capacity for export and innovation activities of the company. However, it has been shown that organizational capacity for export activities in some segments has a negative impact (management support in making risky decisions), and in some segments a positive impact (standard procedures management and encouragement of creativity and employee efficiency) on the export performance of Croatian companies on the market Eastern Europe. It has also been shown that the innovation activities of enterprises have a negative impact on export activities in Eastern European countries. At first glance, this result is surprising. However, such a result can be explained by the fact that the markets of the Eastern European countries are economically less developed (in line with the CAGE framework) and are therefore oriented towards products that do not represent innovations for our companies but are relatively new in the markets of Eastern European countries interesting.

The main limitation of this work is the data characteristic that represents a one-time recording of specific data, which does not give insight into long-lasting connectivity and change. However, such research is largely used in social sciences, due to the high costs of longitudinal studies.

Recommendations and possibilities for further research include: (i) analysis of data independently of the export market or for a larger number of countries, (ii) a more detailed analysis of each determinant which has proved to be an important item of export performance (institutional framework, economic diversity, etc.), (iii) change in the number and distribution of independent variables in the model depending on the intentions and goals of future research. Since the administrative-political differences between two countries have proved to have the most negative impacts on export activities, it would be useful to study the policies and legislative frameworks of each of the countries of Eastern Europe, to study the steps taken by Croatian companies to address these problems, observe the time period of the achieved results, as well as progress with respect to the invested funds. A good insight into the problem situation could be achieved by conducting deep interviews with the surveyed companies.

References

[1] Alvarez, R. (2007). Explaining export success: firm characteristics and spillover effects. World development, 35(3), 377-393.

[2] Barney, J. B., Clark, D. N. (2007). Resource-based theory: Creating and sustaining competitive advantage, Oxford University Press, Oxford.

[3] Bozic, L., i Rajh, E. (2008). Procjena psihometrijskih karakteristika ljestvice za mjerenje tržišne orijentacije. Ekonomski pregled, 59(1-2), 38-50.

[4] Bruner, Conroy & Snell (2012). The Development of General Managers Capabilities in a Global Economy, In. Canals, J. (Eds.) Leadership Development in a Global World: The Role of Companies and Business Schools, Palgrave Macmillan, New York.

[5] Chang, W., Lu, L., Su, H., Lin, T., Chang, K. (2010). The mediating eff ect of role stressors on market orientation and organizational commitment. Social Behaviour & Personality: An International Journal, 38(10), 1431-1440.

[6] Croatian Bureau of Statistics, PC- Axis Databases, Foreign trade in goods, http://www.dzs.hr/ [Accessed 20/09/16]

[7] Ghemawat, P. (2011). World 3.0: Global Prosperity and How to Achieve it, Harvard Business Review Press, Boston, Massachusetts.

[8] Jambulingam, T., Kathuria, R. &Doucette, W. R. (2005). Entrepreneurial orientation as a basis for classification within a service industry: the case of retail pharmacy industry. Journal of Operations Management, 23(1), 23-42;

[9] Jaworski, B. J., & Kohli, A. K. (1993). Market orientation: antecedents and consequences. The Journal of marketing, 53-70

[10] Keskin, H. (2006). Market orientation, learning orientation, and innovation capabilities in SMEs: An extended model. European Journal of Innovation Management, 9(4), 396-417.

[11] Kholi, A., Jaworski, B. (1990). Market orientation: The construct, research propositions, and managerial implications. Journal of Marketing, 54(2), 1-18.

[12] Kuratko, D. F., Hornsby, J. S., & Covin, J. G. (2014). Diagnosing a firm's internal environment for corporate entrepreneurship. Business Horizons, 57(1), 37-47.

[13] Leonidou L. C. (1995) Export barriers: non-exporters' perceptions, International Marketing Review, 12(1),4-25.

[14] Leonidou, L. C., Katsikeas, C. S., Piercy, N. F. (1998). Identifying managerial influences on exporting: Past research and future directions. Journal of International Marketing, 6(2), 74-102.

[15] Leonidou L. C. (2004), An analysis of the barriers hindering small business export development, Emerald Management Reviews, Journal of Small Business Management, 42(3), 279-302.

[16] Narver, J. C., & Slater, S. F. (1990). The effect of a market orientation on business profitability. The Journal of Marketing, 20-35.

[17] OECD, (1997). Proposed guidelines for collecting and interpreting technological innovations data: Oslo Manual, OECD, Paris.

[18] Stoian, M. C., Rialp, A., Rialp, J. (2011). Export performance under the microscope: A glance through Spanish lenses. International Business Review, 20(2), 117-135.

[19] Sveiby, K.E., Simons, R. (2002). Collaborative climate and effectiveness of knowledge work: an empirical study, Journal of Knowledge Management, 6(5), 420-429.

[20] Zhang, D., Sivaramakrishnan, S., Delbaere, M., Bruning, E. (2008). The relationship between organizational commitment and market orientation. Journal of Strategic Marketing, 16(1), 55-73.

MODELLING OF TEMPORAL PATTERNS OF HIDDEN ECONOMY IN CONNECTION WITH ENERGY CONSUMPTION

Blagica Novkovska
University of Tourism and Management in Skopje
Blvd. Partizanski odredi 99, 1000 Skopje, Macedonia
E-mail: blagica@novkovski.com

Ksenija Dumicic
Department of Statistics, Faculty of Economics and Business, University of Zagreb,
Trg J.F. Kennedy 6, HR-10000 Zagreb, Croatia
E-mail: kdumicic@efzg.hr

Abstract: In this work methods of estimation of the size of hidden economy are discussed. Particular attention is devoted to the methods using electricity consumption as indicator (the Lackó method and the Kaufmann and Kaliberda method). It has been shown that such methods can be used for determination of the size of hidden economy in small open economies exposed to severe external influences. Results for Macedonia and some comparison with results for Croatia, as good role-model for other states in Western Balkans, are shown.

Keywords: hidden economy, the DYMIMIC method, the Lacko method, the Kaufmann and Kaliberda method, currency demand method, energy consumption elasticity.

1 INTRODUCTION

Hidden economy is particularly significant phenomenon for the modern society. It is important both from the side of its effect on the total economy and from the side of its characterization and analysis [11].

Hidden economy is generally categorized from the point of view of tax-paying and economic reporting to official institutions. Therefore, this is the part that is not visible to official producers of macro-economic data. Main measure of the size of the hidden economy is corresponding Gross Domestic Product (GDP) of hidden economy sector, later in this work described as HY. It is usually considered the total Gross Domestic Product (*TY*) to be composed of the part of regular Gross Domestic Product (*Y*) and that of hidden economy (*HY*), i.e.

$$TY = Y + HY . \tag{1}$$

From the point of view of its characterization, hidden economy presents a big challenge for researchers, since it is not directly measurable quantity. Therefore, various assumptions are to be made before constructing the model for determination of the size of hidden economy, as a measure of its presence and influence on economic activities of the countries.

For small open economies, as are those of Western Balkans, the issue of determination of the size of hidden economy become more complex, since they are exposed to severe external shocks strongly affecting their economy [3].

The methods of determination of the hidden economy size are recognized as direct and indirect methods, as given in [18] and [17]. In this paper, some of the most often used indirect methods are listed and focused. The methods using electricity consumption as an indicator (the Lackó method and the Kaufmann and Kaliberda method) are considered separately. After the Tanzi method, introduced in 80-ties of the last century, as described in [14] and [16], the DYMIMIC method, presented in [12], with an application, as given in [12], is shown. The Tanzi method comparison with others, later introduced, methods is given in [15]. Further, the Lackó method [6] and the Kaufman and Kaliberda method [5] are focused. Finally, an

application to hidden economy for two countries, Macedonia and Croatia over the period from 1990 to 2004, follows.

2   DETERMINATION OF THE SIZE OF HIDDEN ECONOMY

Methods of determination of the size of hidden economy are divided in two groups: direct and indirect methods. Between the direct methods, the method based on statistical surveys is quite often used [18], [17]. In this method activities constituting the sector of hidden economy are immediately targeted, but the deficiencies of the method are significant. Namely, due to the nature of the economic activity studied, only partial reporting is expected and the size of the hidden economy estimated by these methods is substantially underestimated.

Therefore, particular attention is to be paid to indirect methods that do not rely on the reporting by specific subjects (companies or individuals), but on the effects of the hidden economy on the total economy of the country.

2.1   The Tanzi method

First between the indirect methods to be mentioned is the currency demand approach (the Tanzi method) introduced in 1980s [14], [16]. Later, this method was compared with other newest methods [15].

In this method the currency demand is considered to be increased due to the presence of hidden economy, where financial transactions are expected to be mostly done in cash. Therefore, the excess cash flow is to be considered as a measure of the size of hidden economy. The main advantage of this method is that the observed variable represents a financial quantity, and hence the measure for the size of hidden economy is obtained straightforward in units of the national currency. Many authors also nowadays use this method [1].

However, there is a substantial deficiency of this method relying in the fact that currency demand is a complex phenomenon involving multiple factors [13]. Detailed studies on the money demand [10] and monetary transmission mechanism in Croatia [2] support the above finding for the case of Western Balkans.

2.2   The DYMIMIC method

Second, the dynamic multiple indicators multiple causes method (called the DYMIMIC method) is to be regarded [12]. This method is rather complex, involving several causes (direct and indirect taxation, state regulation burden, unemployment and GDP) and indicators of the presence and extent of hidden economy (employment, GDP growth and currency change). Delay between the causes and the effect (hidden economy) is taken into account (see Fig. 1).

We have previously shown that the above method can be efficiently used to determine the size and the variations of the hidden economy in a small open economy in the case of Macedonia [8].

This method is complex and requires powerful analytical tools for its use (hidden variables and delay between the cause and the effect). However, there are some deficiencies that have to be considered when using it, requiring some precautions. First, even if it is rather complex, there are no strong proofs that it is exhaustive, meaning that some important factors can be neglected while using it. Second, even if the delay between the causes and the main consequence is reasonable, its size for 1 year is somehow arbitrarily chosen.
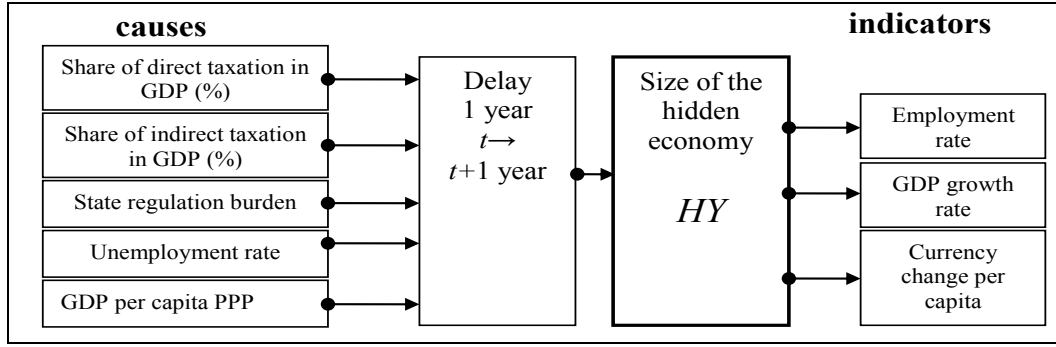
Figure 1: Diagram of connections between the causes and indicators in the DYMIMIC model

2.3 The Lackó method

Particular attention in this work is devoted to the methods involving energy consumption. In this case, the quantities (indicators) used in determination of the hidden economy are obtained by precise measurements of a real physical quantity.

First we discuss the method where household electricity consumption is used as a main indicator. The household electricity approach, or the Lackó method, [6], in a cross-country analysis is described by two simultaneous equations:

$$\ln E_i = a_1 \ln C_i + a_2 \ln PR_i + a_3 G_i + a_4 Q_i + a_5 HY_i + a_6 + u_i, \qquad (2)$$

with coefficients $a_1 > 0$, $a_2 < 0$, $a_3 > 0$, $a_4 < 0$ and $a_5 > 0$ and

$$HY_i = \hat{a}_1 T_i + \hat{a}_2 S_i + \hat{a}_3 D_i. \qquad (3)$$

with coefficients $\hat{a}_1 > 0$, $\hat{a}_2 < 0$ and $\hat{a}_3 > 0$, where: i is the number assigned to the country, E_i is the per capita household electricity consumption in country i, C_i is the per capita real consumption of households without the consumption of electricity in country i in US dollars (at purchasing power parity), PR_i is the real price of consumption of 1 kWh of residential electricity in US dollars (at purchasing power parity), G_i is the relative frequency of months requiring heating in houses in country i, Q_i is the ratio of energy sources other than electricity energy to all energy sources in household energy consumption, HY_i is the per capita output of the hidden economy, T_i is the ratio of the sum of paid personal income, corporate profit and taxes on goods and services to GDP, S_i is the ratio of public social welfare expenditures to GDP, and D_i is the sum of the number of dependants over 14 years and inactive earners, both per 100 active earners.

   In [7] we have shown that the above method can be adapted to the use for a single country and applied to the case of Macedonia, where significant variations due to external shocks are present. We will discuss these results in next section.

2.4 The Kaufman and Kaliberda method

Electricity input method of Kaufmann and Kaliberda [5], uses a single indicator dependant on hidden economy, the total consumption of electricity in the country (E) in conjunction with the official gross domestic product GDP (Y). In [9] we further developed this method using an analytical expression for the size of hidden economy (HY) for a given year (t):

$$HY(t) = r\left(1 - \left(1 - \frac{HY(0)}{r}\right) \frac{Y(t)}{Y(0)} \left(\frac{E(0)}{E(t)}\right)^{1/\mu}\right), \qquad (4)$$

where $HY(0)$ is the size of hidden economy in the base year ($n = 0$) (determined by a different method), and the parameters μ (the elasticity for electricity consumption (E) with respect to GDP (Y)) and r (relative efficiency of the hidden economy compared to the regular economy). We have shown that this method can be efficiently used for determination of the variations of hidden economy in small open economies, while using a limited set of data sources with outstanding precision.

3 APPLICATION TO HIDDEN ECONOMY OF SOME WESTERN BALKANS COUNTRIES

Examples here are given for Macedonia, as a typical small open economy, and Croatia, as a good role-model for other states in Western Balkans.

First, we show the results for the evolution of the hidden economy as a percentage of reported GDP in Macedonia since its independence in 1991, as it was obtained using the Lackó method [7]. It is seen (see Fig. 2) that there are several overshoots over the baseline of about 32 %. All the overshoots are identified and, except the first one, precisely quantised with Gaussians. Each of the Gaussians is described with a given magnitude, standard deviation and a central year (Fig. 2).
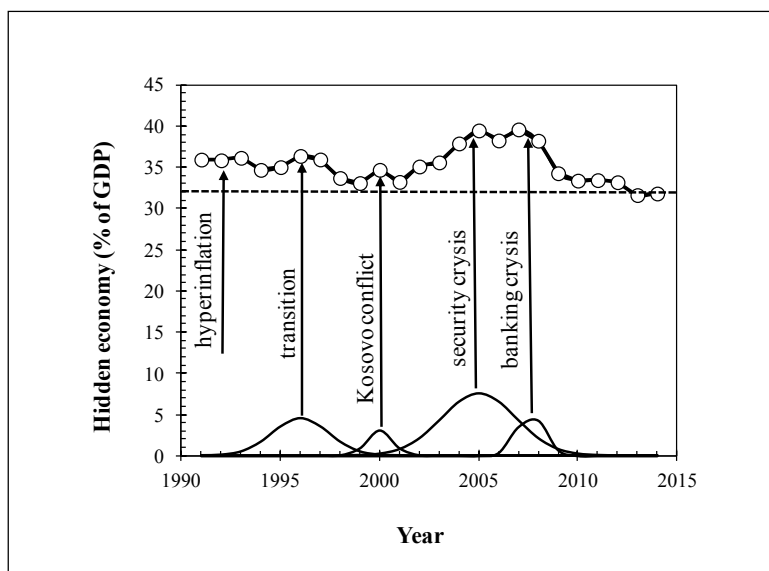
Figure 2: Evolution of the hidden economy in Macedonia since independence until year 2014, as obtained using the Lackó method [7]

First, the peak attributed to the hyperinflation in 1992 is located. It is difficult to be precisely quantised, since in the beginning of independence the data used are not of enough good quality. Then, economic transition from socialist to market economy, lasting roughly 6 years, have caused a temporal increase of about 4 % in the hidden economy. Then, around year 2000, Kosovo conflict influenced shortly Macedonian economy. The most severe shock to the economy was produced by the security crisis (almost war) [4]. Since the crisis peaked in 2002, the damage to the economy lasted exceptionally long (roughly 8 years) and attained very high magnitude of about 8 %. The most recent peak, which is partially mixed with the previous one, is that for banking crisis at around year 2008.

It is to be noted that intensity and duration of the perturbations are bigger in the case of events taking place in the country (transition and security crisis) than in the case of events influencing the country from outside (Kosovo conflict and banking crisis).

Next we show results for the evolution of hidden economy in Macedonia and Croatia in a period close to the independence, where the economy was strongly disturbed by external factors. The first big fluctuation in Croatia is that caused by the war in the beginning of the nineties. Rather high value of hidden economy (up to 45 %) is obtained in this period. Later it sharply decreased. In Macedonia, in this period a sudden shock of hyperinflation caused a short increase around year 1992. Later, a similar behaviour for both countries is observed, due to transition of the economy. It is to be noted that the increase in the beginning progressed faster for Macedonia that for Croatia. At the end of this period, the hidden economy in Macedonia attained somehow higher level than in Croatia.
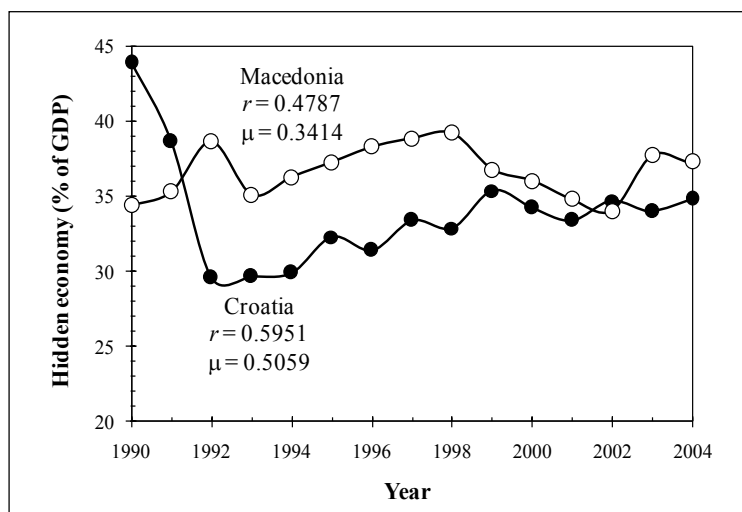
Figure 3: Evolution of the size of hidden economy in Croatia and Macedonia for the period 1990-2004, as obtained by the Kaufmann and Kaliberda method

Based on our studies it can be affirmed that the method of determination of the size of hidden economy based on the energy consumption indicators can provide precise information on the evolution of hidden economy in small open economies, such as Western Balkan countries.

This period, from 1990 to 2004, for comparison of the hidden economies of Macedonia and Croatia was so chosen in order to emphasize the similarities in the case of crisis.

4 CONCLUSIONS

Model methods involving energy consumption are particularly efficient in determination of the size of hidden economic sector in small open economies as are those of the Western Balkan countries. In these models, the quantities (indicators) used in determination of the hidden economy are obtained by precise measurements of a real measurable physical quantity. In addition, instead on *ad hoc* multiple hypotheses, the considered models are based on established economic laws.

Particularly important is the case of the Kaufman and Kaliberda compact method where an analytical nonlinear expression with only two fitting parameters has been obtained, using a single realistic hypothesis. It has been demonstrated that this method provides efficient description of the variation of hidden economy in crisis periods in Western Balkans.

Development of the size of hidden economy in Croatia and Macedonia for the period 1990-2004, as obtained by the Kaufmann and Kaliberda method is studied in details. In Croatia rather high value of hidden economy (45 %) is obtained at the beginning of the war period, but later it declined. In Macedonia, in this period a impulsive shock of hyperinflation caused a short increase of hidden economy around the year 1992. Afterwards, a similar dynamics for both countries is noticed, as the result of transition of the economy. It is to be noted that the increase

in the beginning of the 90-ties progressed faster for Macedonia than for Croatia. At the end of the observed period, the hidden economy achieved somehow higher level in Macedonia than in Croatia.

References

[1] Ardizzi, G., Petraglia, C., Piacenza, M., Turati, G. 2014. Measuring the underground economy with the currency demand approach: a reinterpretation of the methodology, with an application to Italy. Review of Income and Wealth, 60(4): 747–772.

[2] Dumičić, K., Čibarić, I., Horvat, N. 2010. The analysis of monetary transmission mechanism in Croatia using cointegration approach. Croatian Operational Research Review, 1(1): 210–220.

[3] Dumicic, K., Palic, I., Sprajacek, P. 2015. The role of external shocks in Croatia: block exogeneity SVAR approach. Ekonomski i socijalni razvoj, 2(1): 44-54.

[4] Hislope, R. 2003. Between a bad peace and a good war: insights and lessons from the almost-war in Macedonia. Ethnic and Racial Studies, 26(1): 129–151.

[5] Kaufmann, D., Kaliberda, A. 1996. Integrating the unofficial economy into the dynamics of post socialist economies: a framework of analyses and evidence (The World Bank, Policy Research Working Paper 1691, doi:10.1596/1813-9450-1691).

[6] Lackó, M. 2011. The hidden Economies of Visegrad Countries in International Comparison: A Household Electricity Approach. In Halpern, L., Wyplosz, C. (Eds.). Hungary: Towards a Market Economy (pp. 128–152). Cambridge: Cambridge University Press.

[7] Novkovska, B. 2016. How Strongly the Hidden Economy of a Small Country can be Influenced by Drastic Events: Case of Macedonia. UTMS Journal of Economics, 7 (2): 187–195.

[8] Novkovska, B. 2016. The size of the hidden economy in Macedonia: Tendencies and challenges. Proceedings of the ISCCRO: International Statistical Conference in Croatia, 05-06 May 2016, Zagreb, Croatia, 182–189.

[9] Novkovska, B. 2017. Compact total energy consumption method for estimation of the size of hidden economy based on Kaufmann and Kaliberda approach, submitted for publication

[10] Palić, I., Dumičić, K., Barbić, D. 2016, January. The estimation of money demand elasiticity: case of Croatia. In ICHSS 2016: International Conference on Humanities and Social Sciences.

[11] Schneider, F., 2017. Estimating a Shadow Economy: Results, Methods, Problems, and Open Questions. Open Economics, 1(1): 1–29.

[12] Schneider, F., Buehn, A., Montenegro, C.E. 2010. New Estimates for the Shadow Economies all over the World. International Economic Journal, 24(4): 443–461.

[13] Takala, K. Viren, M. 2010. Is cash used only in the shadow economy?. International Economic Journal, 24(4): 525–540.

[14] Tanzi, V. 1980. The underground Economy in the United States: Estimates and Implications. Banca Nationale del Lavoro Quartely Rewiew, 135(4): 427–453.

[15] Tanzi, V. 1999. Uses and abuses of estimates of the underground economy. The Economic Journal, 109(456): 338–347.

[16] Tanzi, V., Blejer, M. I. 1982. Inflation, interest rate policy, and currency substitutions in developing economies: A discussion of some major issues. World Development, 10(9): 781-789.

[17] Williams, C.C. and Horodnic, I.A. 2016. Tackling the undeclared economy in the European Union: an evaluation of the tax morale approach. Industrial Relations Journal, 47(4): 322-340.

[18] Williams, C.C., Horodnic, I.A. 2015. Evaluating the prevalence of the undeclared economy in Central and Eastern Europe: an institutional asymmetry perspective. European Journal of Industrial Relations, 21(4): 389–406.

THE ANALYSIS OF DOMESTIC BALASSA-SAMUELSON EFFECT IN CROATIA: EVIDENCE FROM LONG RUN MODEL

Irena Palić
Faculty of Economics and Business, University of Zagreb, Department of Statistics
Trg J. F. Kennedya
E-mail: ipalic@efzg.hr

Abstract: The domestic Balassa-Samuelson effect in Croatia in long run is assessed in this paper. Previous research shows twofold results of Balassa-Samuelson effect estimation for Croatia. The Johansen cointegration analysis of relative productivity of tradable goods to nontradable goods sector and relative price of nontraded to traded goods is conducted, and the existence of one cointegration relation is estimated. Relative productivity is shown to have positive statistically significant impact on relative price in long run, what confirms domestic Balassa-Samuelson effect in Croatia.

Keywords: Balassa-Samuelson effect, Johansen cointegration approach, relative productivity of tradable goods to nontradable goods sector, relative price of nontradable to tradable goods

1 INTRODUCTION

The impact of relative productivity on relative prices and exchange rate has been widely researched in economic literature for decades. The papers of [1] and [16] recognized the differential of productivity in traded and nontraded goods sector as one of the key determinants of relative prices and real exchange rates. According to Balassa-Samuelson theory, productivity growth in the tradable goods sector will increase wages in that sector and, because of labour mobility between sectors, wages in the nontradable goods sector will also increase. Producers of nontradable goods then increase the prices due to higher wages, which in turn leads to an increase in the overall price level in the economy [12]. Over the past three decades, Central and East European countries experienced the periods of high inflation and real appreciation of domestic currency [10]. Although Croatia is a member of European Union, is not yet the member of Euro area. The Euro area admission criteria, known as the 'convergence criteria' (or 'Maastricht criteria'), are proposed to ensure economic convergence [6]. Among others, the criteria for the introduction of the euro as formal currency are price and exchange rate stability [7]. The Croatian kuna is used as nominal exchange rate anchor for keeping inflation stable. Therefore, the relationship of relative productivity between traded and nontraded goods on the one side, and relative prices of nontraded and traded goods on the other, is important due to its possible implications for real exchange rate, inflation and economic growth.

This paper aims to estimate the domestic Balassa-Samuelson effect in Croatia. The domestic Balassa-Samuelson effect points to the fact that increase in relative productivity of tradable goods sector to nontradable goods sector will lead to increase in relative price of nontraded to traded goods, which will result in real appreciation of domestic currency. Results of previous research of Balassa-Samuelson effect in Croatia are twofold and there is no consensus on the impact of productivity on relative prices.

While the research of [12] find the significant domestic Balassa-Samuelson effect for Croatia, the estimated effect is insignificant in [10]. However, both mentioned papers use linear regression modelling for analysing Balassa-Samuelson effect. The possible shortcoming of using linear regression model is that there is a possibility of misleading results if time series included in the regression model are integrated of order one [11].

This research revisits the analysis of domestic Balassa-Samuelson effect in Croatia and contributes to the existing literature by employing Johansen cointegration approach to analyse domestic Balassa Samuelson effect in Croatia in long run. Although cointegration approach is widely used in testing Balassa-Samuelson effect in international literature, it has not been used for empirical analysis of the mentioned effect in Croatia. Moreover, latest available research of Balassa-Samuelson effect for Croatia by [10] refers to data from 1998 to 2006, and this research will, in addition to using different econometric method, concern data from 2000 to 2016. After the literature review, the long run impact of relative productivity of traded and nontraded goods sector on relative prices of nontraded to traded goods is tested using cointegration approach.

2 PREVIOUS RESEARCH ON BALASSA-SAMUELSON EFFECT IN CROATIA

The Balassa-Samuelson effect is extensively researched in economic literature for decades. The paper of [17] presents numerous papers in this field published from 1964 (when pioneers [1] and [16] published their papers) to 2004. They conclude that statistically insignificant coefficients or coefficients which are not in line with Balassa-Samuelson effect can be found in only six papers out of fifty eight papers analysed.

The comprehensive analysis of Balassa-Samuelson effect in Croatia is provided in research of [10] and [12]. The research of [12] analyses the Balassa-Samuelson effect in six central European economies including Croatia. Both aforementioned papers analyse both domestic and international version of Balassa-Samuelson effect in Croatia using linear regression modelling. The research of [13] estimates the impact of relative labour productivity on price level for 27 European countries in 1999 using cross-section regression analysis. Although the Balassa-Samuelson effect is not estimated individually for Croatia, the author considers that the higher price level in Croatia compared to other transition countries is partially affected by relative labour productivity in traded to nontraded goods sector. Moreover, research of [4] uses the dynamic ordinary least squares regression model and the autoregressive distributed lag model in order to test the Balassa-Samuelson effect in group of countries including Croatia and concludes that, in contrast to the other countries under observation, the Balassa-Samuelson effect is present in Croatia in period 1991-2004.

The domestic version of Balassa-Samuelson effect is represented by equation (1), what is explained in detailed in [12]:

$$\Delta p^{NT} - \Delta p^{T} = \left(\frac{\delta}{\gamma}\right)\Delta a^{T} - \Delta a^{NT}, \tag{1}$$

where p^{NT} refers to price of nontradable goods, p^{T} is price of tradable goods, and a^{T} and a^{NT} refer to total factor productivity in traded and nontraded goods sectors, respectively. Lower-case letters denote logarithmic values.

3 THE EMPIRICAL ANALYSIS OF THE IMPACT OF RELATIVE PRODUCTIVITY ON RELATIVE PRICES IN CROATIA

3.1 Data

Prior to calculating relative productivity in traded goods sector relative to nontraded goods sector, it is defined which activities comprise traded and nontraded sector using NACE statistical classification of economic activities in the European Community [8]. According to [10] the tradables sector usually includes industry, while the nontradable sector is comprised

of services. Agriculture is excluded from the analysis because of high dependency on government subsidies and intervention.

For the purpose of empirical analysis, the traded sector refers to following activities: mining and quarrying (B); manufacturing (C); electricity, gas, steam and air conditioning supply (D); water supply, sewerage, waste management and remediation activities (E). The nontraded sector incorporates: construction (F); wholesale and retail trade repair of motor vehicles and motorcycles (G); transportation and storage (H); accommodation and food service activities(I); information and communication (J); financial and insurance activities (K); real estate activities (L); professional, scientific and technical activities (M); administrative and support service activities (N); public administration and defence and compulsory social security (O), education (P) and human health and social work activities (Q). The mentioned division is in line with [13] and largely in line with [10]. For detailed review of classifying activities in tradable and nontradable sectors, see [10]. The productivity of each sector is approximated by average labour productivity, due to the fact that estimation of capital for each activity, which is essential for calculation total factor productivity, is not available. Quarterly data from 2000Q1 to 2016Q4 are used in cointegration analysis.

Average labour productivity is calculated as the ratio of gross value added (in million euros, in 2010 based chain linked volumes) and number of employed persons. Quarterly data on gross value added are available at [9]. Quarterly employment is calculated as the average of monthly values of number of employed persons in legal entities, according to NACE classification of activities, what is provided by [2]. The relative price of nontradables to tradables are calculated as the difference of logarithmic values of consumer price indices (CPI) and producer price indices (PPI), which are available at [3]. The reference year for both price indices is 2010.

The logarithmic value of relative productivity is calculated as the difference of logarithmic values of productivity in traded sector and productivity in nontraded sector. Seasonal adjustment of logarithmic values of relative productivity and relative price is conducted using the X-13 ARIMA-SEATS quarterly seasonal adjustment method developed by [18]. Hence, seasonally adjusted logarithmic values of relative productivity in traded sector to nontraded sector are denoted by A and seasonally adjusted logarithmic values of relative price of nontradables to tradables are denoted by $PNTT$ are included in cointegration analysis.

Prior to cointegration analysis, the stationarity of both time series is tested using the Augmented Dickey-Fuller (ADF) unit root test. The results of the ADF test are provided in Table 1. Both relative productivity A and relative price $PNTT$ are non-stationary in levels, but stationary in first differences at 5% significance.

Table 1: ADF unit root test *t*-test statistics for variables A and PNTT

Variable	Constant	Constant and trend	No deterministic components
A	0.3461	-1.1869	-1.6487
$PNTT$	-1.6650	-1.7701	-1.6482
ΔA	-9.6189*	-9.5591*	-1.0887
$\Delta PNTT$	-2.1394	-2.2065	-2.1640*

Note: *denotes the stationarity of time series at 5% significance
Source: Author's calculation (EViews 8)

Since both time series are integrated of order (1) at 5% significance. If a linear combination of non-stationary variables is stationary, the variables are cointegrated and the long run relationship exists among variables [5]. Therefore, the Johansen cointegration approach is used in order to examine the existence of long run relationship between relative productivity and relative prices. However, long-run in econometric sense refers to the long-run

relationship between non-stationary variables, thus cointegration does not require the long-run equilibrium to be the result of a market mechanism or behaviour of individuals [14].

3.2 Cointegration analysis of relative productivity and relative price in Croatia

Before conducting cointegration tests, the appropriate model concerning the existence of deterministic components (trend and constant) is selected. The lowest value of Akaike criteria is recorded for model in which linear trend is present in both cointegrating equation and error correction model, while constant is present in cointegrating equation. After choosing the most appropriate model, it is necessary to test the existence of cointegration and determine the number of cointegrating relations.

The number of cointegration relations is examined using trace test and maximum eigenvalue test. The results of the both tests are presented in Table 2. When null hypothesis is rejected for the first time, the conclusion about the number of cointegration vectors can be drawn. The conclusion is brought by comparing empirical test statistics and critical values of the tests. The cointegration among selected variables exists at 10 % significance. After the determination of number of cointegrating relations, the long run equation is estimated. It is necessary to note that lag number for the estimated model equals 9, what is necessary in order to eliminate residual autocorrelation problem, what is explained later in discussion of model diagnostics tests.

Table 2: The results of the trace test and the maximum eigenvalue test

Hypothesized number of cointegrating equations	Eigenvalue	Trace statistic	0.10 Critical value (trace statistic)	Max eigenvalue statistic	0.10 Critical value (max eigenvalue statistic)
0	0.2729	23.5448*	23.3423	18.4890*	17.2341
1	0.0835	5.0558	10.6664	5.0559	10.6664

Source: Authors' calculation
* denotes rejection of the hypothesis at the 0.10 level

The long-run cointegrating equation obtained on the basis of estimated vector error correction model (VECM), with corresponding *t*-values in brackets is given by equation (2):

$$PNTT = 0.2284 - 0,0029 trend + 1.0657 A$$
$$(-6.1795)\ (4.0198)$$

(2)

Relative productivity has positive statistically significant impact on relative price of nontradables to tradables, what confirms the Balassa-Samuelson effect is present in Croatian economy. Furthermore, the research of [15] analyses exchange rate misalignment in Croatia and uses price ratio of tradables to nontradables as one of possible determinants of real effective exchange rate. They concluded that the price ratio of tradables to nontradables has positive impact on real effective exchange rate, i.e. it causes exchange rate depreciation, However, in this research the reciprocal value of price ratio of tradables to nontradables is used, namely price ratio of nontradables to tradables. Therefore, an increase in relative productivity increases the price of nontradables to tradables, what is expected to decrease real effective exchange rate, i.e. cause appreciation of Croatian kuna.

The error correction term (ECT) obtained on the basis of the equation equals -0.4256, with corresponding *t*-statistics equal to -3.7328, what points to the significance of ECT. The negative sign of calculated ECT indicates that variables return to equilibrium, while its value

provides information about the adjustment speed. Therefore, 42.56% of disequilibrium is corrected in each quarter and relative price of nontradables to tradables returns to the equilibrium level for 2.35 quarters, what is approximately 7 months.

Regarding the diagnostics of the model residuals, White heteroskedasticity test is conducted. The White test chi-square test statistic equals 94.1408, with corresponding *p-value* of 0.9124, what indicates that the null hypothesis of homoscedasticity cannot be rejected at any reasonable significance level. Regarding the residual autocorrelation test, the LM test is conducted. The null hypothesis of no autocorrelation of residuals cannot be rejected up to lag length k=12 at 5% significance level, since all corresponding empirical significance levels are higher than 0.05.

The stability of model is checked by calculating the inverse roots of characteristic AR polynomial using EViews 8. The estimated VECM with *r* cointegrating relations is stable if *k-r* roots are equal to unity and the remaining roots have modulus less than one and lie inside the unit circle, where *k* is the number of endogenous variables and *r* is the number of cointegrating relations. The analysis has shown that VEC specification imposes 1 unit root and the remaining roots have modulus less than one. Since there are two variables in the model and one cointegrating relation, the existence of one unit root shows that the system is stable. Therefore, the ECM diagnostic tests show that the estimated model is appropriate. For detailed explanation of problems of heteroskedasticity, autocorrelation as well as AR roots calculation, see [5].

The obtained result is in line with Balassa-Samuelson theory and is interesting in context of joining Euro area and adopting euro, since stable exchange rate and price inflation belong to convergence criteria for euro adoption in Croatia [7]. However, although this research confirms the existence of Balassa-Samuelson effect in long-run, it is necessary to state that changes in relative productivity in Croatia are not of considerable size. The calculated coefficient of variation of rate of change in Croatian relative productivity in the observed period from 2000Q1 to 2016Q7 is 6.71%. Therefore, the Balassa-Samuelson effect does not cause big variations in relative prices and it is not disrupt the fulfilment of Maastricht criteria.

4 CONCLUSIONS

The analysis of domestic Balassa-Samuelson effect in Croatia in long run is conducted in this paper. The domestic Balassa-Samuelson indicates that and increase in relative productivity of tradable goods to nontradable goods sector causes the relative price of nontraded to traded goods to increase, what in turns leads to real exchange rate appreciation. The analysis of relative productivity and relative prices from first quarter 2000 to fourth quarter 2016 has shown that cointegration, i.e. long-run equilibrium exists among two mentioned variables. The relative productivity of tradable goods to nontradable goods has statistically significant positive impact on the relative price of nontraded to traded goods, what is in line with Balassa-Samuelson theory. According to relevant previous research, higher relative price is related to the real appreciation of Croatian kuna. The result of the analysis is interesting in the context of Maastricht criteria for accession into Euro area, namely the stability of exchange rate and inflation. Even though the research points to the presence of domestic Balassa-Samuelson effect in Croatia, the productivity changes in Croatia are not so sizeable to cause big changes in relative prices and exchange rate, and thus the estimated Balassa-Samuelson effect does not disturb fulfilment of Maastricht criteria in Croatia. However, this research analyses the domestic version of Balassa-Samuelson effect and the question of the existence of international Balassa-Samuelson effect in long run in Croatia still remains unanswered. Therefore, future research will be aimed at analysing the long-run impact of

differential of relative productivity between Croatia and Euro area on the differential of relative prices between Croatia and Euro area.

Acknowledgement

References

[1] Balassa, B. 1964. The Purchasing Power Parity Doctrine: A Reappraisal. Journal of Political Economy, 72, 584–596.

[2] Croatian Bureau of Statistics. 2017. Released Data, Employment and Wages, www.dzs.hr [Accessed 12/04/2017].

[3] Croatian National Bank. 2017. Statistical data, Selected non-financial statistics, Table J1: Consumer price and industrial producer price indices, www.hnb.hr [Accessed 20/04/2017].

[4] Egert, B. 2005. Balassa-Samuelson Meets South Eastern Europe, the CIS and Turkey: A Close Encounter of the Third Kind?. The European Journal of Comparative Economics, 2 (2): 221–224

[5] Enders, W. 2015. Applied Econometric Time Series (4th Ed.). London: John Wiley & Sons.

[6] European Commission. 2017a. https://ec.europa.eu/info/business-economy-euro/euro-area/enlargement-euro-area/who-can-join-and-when_en [Accessed 23/05/17]

[7] European Commission. 2017b. https://ec.europa.eu/info/business-economy-euro/euro-area/enlargement-euro-area/convergence-criteria-joining_en [Accessed 23/05/17]

[8] Eurostat. 2017a. Metadata: Statistical Classification of Economic Activities in the European Community, http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrN om=NACE_REV2&StrLanguageCode=EN [Accessed 02/05/2017].

[9] Eurostat. 2017b. Quarterly National Accounts, Gross Value Added, http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=namq_10_a10&lang=en [Accessed 02/05/2017].

[10] Funda, J., Lukinić, G., Ljubaj I. 2007. Assessment of the Balassa-Samuelson Effect in Croatia. Financial Theory and Practice, 31 (4): 321–351.

[11] Granger, C.W.J., Newbold, P. 1974. Spurious regressions in econometrics, Journal of Econometrics, 2 (2): 111–120.

[12] Mihaljek, D., Klau, M. 2003. The Balassa-Samuelson Effect in Central Europe: A Disaggregated Analysis, Bank for International Settlements. Working paper No. 143, Basel

[13] Nestić, D., 2005. Price Level Convergence: Croatia, Transition Countries and the EU. Croatian National Bank. Working paper No. W-13, Zagreb,

[14] Palić, I., Dumičić, K., Barbić, D. 2016. The estimation of money demand elasticity: case of Croatia. International Journal of Research in Business, 4 (9): 89–98.

[15] Palić, I., Dumičić, K., Šprajaček, P. 2014. Measuring real exchange rate misalignment in Croatia: cointegration approach. Croatian Operational Research Review (CRORR), 5: 135–148.

[16] Samuelson, P.A. 1964. Theoretical Notes on Trade Problems. Review of Economics and Statistics, 46, 145–154.

[17] Tica, J., Družić, I. 2006. The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence. Faculty of Economics and Business, University of Zagreb. Working Paper No. 0607, Zagreb

[18] U.S. Census Bureau 2013. X-13 ARIMA-SEATS Reference Manual, http://www.census.gov/ts/x13as/docX13ASHTML.pdf [Accessed 13/03/16].

ANTECEDENTS OF ENTREPRENEURIAL BEHAVIOR: STUDY OF SLOVENIAN AND CROATIAN ECONOMICS AND BUSINESS STUDENTS

Vanja Šimićević
Croatian Studies
Kampus Borongaj, Borongajska cesta 83d, 10000, Zagreb
E-mail: vanja.simicevic@hrstud.hr

Mirjana Pejić-Bach
University of Zagreb, Faculty of Economics and Business
Trg J. F. Kennedy 6, 10 000 Zagreb, Croatia
E-mail: mpejic@efzg.hr

Ana Aleksić
University of Zagreb, Faculty of Economics and Business
Trg J. F. Kennedy 6, 10 000 Zagreb, Croatia
E-mail: aaleksic@efzg.hr

Abstract: By using structural equation modeling and empirical data on the sample of students from Croatia and Slovenia, the purpose of this paper is to examine current antecedents of entrepreneurial behavior, i.e. how an innovative cognitive style and attitudes towards entrepreneurship can influence one's intentions to become an entrepreneur. The results suggest that individual cognition and attitudes towards entrepreneurship, as well as the personal attraction to entrepreneurship, social norms and perceived self-efficacy, can have a positive influence on emergence of entrepreneurial intentions.

Keywords: structural equation modelling, entrepreneurship, entrepreneurial intentions, innovative cognitive style, attitudes towards entrepreneurship

1 INTRODUCTION

Intentions-based models have been successful in investigating the cognition of individuals and their resultant behavior, offering practical framework for understanding and prediction of entrepreneurial behavior [16]. Using the theory of planned behavior (TPB) as their theoretical basis [1], they emphasize the importance of entrepreneurial intentions in predicting one's behavior since intentions are seen as immediate antecedents of actual behavior, i.e. organizational founding. While we have learned that intentions are central to thus entrepreneurial actions, we have not yet explored the pathways to intent [12]. Additionally, further research in this field is necessary since entrepreneurship concerns itself with distinctive ways of thinking and behaving [20]. We need a better, richer understanding of how a cognitive style influences a nascent entrepreneur's development of his or her own perceptions of intentionality [12]. In this study, we examine the complex interaction of a cognitive style, attitudes towards entrepreneurship and entrepreneurial intentions. Taking into consideration that entrepreneurship is a multidimensional phenomenon, the obtained data will be used to test the entrepreneurial intention model using structural equation techniques. The paper is organized into five sections. After an introduction, the research model is set and research propositions are developed. This is followed by the presentation of the research methodology, including sample and research instrument. Research data analysis and main findings are presented afterwards, followed by conclusion and main limitations of the study.

2 THEORY AND RESEARCH PROPOSITIONS DEVELOPMENT

With the intention of better understanding possible relationships between attitudes towards entrepreneurship, an innovative cognitive style and entrepreneurial intentions, we have developed the research model as shown in Figure I. It was particularly important to address complex interactions between analyzed variables from a multidimensional perspective, and their influence on the dependent variable (entrepreneurial intentions) in order to examine multiple pathways and influences [15]. In that sense, we presume that entrepreneurial intentions are under the positive influence of individual attitudes towards entrepreneurship (defined by the theory of planned behavior) (RP1) and an innovative cognitive style (RP2). We also presume that an innovative cognitive style has a positive influence on attitudes towards entrepreneurship (RP3).
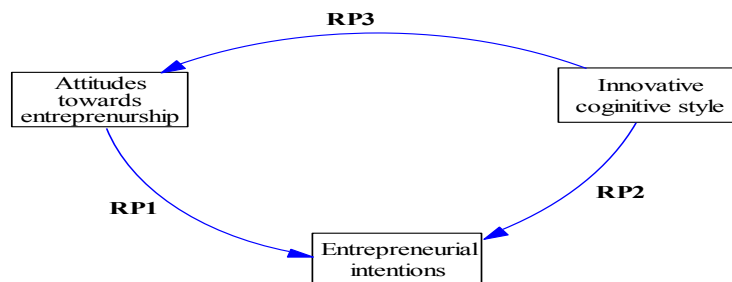
Figure I:. The research model of the impact of attitudes towards entrepreneurship and innovative cognitive style on entrepreneurial intentions

Entrepreneurial intention can be defined as a state of mind that directs individual attentions and actions towards development and creation of a new business venture [6]. Previous empirical analyses indicate the importance of entrepreneurial intentions and how variance in behavior can be attributed to intentions (e.g. [3], [5], [22]). According to the theory of planned behavior, intentions can be significantly predicted by a specific attitude towards behavior and a set of beliefs that an attitude is based on [1]. This attitude is developed on the basis of: two beliefs that reflect the perceived desirability of performing the behavior: personal attitude toward outcomes of the behavior (PA) and perceived normative expectations and social pressures of other people (SN), and a third belief that reflects perceptions that the behavior is personally controllable and forms perceived behavioral control [1], [16], [17]. As such perceived behavioral control reflects individual perception of feasibility of performing and is thus related to perceptions of self-efficacy (PSE) [16]. However, although the link between attitude - intention and intention - behavior has been recognized, empirical findings regarding the scope of connection and influence of different factors leading to intentions are still mixed. Therefore, we aim to test the presence of positive connection between attitudes and intentions. In that sense, we propose our first research proposition (RP$_1$): *"Attitudes towards entrepreneurship positively influence entrepreneurial intentions"*

Entrepreneurial cognition refers to the knowledge structures that people use to make assessments, judgments or decisions involving opportunity evaluation, venture creation, and growth [19]. Innovativeness is especially attributed to entrepreneurs as they are the ones who have to sense opportunities and provide creative and innovative solutions [4]. The Kirton Adaptation-Innovation Inventory (KAI) is a widely accepted measure of an individual cognitive style that posits an individual on a continuum with adaptors and innovators at the extreme ends in terms of creative, problem-solving and decision-making behavior [18], [21]. Adaptors are described as disciplined, conservative, efficient, methodical and value themselves for doing things better and being efficient; innovators are impulsive and quick to

change in their search for a different situation [18]. Several authors (e.g. [2], [4]) have tried to explore the relationship between an individual cognitive style and entrepreneurial behavior. However, little research has been done in the field of cognitive styles and entrepreneurship intentions directly. We aim to analyze whether individuals with an innovative cognitive style will show more intention towards entrepreneurship, and declare our second research proposal (RP$_2$): *"Innovative cognitive style positively influences entrepreneurial intentions."*

It is also possible to expect that an individual's cognitive style can have significant influence on one's attitudes towards entrepreneurship. As Krueger [14] states behind entrepreneurial attitudes are deep cognitive structures. More specifically, an individual cognitive style could have different influences on one's personal attitudes or self-efficacy in different phases of the entrepreneurial process [13]. Based on previous theoretical and empirical findings we propose the third research proposition (RP$_3$): *"Innovative cognitive style positively influences attitudes towards entrepreneurship."*

3 RESEARCH METHODOLOGY

In order to test our research propositions we used a self-report questionnaire as a method to collect the responses on a sample of bachelor and master students in economics and business studies in Croatia and Slovenia. Questionnaires were completed by the total of 400 students, 39.5% from Croatia and 60.5% from Slovenia. 62,5% of them were female, and 75,3% at master level. More than 73% of them stated they have considered becoming an entrepreneur.

Our research instrument consisted of three parts. The first two parts measured independent variables: (1) attitudes towards entrepreneurship and (2) innovative cognitive style. Attitudes towards entrepreneurship, based on the Ajzen's TPB, were presented as an aggregate measure comprising of personal attraction (PA-consisting of 5 items), social norms (SN-consisting of 3 items) and perceived self-efficacy (PSE-consisting of 4 items). The innovative cognitive style, measured by KAI was also presented as an aggregate measure comprised of willingness to try (WTT-consisting of 6 items), creative-original (CO-consisting of 6 items) and opinion-leader & ambiguities problems (OL&AP-consisting of 5 items). The third part of the research instrument included questions regarding personal characteristics and entrepreneurial intention, seen as a dependent variable and binomial with the yes/no response (it measures if a respondent has ever seriously considered becoming an entrepreneur). This type of research item has been widely accepted and previously used (e.g. [15]). The independent variables include gender, the year of the study and the country of the study. All items in the questionnaire are based on the theory and have taken previous work as a reference, i.e. Goldsmith. [10] for cognitive style, and Liñán and Chen [17] for different attitudinal antecedents of entrepreneurial intention. All the used multi-item measures were based on the 7-point Likert scales from 1 (strongly disagree) to 7 (strongly agree).

4 DATA ANALYSIS AND RESEARCH FINDINGS

Data collected were analyzed using various statistical methods. The validity analysis was first conducted. The explanatory factor analysis was performed in order to test convergent validity. Iterated principal axis factor combined with varimax rotation for extracting six factors was conducted by means of statistical package SAS. The approach suggested by Costello and Osborne [7] was used by applying a loading cut-off value in the magnitude from 0.40 to 0.70. According to the defined criteria, all of the measurement factors were to be retained. Our factor analysis confirmed the existence of six factors. However, it has to be declared that KAI originally consists of four factors (willingness to try, creative-original, opinion-leader and ambiguities-problems), and for the purpose of this survey two factors, i.e.

opinion-leader and ambiguities-problems, were merged into one factor, because the factor analysis consisting of seven factors did not reveal interpretable results. The discriminate validity was also checked with the usage of the confirmatory factor analysis [8]. All of the t-values exceed the threshold value of 1.96, as proposed by Costello and Osborne [7]. We can conclude that they are statistically significant at the level of .01. Therefore, it can be further concluded that the observed loading paths are operationally precise and appropriate. All of the items were statistically significant at the level of .01 Second, the reliability analysis was conducted. The questionnaire, originally in the English language, was translated into the Croatian and the Slovenian language and its validity confirmed. To test the reliability of the scales we computed Cronbach's alpha coefficients. As all of the calculated Cronbach's alpha coefficients were larger than .70, we could confirm the internal consistency of the scales' items [9]. Third, the descriptive data analysis and the non-parametric correlation analysis were conducted in order to check if problems in data due to validity exist, which could be concluded based on negative or low correlations [8]. Most coefficients showed that there was a medium to low correlation between items representing attitudes towards entrepreneurship and an innovative cognitive style. The results, although moderately, emphasize the connection between examined items which indicates to a positive connection between attitudes to entrepreneurship and an innovative cognitive style. Fourth, the model fit was measured. The SAS module was used for developing the structural equations model according to the proposed conceptual model, and produced a chi-square of 1281.773 with 418 degrees of freedom. Table I presents the indices used for assessing the overall model validity. Goodness of fit statistics (GFI) was used for measuring the correspondence between observed and hypothesized variance. According to the recommendations of Hooper et al. [11], GFI should be higher than 0.90. The GFI in our model was 0.860, which is quite near to the advocated norm, and adjusted goodness-of-fit index was 0.834, which could also be considered appropriate. In addition, the values of Normed-fit index (NFI) and Non-normed-fit index (NNFI) indicated a good fit since they were near the proposed value of 0.9. The value of comparative-fit index (CFI), that is .914, was also satisfactory. In addition, Root-mean-square-error (RMSEA) revealed the acceptable 0.063 value. In short, it is possible to make a conclusion that the research model is valid, and it is up to the before mentioned methodological prerequisites.

Table I: Fit indices for the research model

Fitness indicator	Model estimated	Explanations
Chi-square (χ^2)	1281.773	χ^2 is not significant
Degrees of freedom (df)	418	
p-value	0.000	
χ^2/df	3.067	Good, close to 3
GFI	.860	Good, close to .9
AGFI	.834	Good
NFI	.878	Good, close to .9
NNFI	.904	Very good result
CFI	.914	Very good result
RMSEA	.063	<.07, good result
90% confidence interval of RMSEA	(.059 - .067)	Upper limit <.07, very good result

Finally, statistical testing of the research propositions, statistical significance of the parameters and the amount of variance of endogenous constructs accounted for by independent constructs were tested using the structural equations model. The results of the path analysis are shown in Figure II.
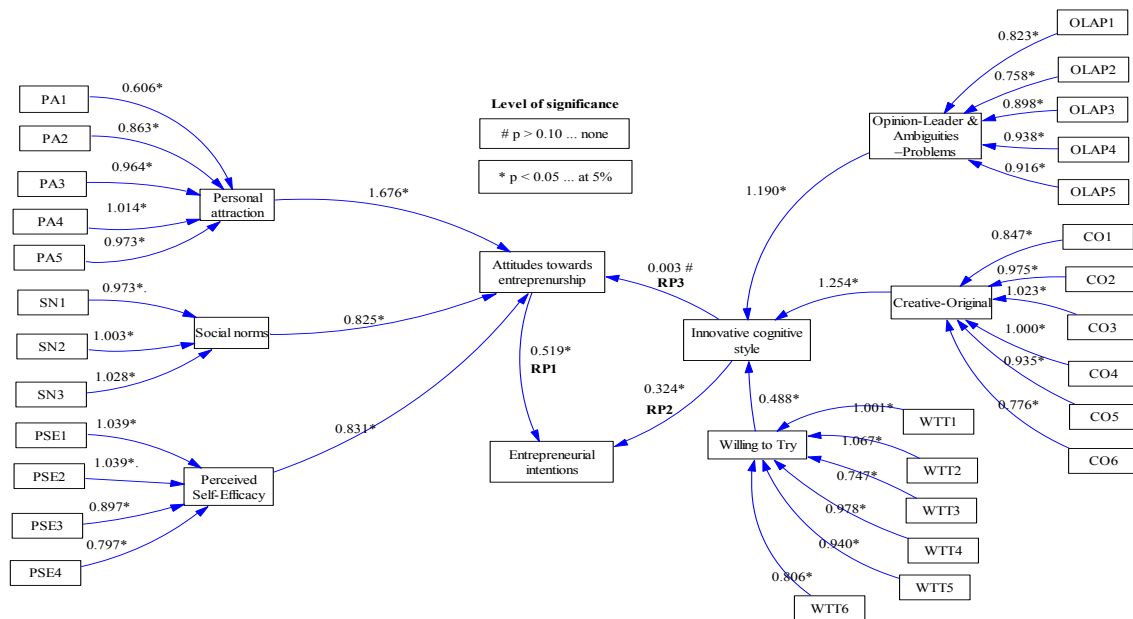
Figure II: Path diagram with path coefficients estimates and their significance levels

The first research proposition (RP_1) was accepted. The standard solution of the path coefficient estimate from attitudes towards entrepreneurship and entrepreneurial intention was .519 with the t-value of 7.636, which indicated the existence of a positive effect at the 5% significance level. The second research proposition (RP_2) yielded the standard solution of the path coefficient estimate from an innovative cognitive style to entrepreneurial intentions of .324 with the t-value of 3.287. In addition, R^2 value was .390 indicating that 39.0% of variations in entrepreneurial intentions could be explained by variations in attitudes toward entrepreneurship, and variations in an innovative cognitive style. For the third research proposition (RP_3) standard solution of the path coefficient estimate is .003 with the t-value of 8.594, which indicated that the effect was not present between the constructs. R^2 value was .279 indicating that only 27.9% of variations in attitudes towards entrepreneurship could be explained by variations in an innovative cognitive style, which eventually led to the rejection of the third research proposition.

5 CONCLUSION

This paper was motivated by the need to additionally explore entrepreneurial intention and its antecedents in Croatia and Slovenia. Based on our research two out of three research propositions have been supported with empirical data suggesting a positive connection between (1) attitudes towards entrepreneurship and entrepreneurial intentions and (2) an innovative cognitive style and entrepreneurial intentions. Besides, the paper sheds a new light on the relationship between an innovative cognitive style and attitudes towards entrepreneurship. Although with a medium effect of innovative cognitive style on intentions, a positive nature of the relationship has been found. This indicates that cognition-based perspectives can have a significant influence on entrepreneurial behavior. Our research, has several limitations. First, data are self-reported and from a common source, thus more associated to subjectivity and individual perception. By using structural equation modeling, we tried to minimize the problem of a common method variance and multicollinearity. Second, only direct relationships in the model have been analyzed. Future research should investigate the influence of additional variables such as economic and situational conditions or cultural values and norms that could mediate and affect analyzed relationships.

References

[1] Ajzen, I. 1991. The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2): 179-211.

[2] Allinson, C. W., Chell, E., Hayes, J. 2000. Intuition and entrepreneurial performance. European Journal of Work and Organizational Psychology, 9(1): 31–43.

[3] Armitage, C. J., Conner, M. 2001. Efficacy of the theory of planned behaviour: A meta-analytic review. British Journal of Social Psychology, 40(4): 471–499.

[4] Armstrong, S. J., Hird, A. 2009. Cognitive style and entrepreneurial drive of new and mature business owner-managers. Journal of Business Psychology, 24(4): 419–430.

[5] Autio, E., Keeley, R. H., Klofsten, M., Parker, G. G. C., Hay, M. 2001. Entrepreneurial intent among students in Scandinavia and in the USA. Enterprise and Innovation Management Studies, 2(2): 145–160.

[6] Bird, B. 1988. Implementing entrepreneurial ideas: The case for intention. The Academy of Management Review, 13(3): 442-453.

[7] Costello, A. B., Osborne, J. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research & Evaluation, 10(7): 1-9.

[8] de Vaus, D. 2001. Research design in social research. London: Sage Publications.

[9] Feldt, L. S., Kim, S. 2008. A comparison of tests for equality of two or more independent alpha coefficients. Journal of Educational Measurement, 45(2): 179-193.

[10] Goldsmith, R. E. 1991. The validity of a scale to measure global innovativeness. Journal of Applied Business Research, 7(2): 89-97.

[11] Hooper, D., Coughlan, J., Mullen, M. R. 2008. Structural equation modelling: Guidelines for determining model fit. Electronic Journal of Business Research Methods, 6(1): 53-60.

[12] Kickul, J., Krueger, N. 2005. Toward a new model of intentions: The complexity of gender, cognitive style, culture, social norms, and intensity on the pathway to entrepreneurship. Center for Gender in Organizations. Working paper No. 20., Boston.

[13] Kickul, J., Gundry, L. K., Barbosa, S. D., Whitcanack, L. 2009. Intuition versus analysis? Testing differential models of cognitive style on entrepreneurial self-efficacy and the new venture creation process. Entrepreneurship Theory and Practice, 33(2): 439-453.

[14] Krueger, N. F. jr. 2007. What lies beneath? The experiential essence of entrepreneurial thinking. Entrepreneurship Theory and Practice, 31(1): 123-138.

[15] Krueger, N. F. jr., Carsrud, A. L. 1993. Entrepreneurial intentions: Applying the theory of planned behaviour. Entrepreneurship and Regional Development, 5(4): 315-330.

[16] Krueger, jr., N. F., Reilly, M. D., Carsrud, A. L. 2000. Competing models of entrepreneurial intentions. Journal of Business Venturing, 15(5-6): 411–432.

[17] Liñán, F., Chen, Y.W. 2009. Development and cross-cultural application of a specific instrument to measure entrepreneurial intentions. Entrepreneurship Theory and Practice, 33(3): 593-617.

[18] Marcic, D., Willey, S., Johnson, L. 1990. Adaptors and innovators: Success in business school. The Journal of Applied Business Research, 6(2): 98-103.

[19] Mitchell, R. K., Busenitz, L., Lant, T., McDougall, P.P., Morse, E. A., Smith, J. B. 2002. Toward a theory of entrepreneurial cognition: Rethinking the people side of entrepreneurship research. Entrepreneurship Theory and Practice, 27(2): 93-104.

[20] Mitchell, R. K., Busenitz, L., Bird, B., Gaglio, C. M., McMullen, J., Morse, E., Smith, J. 2007. The central question in entrepreneurial cognition research. Entrepreneurship Theory and Practice, 31(1): 1-27.

[21] Stum, J. 2009. Kirton's adaption-innovation theory: Managing cognitive styles in times of diversity and change. Emerging Leadership Journeys, 2(1): 66-78.

[22] Tkachev, A., Kolvereid, L. 2010. Self-employment intentions among Russian students. Entrepreneurship & Regional Development: An International Journal, 11(3): 269-280.

IMPACT OF PICTURES ON RESPONSE RATES IN BUSINESS WEB SURVEYS: CROATIAN CASE

Berislav Žmuk
University of Zagreb, Faculty of Economics and Business – Zagreb
Trg J. F. Kennedyja 6, HR-10000 Zagreb, Croatia
E-mail: <bzmuk@efzg.hr>

Abstract: Response rates in web surveys have been very low in the last decade. The aim of the paper is to inspect whether pictures have a statistically significant impact on response rates in business web surveys. In order to investigate the issue, three versions of a questionnaire have been developed and sent to a sample of Croatian enterprises. The only difference among the questionnaire versions refers to whether they have included pictures or not. The conducted statistical analysis has shown that there is no statistically significant difference in response rates to questionnaires without and with pictures.

Keywords: Chi-square test, Croatian enterprises, pictures, proportion difference test, response rate, web survey.

1 INTRODUCTION

Internet penetration rates in countries worldwide are rising and in some countries, the maximum has almost been reached [see 10]. Consequently, web surveys are commonly used by researchers, enterprises, students and other institutions as a data collection method [9]. The main advantages of web surveys are low costs and the speed of data collection. Furthermore, web surveys can include different multimedia items, such as pictures, videos, music and similar, which make web surveys more attractive to respondents [3]. However, response rates to web surveys in most cases are quite low [8]. This problem can be softened by careful design of questionnaires, by giving some incentives to respondents and by well-designed invitation letters. Furthermore, the response rate could be higher if some reliable and well-known organisations or enterprises support the research [6].

In this paper, the emphasis will be place on the questionnaire design problems. To be more precise, the dilemma to include or not to include pictures in questionnaires in business web surveys will be observed. On the one hand, if pictures are carefully selected and placed on the right places in the questionnaire, pictures can improve questionnaire design by increasing concertation of respondents and by helping them answer the questions [5]. However, on the other hand, if pictures are not carefully selected and placed on the right places in the questionnaire, they can also be a hindrance to a web survey. So, respondents could have different technical difficulties with pictures, respondents could consider them inappropriate and pictures could even be biased by suggesting to respondents, directly or indirectly, what they should answer [4].

Whereas the impact of pictures on response rates in web surveys in which respondents are individual persons is investigated [2, 5, 11, 15], no attention is given to business web surveys in which respondents are employees who provide answers on behalf of their enterprise. Consequently, the research hypothesis of the paper is that if pictures are carefully chosen and placed on right places in the questionnaire, they do not have an impact on response rates in business web surveys. For the purpose of the research, a web survey in Croatian enterprises will be conducted and response rates will be observed and compared.

After a brief introduction, in the second chapter, main characteristics of the conducted web survey in Croatian enterprises are given and the approach to the analysis is explained. In the third chapter, response rates according to different characteristics of enterprises are calculated and presented whereas in the fourth chapter, statistical tests are used for comparison of

response rates. The fifth, final, chapter presents conclusions and guidelines for further research.

2 DATA AND METHODS

In order to determine whether the impact of pictures in business web surveys on response rates is present, a business web survey about statistical methods use in Croatian enterprises is conducted. Overall 37,855 Croatian enterprises were invited by e-mail to participate in the web survey in October 2016. Two reminders were also sent and the web survey was closed at the end of December 2016.

In the web survey, information about different characteristics of enterprises was provided. So, enterprises were observed according to their size, main activity, legal form and geographical location of their headquarters. According to their size, enterprises are stratified into small, medium and large. In order to group enterprises according to their size the Accounting Act [14] was used. The National Classification of Economic Activities [12] was used to stratify enterprises according to their main activity and four groups of main activities of enterprises are recognized: industrial, trade, service and other enterprises. Only limited liability enterprises as defined in the Enterprises Act [13] participated in the survey. Consequently, according to their legal form the distinction was made between joint stock enterprises, limited liability enterprises, and simple limited liability enterprises. Finally, enterprises are going to be observed according to the location of their headquarters. The Nomenclature of territorial units for statistics – second level (NUTS 2) will be used and consequently, enterprises from the Continental Croatia and from the Adriatic Croatia will be observed [7].

Three different versions of a questionnaire were developed. In the first questionnaire version no pictures were shown to enterprises. In the second questionnaire version "positive" pictures were shown to enterprises whereas the third questionnaire version had "negative" pictures. Positive pictures show something positive, for instance a positive trend line or a table full of statistical books. Negative pictures show opposite things (a negative trend line, a table with only two statistical books). Pictures are carefully chosen and only five of them were included in the questionnaire. It has to be emphasized that only one questionnaire version has been offered to each enterprise. In the analysis, the minimum response rate or Response Rate 1 [1] will be observed. The Response Rate 1 is calculated as a ratio of fully completed questionnaires and contacted respondents. The statistical tests, the proportion difference test and the chi-square test, will be used to inspect whether there is a statistically significant difference in response rates between different questionnaire versions.

3 RESPONSE RATES IN THE BUSINESS WEB SURVEY

In this paper, response rates are calculated by taking into account only fully completed questionnaires. This ensures that a respondent (an employee in an enterprise) has seen all pictures in the provided questionnaire. Consequently, the impact of pictures on respondent's decision to complete the questionnaire can be observed. Table 1 provides response rates obtained when all enterprises are observed. According to Table 1, overall 780 enterprises have participated in the web survey and completed the provided questionnaire. Consequently, the Response Rate 1 of the web survey is 2.06%. If the number of contacted enterprises is observed, it can be seen that almost the same number of enterprises got a questionnaire without pictures, with positive pictures and with negative pictures. Furthermore, the response rates seem to be quite similar among the observed questionnaire versions. However, the presence of a statistically significant difference will be inspected later.

Statistics	Web survey:				Total (without + with pictures)
	Without pictures	With pictures	With positive pictures	With negative pictures	
Contacted enterprises	12,619	25,236	12,619	12,617	37,855
Completed web surveys	268	512	259	253	780
Response Rate 1 (in %)	2.12	2.03	2.05	2.01	2.06

Table 2: Response rates, enterprises stratified according to their size

Size	Statistics	Web survey:				Total (without + with pictures)
		Without pictures	With pictures	With positive pictures	With negative pictures	
Small	Contacted enterprises	12,128	24,258	12,130	12,128	36,386
	Completed questionnaires	256	492	248	244	748
	Response Rate 1 (in %)	2.11	2.03	2.04	2.01	2.06
Medium	Contacted enterprises	379	755	378	377	1,134
	Completed questionnaires	8	16	8	8	24
	Response Rate 1 (in %)	2.11	2.12	2.12	2.12	2.12
Large	Contacted enterprises	112	223	111	112	335
	Completed questionnaires	4	4	3	1	8
	Response Rate 1 (in %)	3.57	1.79	2.70	0.89	2.39

In Table 2 response rates for enterprises stratified according to their size are observed. Despite the small number of medium and large enterprises that participated in the web survey, the response rates are quite similar to response rates at small enterprises. The reason for that can be found in the fact that the number of medium and large enterprises is very small in comparison to the number of small enterprises in the country.

Table 3: Response rates, enterprises stratified according to their main activity

Main activity	Statistics	Web survey:				Total (without + with pictures)
		Without pictures	With pictures	With positive pictures	With negative pictures	
Industrial	Contacted enterprises	3,762	7,507	3,754	3,753	11,269
	Completed questionnaires	66	148	68	80	214
	Response Rate 1 (in %)	1.75	1.97	1.81	2.13	1.90
Trade	Contacted enterprises	3,197	6,397	3,196	3,201	9,594
	Completed questionnaires	54	105	52	53	159
	Response Rate 1 (in %)	1.69	1.64	1.63	1.66	1.66
Service	Contacted enterprises	5,259	10,531	5,269	5,262	15,790
	Completed questionnaires	134	243	132	111	377
	Response Rate 1 (in %)	2.55	2.31	2.51	2.11	2.39
Other	Contacted enterprises	401	801	400	401	1,202
	Completed questionnaires	14	16	7	9	30
	Response Rate 1 (in %)	3.49	2.00	1.75	2.24	2.50

If response rates for the questionnaire version without pictures are observed, Table 3 reveals that the category of other enterprises has the highest response rate. On the other side, if the questionnaire version with pictures is observed, service enterprises have the highest response rate.

According to Table 4, when all questionnaire versions are observed, simple limited liability enterprises have the highest response rates. On the other hand, limited liability

enterprises seem to have the lowest response rates. However, limited liability enterprises have convincingly the highest number of completed questionnaires.

Table 4: Response rates, enterprises stratified according to their legal form

Legal form	Statistics	Web survey:				Total (without + with pictures)
		Without pictures	With pictures	With positive pictures	With negative pictures	
Joint stock enterprises	Contacted enterprises	241	480	240	240	721
	Completed questionnaires	5	11	7	4	16
	Response Rate 1 (in %)	2.07	2.29	2.92	1.67	2.22
Limited liability enterprises	Contacted enterprises	11,879	23,761	11,881	11,880	35,640
	Completed questionnaires	246	476	239	237	722
	Response Rate 1 (in %)	2.07	2.00	2.01	1.99	2.03
Simple lim. liability enterprises	Contacted enterprises	499	995	498	497	1,494
	Completed questionnaires	17	25	13	12	42
	Response Rate 1 (in %)	3.41	2.51	2.61	2.41	2.81

Table 5: Response rates, enterprises stratified according to their NUTS 2 region

NUTS 2 region	Statistics	Web survey:				Total (without + with pictures)
		Without pictures	With pictures	With positive pictures	With negative pictures	
Continental Croatia	Contacted enterprises	7,815	15,635	7,803	7,832	23,450
	Completed questionnaires	171	338	171	167	509
	Response Rate 1 (in %)	2.19	2.16	2.19	2.13	2.17
Adriatic Croatia	Contacted enterprises	4,804	9,601	4,816	4,785	14,405
	Completed questionnaires	97	174	88	86	271
	Response Rate 1 (in %)	2.02	1.81	1.83	1.80	1.88

Table 5 shows that enterprises from the Continental Croatia are more willing to participate in a web survey than the enterprises from the Adriatic Croatia. Moreover, it has to be taken into account that more enterprises have headquarters in the Continental Croatia than in the Adriatic Croatia, which resulted in the fact that the number of enterprises from the Continental Croatia which participated in the survey was almost twice the number of enterprises from the Adriatic Croatia.

4 ANALYSIS OF THE IMPACT OF PICTURES ON RESPONSE RATES

After the response rates are calculated the differences among them will be observed. What is compared first is whether response rates achieved by using the questionnaire version without pictures are equal to those achieved by using the questionnaire version with pictures. For the purpose of this analysis, questionnaire versions with positive and with negative pictures will be observed together. In order to determine whether the difference in response rates is present, the statistical proportion difference test will be used. Table 6 shows main results of conducted statistical tests for all the observed characteristics of enterprises.

According to results from Table 6 the largest difference in response rates, when response rates for questionnaire versions without and with pictures are compared, is present at large enterprises (0.0178) and at other enterprises (0.0149). However, at the significance level of 5%, all conducted statistical tests led to the conclusion that the null hypothesis cannot be rejected. In other words, it can be concluded that there is no statistically significant difference in response rates at each of the observed enterprises characteristics when response rates to questionnaire versions without and with pictures are compared.

Table 6: Proportion difference test, without pictures Response Rate 1 vs. with pictures Response Rate 1

Characteristic of enterprises	Proportion difference	Common proportion	Standard error	z-score	p-value
All					
All	0.0009	0.0206	0.0015	0.61	0.5399
Size					
Small	0.0008	0.0206	0.0016	0.52	0.6006
Medium	-0.0001	0.0212	0.0091	-0.01	0.9926
Large	0.0178	0.0239	0.0177	1.01	0.3147
Main activity					
Industrial	-0.0022	0.0190	0.0027	-0.80	0.4259
Trade	0.0005	0.0166	0.0028	0.17	0.8631
Services	0.0024	0.0239	0.0026	0.93	0.3507
Other	0.0149	0.0250	0.0095	1.57	0.1175
Legal form					
Joint stock enterprises	-0.0022	0.0222	0.0116	-0.19	0.8520
Limited liability enterprises	0.0007	0.0203	0.0016	0.43	0.6694
Simple limited liability enterprises	0.0089	0.0281	0.0091	0.99	0.3240
NUTS 2 region					
Continental Croatia	0.0003	0.0217	0.0020	0.13	0.8964
Adriatic Croatia	0.0021	0.0188	0.0024	0.86	0.3890

Table 7: Chi-square test of proportions equality, without pictures Response Rate 1 vs. with positive pictures Response Rate 1 vs. with negative pictures Response Rate 1, degrees of freedom = 2

Characteristic of enterprises	Chi-square	p-value
All		
All	0.4454	0.8003
Size		
Small	0.3062	0.8580
Medium	0.0001	0.9999
Large	1.7941	0.4078
Main activity		
Industrial	1.6671	0.4345
Trade	0.0378	0.9813
Services	2.6401	0.2671
Other	2.6513	0.2656
Legal form		
Joint stock enterprises	0.8989	0.6380
Limited liability enterprises	0.1906	0.9091
Simple limited liability enterprises	1.0077	0.6042
NUTS 2 regions		
Continental Croatia	0.0814	0.9601
Adriatic Croatia	0.7538	0.6860

In order to inspect whether questionnaire versions with positive or with negative pictures have different impacts on response rates than the questionnaire version without pictures, the chi-square test of proportions equality is conducted. Under the null hypothesis of the chi-square test, it is assumed that all proportions are equal. According to results from Table 7, at the significance level of 5%, the null hypothesis cannot be rejected at none of the observed characteristics of enterprises. Consequently, it can be concluded that response rates to the observed questionnaire versions are the same no matter which characteristic of enterprises is observed. That way the research hypothesis that pictures do not have an impact on response rates in business web surveys can be accepted.

5 CONCLUSIONS

Web surveys can be enriched with different multimedia items. However, if those multimedia items are not necessary to be included in the questionnaire (for specific research aims), the question whether multimedia items should be included in the questionnaire or not remains open. Because of that, the paper analyses whether pictures should be included in a questionnaire of a business web survey or not. The answer to that research question has been reached by comparing response rates. Namely, response rates to web surveys are very low nowadays. Therefore, any improvement of questionnaire design that would lead to higher response rates is highly desirable.

If response rates of enterprises that completed a questionnaire without pictures and of enterprises that completed a questionnaire with pictures are the same, it can be concluded that pictures in questionnaires do not have an impact on response rates. The conducted analysis has shown that this is true for each of the observed characteristics of enterprises.

However, it has to be emphasized that the pictures that have been used in the questionnaires have been carefully chosen and placed in the questionnaires. Furthermore, only five pictures have been included in the questionnaires. In the further research, the impact of pictures on answers in business web surveys should be investigated.

References

[1] AAPOR. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf [Accessed 15/02/2017].

[2] Cape, P., Phillips, K. 2015. Questionnaire Length and Fatigue Effects: The Latest Thinking and Practical Solutions. https://www.surveysampling.com/site/assets/files/1586/questionnaire-length-and-fatiigue-effects-the-latest-thinking-and-practical-solutions.pdf [Accessed 15/02/2017].

[3] Couper, M. P. 2008. Designing Effective Web Surveys. New York: Cambridge University Press.

[4] Couper, M. P., Conrad, F. G., Tourangeau, R. 2007. Visual Context Effects in Web Surveys. Public Opinion Quarterly, 71(4): 623–634.

[5] Couper, M. P., Tourangeau, R., Kenyon, K. 2004. Picture This!: Exploring Visual Effects in Web Surveys. Public Opinion Quarterly, 68(2): 255–266.

[6] Dillman. D. A., Smyth, J. D., Christian, L. M. 2014. Internet, Phone, Mail and Mixed-mode Surveys: The Tailored Design Method. Hoboken: John Wiley & Sons.

[7] Eurostat. 2015. Regions in the European Union: Nomenclature of territorial units for statistics NUTS 2013/EU-28. Luxembourg: Publications Office of the European Union.

[8] Evans, J. R., Mathur, A. 2005. The value of online surveys. Internet Research, 15(2): 195–219.

[9] Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., Tourangeaeau, R. 2004. Survey Methodology. Hoboken: John Wiley & Sons.

[10] Internet World Stats. 2017. World Stats. http://www.internetworldstats.com/stats.htm [Accessed 15/02/2017].

[11] Kivu, M. 2010. Long Questionnaires: Impact on Abandon Rate. Bucharest: IPSOS-Romania.

[12] Official Gazette. 2007. The National Classification of Economic Activities. Zagreb: Narodne novine d.d. 16(58).

[13] Official Gazette. 2011. The Enterprises Act. Zagreb: Narodne novine d.d. 20(152).

[14] Official Gazette. 2015. The Accounting Act. Zagreb: Narodne novine d.d. 24(78).

[15] Toepoel, V., Couper, M. P. 2010. Can Verbal Instructions Counteract Visual Context Effects in Web Surveys?. Public Opinion Quarterly, 75(1): 1–18.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 2:
High-Performance Computing and Big Data and General OR Topics

OPERATIONS RESEARCH AS THE BRIDGE OVER TECHNOLOGICAL VALLEY OF DEATH

Drago Bokal,
Anja Goričan,
Faculty of Natural Sciences and Mathematics, University of Maribor,
Koroška cesta 160, Maribor, Slovenia,
d@bokal.net, gorican.a@gmail.com

Abstract: In the paper, we discuss the process of knowledge transfer from basic academic knowledge to marketable off-the-shelf products and services. Using the scale of technology readiness levels, introduced by NASA and adopted by European Comission for H2020 projects, we address the problem of the valley of death appearing at intermediate technology readiness levels, where the scientific challenges have been solved, but the technology is not yet operational. We propose an abstract tool – Cartesian product of two ontologies, the one describing methodologies originating at low technology readiness levels and the other describing challenges at high technology readiness levels. We argue that this approach has properties that help combating grade inflation and support achieving psychologically optimal experience – flow – to participants in scientific research, technological development, and education.

We present a case study of the ontology of photovoltaic electricity quantities forecasting. The prototype builds on surveying existing bibliography and on our own experiments.

Keywords: knowledge transfer ontology, Cartesian product, technology readiness level, valley of death, photovoltaic electricity production, forecasting.

1   INTRODUCTION

Science in general and mathematics in particular is expanding the knowledge of mankind. Yet it is fair to say that the process of how their knowledge actually benefits the mankind is fairly complex and obscured to most participants in this process. As clear goals, instantaneous feedback, experience of control, and balance between ability and opportunity are key elements of flow – psychologically optimal experience, improving understanding of the process would contribute to this state in many stakeholders of the knowledge discovery process.

For some domains where the operational environment causes severe risks, such as space exploration [8], or where there is greater public interest in measuring the progress of knowledge applicability, such as EU's H2020 public grants scheme [15], the progress has been codified into a scale of technology readiness levels. As an example, H2020 projects use the following: (1) basic principles observed, (2) technology concept formulated, (3) experimental proof of concept, (4) technology validated in lab, (5) technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies), (6) technology demonstrated in relevant environment, (7) system prototype demonstrated in operational environment, (8) system complete and qualified, (9) actual system proven in operational environment.

Although the above are not universal, it can be argued that there is a point in the development of knowledge or technology where it stops being interesting for scientists, as "now only the trivial technical problems need to be solved" but is not yet interesting for business to be applied, as "this technology has not solved any problems yet". The stage(s) where knowledge development stops being funded through research grants and is not yet funded by market-focused businesses is termed "the valley of death" [2], a term explained by the valley in the chart plotting amount of resources used at various technology readiness levels, see Figure 1. In this paper, we propose a systematic approach to bridging the valley of death and argue that this approach has several advantages for the stakeholders in the process: researchers, students, and businesses. In Section 2, we describe the structure of each side of the valley of death and taxonomize the inputs that side gives into the knowledge maturing process.

In Section 3, we propose to combine the two ontologies from each side of the process using a plain Cartesian product of the instance sets of ontologies' knowledge bases, and act on this Cartesian product using several criteria that compare the elements using application-related criteria. We continue by discussing the impact of the proposed approach to knowledge
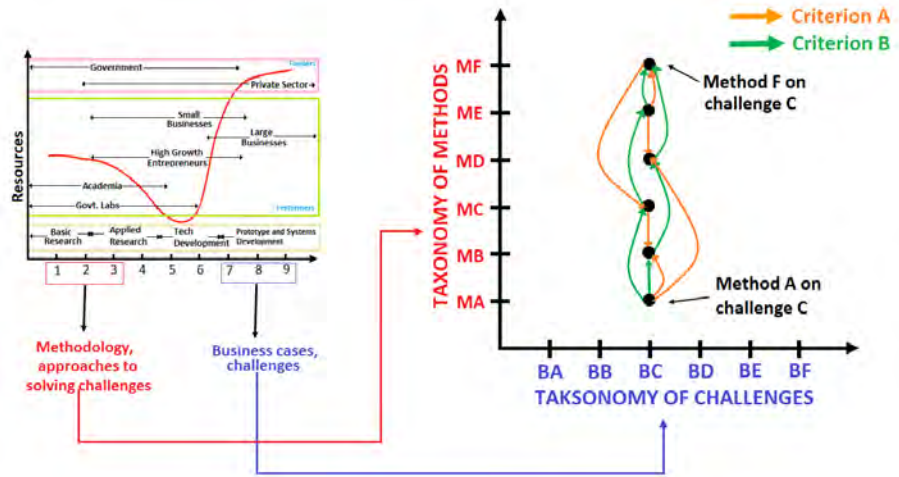
Figure 1: Figure of death valley chart.

discovery and maturing process in Section 4 and conclude by a case study on electric energy quantities forecasting from [13]. Big Data applications encompassing many modelling applications seem to be among the domains, most benefiting from the proposed systematization.

2 TWO SIDES OF THE VALLEY OF DEATH

We define the two sides of the valley of death as (a) the contributing one spanning from the birth of new knowledge till the stage it stops being interesting to novelty-oriented scientists, and (b) the benefiting one that spans over the stages where the stakeholders applying the knowledge can benefit from the application. The earlier is mostly characterized by the ability of scientists to produce publishable results, and the latter is characterized by the ability of the users of knowledge to apply it at a lesser cost than is the cost of not applying it and using the existing solutions. When the new technology introduces changes into the established processes of its application, it is disrupting, and its application is hindered by the cost of learning and adopting new processes and the risks of unforeseen impacts, as well as by the pressure of the users benefiting most from the current system against loosing their benefits.

The two sides of the death valley provide different inputs to the knowledge discovery process: the benefiting side provides challenges that can be solved using the methodologies developed by the contributing, the scientific side. Operations research and data science constitute the contributing side of many disciplines, ranging from engineering, logistics, medicine, social sciences to the very direct business applications. Hence development of clear models of knowledge transfer and improving our understanding the process is of interest to operations researchers and data scientists. We emphasize that most obvious application of the model proposed in this contribution is in bridging the valley of death when transferring knowledge from academic to business sector, but it can also be used within sciences when applying methods from one domain (such as data science) in a different scientific field (such as medicine).

3 KNOWLEDGE TRANSFER ONTOLOGY

Ontologies are models that computer scientists use to represent knowledge. The limited space of this paper does not allow for a detailed treatment of formalities – in terms of technology readiness levels, it is a report on the basic principles observed – and the reader can find rigorous definitions of ontologies in [4], and a detailed treatment of our case study in a parallel submission to SOR 2017 [6]. Our ontology consists of three subontologies. Two are in the core: one

taxonomizing challenges of the benefiting side, and another taxonomizing methodologies of the contributing side. Already earlier, taxonomies have been proposed for introducing systematics into operations research: one of the earliest applications has been the three-letter-notation used in scheduling [7]; it has been adapted to the employee timetabling problema [3]. Recently, we experimented with systematic applications in several domains, such as [14], [9], [5], [11], [12], [13].

Ontology of use cases. The ontology of use cases has knowledge base instances that constitute problem sets for the methods that are being researched. The minimum requirement for this ontology is that its concept set contains the identifiers for use cases that practitioners meet in their domain. In this minimum setting, the instances of the ontology constitute the set of all possible use cases and the concept hierarchy of the ontology is their taxonomy. However, the ontology of use cases can be extended from pure taxonomy to encompass also problem instances, and can hence constitute an exchange standard for these instances, allowing the practitioners at high levels of technology readiness to contribute data to research.

Ontology of methodologies. As for use cases, the knowledge base of the instances in the ontology of methodologies are implementations of methodologies used to address the aforementioned use cases. As the minimum, it has the methodologies for solving the problems as its concept set. It can be augmented to encompass interfaces between implementations and data: respecting such interfaces allows new methodologies to be easily applicable to existing data or new data to be easily tested on existing methodologies.

Cartesian product of ontologies. If we assume that each methodology is applicable to each use case or allow non compatible elements to be paired in the (abstract) instance set, then the Cartesian product of the methodology knowledge base's instance set and the use case knowledge base's instance set constitutes a formal space describing the experiment of applying all instances of methodologies to all instances of the use cases – constituting the knowledge base of the elementary Cartesian product of the two ontologies. Each element of this knowledge base instance set is a pair of a methodology that can theoretically be applied to the use case in the same pair. On a selected use case instance, the methodologies perform better or worse, which can be assessed using the comparison criteria in the next section.

Taxonomy of comparison criteria. When deciding which methodology instance to use in a specific application, one is faced either with a research problem of no known methodology fitting directly to that use case and the need to adopt one (or several) of the methodologies performing well on similar use cases, or one is faced with a (possibly large) body of bibliography describing applications of various methodologies to a popular use case. Not studying the cases in greater detail, one runs the risk of a competitor offering better service, yet the studying is a burden on the resources that again may expose the user to be outrun by competition. Our ontology already encompasses one key information, required by practitioners: the pairs of use cases and methodologies can be used to label papers that study the corresponding applications. The set of all used labels containing an use case would then correspond to all the methods that have been applied to that use case, and the practitioner would be able to select the most appropriate one among them. However, papers also compare the methods on the use cases, defining an implicit directed graph whose vertices are the used labels and whose arcs are directed from worse methodologies to the better ones. The criteria for comparing the instances label the arcs. An instance pair of an use case and a methodology from the knowledge base is in the Pareto frontier, if and only if whenever there is an instance that improves a criterion if compared to the given pair, it also worsens another criterion. The set of Pareto optimal instance pairs is of interest to practitioners: for a specific use case, it contains the methodologies that can form a suitable compromises for particular application.

4 IMPACT ON THE KNOWLEDGE MATURING PROCESS

We foresee three directions of impact on knowledge discovery and maturing process if the ontological model of knowledge transfer is applied at a greater extent. First, the taxonomized problem domain would allow researchers to focus on problems that have greater similarity among each other, pinpoint best approaches through applying taxonomized criteria, and focus on improving the pareto frontier defined by these criteria. As the data on current knowledge would be better organized, literature search would be more efficient – not only citations, but also relevance of the contribution is recorded by the proposed system.

Second, taxonomized methodology domain would help practitioners efficiently identify best approaches and the researchers whose work is in the pareto frontier of their specific problem and whose consulting would be most beneficent to the practitioner.

Third, experience shows that students of mathematics desire courses, where they apply their theoretical knowledge in a higher technology readiness setup. This is understandable, as for most of them, their career perspective lies on the benefiting side of the death valley. In the perspective of empowering students with knowledge directly relevant for their career, such summative courses contribute to the understanding of the purpose of the basic, formative courses that teach definitions and examples of concepts. Using the approach of knowledge transfer taxonomy, the overview of the challenge landscape and the methodology landscape would be more easily understood to students, and the choice of most suitable method for a problem they meet more informed. They could also contribute meaningful knowledge: BSc students could read papers and understand them sufficiently to detect the use case and the methods the paper describes, and which criteria are used to compare the methods in the paper. This would empower BSc students to find suitable methods and researchers when employed. Masters students could partake in a more difficult task of reproducing results of the papers on existing or new data samples, and strengthen or weaken evidence on comparisons of various methods. All these would empower students to use their learning process to contribute new elements to the knowledge base of the topic they are studying, thus aligning their interests with the interests of the professors and their future employers, which is according to recent models of grade inflation a key element in its elimination [10]. Furthermore, the clarity and efficiency of represented information on applications of specific methodologies on specific cases contributes to flow – psychologically optimal experience – in education, research, and practice, which is elaborated in [1]. It addresses all four content related (out of eight total) conditions for flow: (i) it sets **clear goals** to researchers (find a method that is on a specific use case in its pareto frontier), to students (study a paper and understand the instances, criteria, and comparisons it addresses), and to practitioners (select the use case and the criteria, and investigate the pareto frontier). (ii) it contributes to **instantaneous feedback** to all stakeholders, as their contribution is easily (and algorithmically) comparable to the existing knowledge, (iii) it contributes to the experience of control, as the stakeholders have better overview over the development of the field prior, during, and after their research; technological augmentations, such as regular updates would enable them to follow the state of their field more efficiently, and (iv) the balance between ability and opportunity is more available: from the easiest tasks, such as adding an arc to the ontology to the most complex ones, such as adding a new methodology, or improving a pareto frontier: challenges range from BSc students to significant research projects.

5 PHOTOVOLTAIC ENERGY PRODUCTION FORECASTING: A CASE STUDY

While we briefly over viewed several knowledge transfer ontologies in Section 3, we will summarize a recent case study. In a greater detail, it is presented in a parallel submission to SOR 2017 [6]. It addresses the problem of forecasting electricity production of photovoltaic power plants. We summarize the taxonomy of use cases, methodologies, and comparison criteria, and present a fraction of the knowledge instance set as a figure revealing its structure.
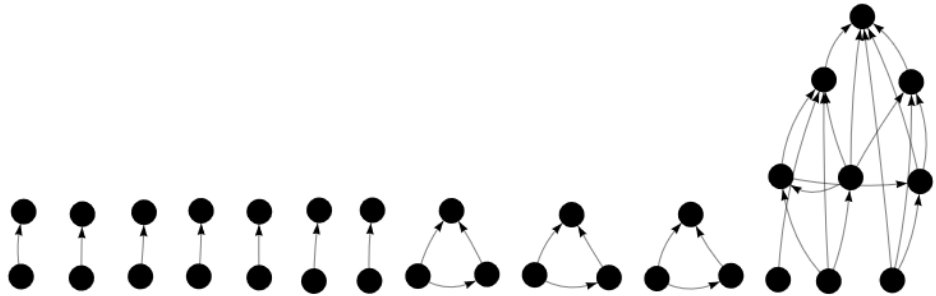
Figure 2: Taxonomy of approaches to forecasting quantities in electricity energy system.

The use cases depend on the process in which the forecast is used. For photovoltaic production, the process is classified by the purpose of forecasting (ie. balancing, market, research), forecasting quantity, forecasting group, voltage level, maturity of forecast, the delay between the last measurement and the first forecast.

Given the use case, we can use the taxonomy to examine which modelling approaches have been successful forecasting data for it. The sources of this information are published papers, internal technical reports, or experiments with the target data set. Studying them is very time consuming, implying that a tool summarizing that information and pointing out the best approach in a given use case seems of high relevance to the research and practice community. We propose to systematically record the following most relevant parameters of the modelling approach: the available data about electricity production, weather that will be used for forecasting (usually, the temperature and solar irradiation), resolution of the input data, length of the training period, modelling technique, criterion optimized by the model fitting, software implementation, hardware platform.

When forecasting with different modelling techniques, we need to have criteria to compare models to each other. One dimension describing criteria used to compare instances is model comparison criteria. This criteria can be one of the errors or coefficient of determination R^2 or information criteria like BIC and AIC or the cost or errors or execution speed of the algorithm. Most common criteria are different errors such as MAE, MAPE, SMAPE, RMSE. Second dimension of criteria for comparing instances is difference, which determines difference between criteria of evaluation, and it can be high (more than one percentage point), medium (between 0.5 and 1 percentage point), low (between 0.1 and 0.5 percentage point) or negligible (less than 0.1 percentage point).

All methodologies that form a connected graph belong to a set of papers that study data from the same business cases and are therefore comparable. A vertex that is at the end of an arrow represents a methodology that is better according the criterion labelling that arc than the methodology at the beginning of the arrow; usually this implies a more accurate forecast, but it can also mean greater technology readiness level of the approach, or more efficient algorithm implementation. Figure 2 shows all articles entered during the project "Po kreativni poti do praktinega znanja" and during previous studies, which were presented in [12].

6 Acknowledgements

References

[1] Bokal D. (2017). Optimizing enjoyment of mathematics and OR education with introducing psychological concepts flow and grit using simulation-based model of emotional states of learning. Submitted to SOR 2017.

[2] Butter M., Fischer N., Gijsberts G. Hartmann Ch., de Heide M. and van der Zee F. (2014). Horizon 2020: Key Enabling Technologies (KETs), Booster for European Leadership in the Manufacturing Sector. Study for the ITRE Committee. Brussels, Belgium.

[3] De Causmaecker, Patrick and Vanden Berghe, Greet. (2011). A categorisation of nurse rostering problems. Journal of Scheduling, Vol(14): 3–16.

[4] Ehrig Marc. (2007). Ontology Alignment: Bridging the Semantic Gap. New York: Springer Science+Business Media.

[5] Gologranc T., Jerebic J., Kranjc J., Lužar B., Mali L., Povh J., Bokal D. (2015). On applying mathematical models of frequency assignment to Wi-Fi throughput optimization. 13th International Symposium on Operational Research in Slovenia: 137-142.

[6] Goričan A., Bratuša A. and Bokal D. (2017). Knowledge transfer ontology of photovoltaic electricity. Submitted to SOR 2017. production forecasting.

[7] Graham R.L., Lawler E.L., Lenstra J.K. and Rinnooy-Kan A.H.G. (2008). Optimization and approximation in deterministic sequencing and scheduling: a survey. Annals of discrete mathematics, Vol(5): 287–326.

[8] Mankins, John C. (1995). Technology readiness levels. White Paper, Vol(6): 1995.

[9] Repnik M. and Bokal D. (2015). A basis for taxonomy of fuzzy linear programming methods. 13th International Symposium on Operational Research in Slovenia: 188–192.

[10] Smole, A., Kent, D., Jagrič, T., and Bokal, D. (2017). Social dilemma model of grade inflation calls to end focusing on grades in favor of knowledge. Submitted.

[11] Šmigoc S. (2016). Primerjava pristopov k napovedovanju porabe električne energije. Univerza v Mariboru, Fakulteta za naravoslovje in matematiko. 57593. Maribor.

[12] Tajnik M. (2016). Integrirani avtoregresijski modeli s premikajočimi sredinami za napovedovanje porabe električne energije. Univerza v Mariboru, Fakulteta za naravoslovje in matematiko. 64888. Maribor.

[13] Tajnik M., Bokal D. (2017). Taksonomija primerjav tehnik za napovedovanje porabe električne energije. 13. konferenca CIGRE.

[14] Tavčar R., Dedič J., Bokal D. and Žemva A. (2014). Towards a Decision Support System for Automated Selection of Optimal Neural Network Instance for Research and Engineering. Citeseer.

[15] G. Technology Readiness Levels. (2015). European Commission. $https$: $//ec.europa.eu/research/participants/data/ref/h2020/wp/20142015/annexes/h2020$ $-$ $wp1415 - annex - g - trlen.pdf$ [Accessed 27/03/2017].

REDUCTIONS OF GROUP SYMMETRIC SEMIDEFINITE PROGRAMS

Igor Đukanović

University of Maribor, Razlagova 14, 2000 Maribor, Slovenia

Abstract: Dramatic reductions of size can be achieved on the semidefinite programs on which a permutation group is acting. Numerous examples arise in combinatorial optimization problems on graphs with rich automorphism groups. This paper is a selfcontained presentation of the implicit version of de Klerk-Pasechnik-Schrijver reduction method.

Keywords: group symmetric semidefinite programs, interior point methods.

1 INTRODUCTION

In this paper, inspired by the breakthrough in reducing symmetric SDPs developed in the sequence of papers [5], [3] and [2], the method described in [2] is "rediscovered" from the implicit point of view while applying only elementary tools. The method is particularly interesting for combinatorial optimization problems on vertex transitive or other graphs with large automorphism groups (e.g., computing crossing number[2]).

2 GROUP SYMMETRIC SEMIDEFINITE PROGRAMS

We will denote the space of symmetric $n \times n$ matrices by $\mathcal{S}^n$ and the cone of positive semidefinite matrices by $\mathcal{P}$. Trace inner product

$$\langle A, B \rangle = \operatorname{tr} AB^T = \sum_{i,j=1}^{n} a_{ij}b_{ij} \tag{1}$$

is a well-known matrix inner product. By $A \succeq 0$ and $A \succ 0$ we will denote that matrix A is positive semidefinite, respectively positive definite. A semidefinite program can be formulated as

$$
\begin{aligned}
P := \quad & \min \langle C, X \rangle \\
& \langle A_i, X \rangle = b_i \qquad i = 1, ..., m \\
& X \succeq 0
\end{aligned} \tag{2}
$$

where $C \in \mathcal{S}^n$, $b \in \mathbb{R}^m$ and $A_i \in \mathcal{S}^n$, $i = 1, ..., m$. As we want to apply reliable primal-dual methods we will be interested only in the case when (2) equals its dual

$$
\begin{aligned}
D := \quad & \max b^T y \\
& Z := C - \textstyle\sum_{i=1}^{m} y_i A_i \succeq 0.
\end{aligned} \tag{3}
$$

Let H be a group of permutations of the set $V := \{1, ..., n\}$. Let $\pi \in H$. Denote by e_i the i-th column of identity matrix and by P_π the permutation matrix such that

$$P_\pi e_i = e_{\pi(i)} \qquad \forall i \in V.$$

Definition 2.1 *Matrix M is group H symmetric, if $P_\pi^T M P_\pi = M$ for all $\pi \in H$.*

Notice that $m_{ij} = e_i^T M e_j = e_i^T P_\pi^T M P_\pi e_j = e_{\pi(i)}^T M e_{\pi(j)} = m_{\pi(i)\pi(j)}$, so Definition 2.1 is equivalent to

$$m_{ij} = m_{\pi(i)\pi(j)} \qquad \forall i, j \in V, \pi \in H. \tag{4}$$

In the group representation theory this property is refered as *matrix M possesses the symmetry of the group H.* (Notice that this does not imply $M^T = M$.) The advantage of the group symmetric matrix is obvious. As many entries in the matrix are equal the space complexity of the problem is decreased. As we will see shortly it can effect the time complexity as well.

Definition 2.2 *Semidefinite program (2) is group H symmetric, if for each $\pi \in H$ the feasible set*

$$\mathcal{F} := \{X|\ \langle A_i, X\rangle = b_i\ i = 1, ..., m\} = \{X|\ \langle A_i, P_\pi^T X P_\pi\rangle = b_i\ i = 1, ..., m\}$$

is preserved, and the cost

$$\forall X \in \mathcal{F}:\ \langle C, X\rangle = \langle C, P_\pi^T X P_\pi\rangle$$

as well.

Let X be a feasible solution of a group H symmetric semidefinite program (2). Then by Definition 2.2 $P_\pi^T X P_\pi$ is also feasible and has the same cost as X. Likewise the *group H-average of X,*

$$\bar{X} = \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T X P_\pi \tag{5}$$

is also feasible and has the same cost as X.

Proposition 2.3 *Group H-average is group H symmetric.*

Proof. We will use a lot the fact that

$$\rho \in H \Rightarrow H \circ \rho = \{\pi \circ \rho|\ \pi \in H\} = H \tag{6}$$

to simplify at the first glance very technical sums. So

$$P_\rho^T \bar{X} P_\rho = \frac{1}{|H|} \sum_{\pi \in H} P_\rho^T P_\pi^T X P_\pi P_\rho = \frac{1}{|H|} \sum_{\pi \circ \rho \in H} P_{\pi \circ \rho}^T X P_{\pi \circ \rho} = \bar{X},$$

i.e., $\bar{x}_{\rho(i)\rho(j)} = \bar{x}_{ij}$. $\qquad \square$

So the primal (2) has a strictly interior group H symmetric initial point $\bar{X}$ whenever it has strictly interior point X.

Next, let us define group H-averages

$$\bar{C} = \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T C P_\pi \qquad \bar{A}_i = \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T A_i P_\pi, \tag{7}$$

and notice that the group H-average of Z

$$\bar{Z} = \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T Z P_\pi = \bar{C} - \sum_{i=1}^m y_i \bar{A}_i \tag{8}$$

does not need to be feasible for dual.

This motivates reformulation of (2) so that the new dual allows group symmetric solution, too.

$$
\begin{array}{ll}
\bar{P} := & \min\langle \bar{C}, X\rangle \\
& \langle \bar{A}_i, X\rangle = b_i \quad i = 1, ..., m \\
& X \succeq 0
\end{array}
\qquad
\begin{array}{ll}
\bar{D} := & \max b^T y \\
& Z := \bar{C} - \sum_{i=1}^m y_i \bar{A}_i \succeq 0.
\end{array}
\tag{9}
$$

Notice that (7) satisfy $\langle \bar{C}, X\rangle = \langle C, X\rangle$ and $\langle \bar{A}_i, X\rangle = \langle A_i, X\rangle$ for any feasible solution X of (2). Likewise any feasible solution $Z = C - \sum_{i=1}^m y_i A_i$ of (3) defines the feasible solution $\bar{Z}$ of (9) with the same cost. Therefore optima can only be improved

$$P \geq \bar{P} \geq \bar{D} \geq D.$$

But since $P = D$, the equality holds throughout.

Since $P_\pi^{-1} = P_\pi^T$, any feasible solution X of (9) satisfies

$$\langle \bar{A}_i, X \rangle = \langle P_\pi \bar{A}_i P_\pi^T, X \rangle = \langle \bar{A}_i, P_\pi^T X P_\pi \rangle,$$

and likewise for $\bar{C}$, therefore (9) is group H symmetric semidefinite program. However as the group H-average $\bar{X}$ is a solution with the same cost as X, it suffices to consider only group symmetric feasible solutions. Likewise the corresponding dual, by definition consisting of group symmetric matrices only, is invariant under the action of the group H, too.

So we have restated the original problem (2) as a primal-dual pair (9) without duality gap which is stated in terms of group symmetric matrices only. This reduces the storage requirements.

3 INTERIOR POINT METHODS

In this section we just recall several well-known facts (see e.g., textbook [1] and paper [5]). Given primal-dual pair of semidefinite programs (e.g., (2)-(3) or (9)) and a feasible pair $X \succeq 0$ and $Z = C - \sum_{i=1}^m y_i A_i \succeq 0$ the duality gap

$$\langle C, X \rangle - b^T y = \langle C, X \rangle - \sum_{i=1}^m y_i \langle A_i, X \rangle = \langle Z, X \rangle \geq 0.$$

When both the primal and the dual have strictly feasible solutions (i.e., when Slater's qualification holds) there are matrices X^* and Z^* such that

$$\langle X^*, Z^* \rangle = 0.$$

Obviously this pair is optimal. For positive-semidefinite matrices this is equivalent to

$$X^* Z^* = 0.$$

Under mild conditions (e.g., if both the primal and the dual have feasible solutions) there is for each $\mu > 0$ a unique primal-dual feasible pair of matrices which satisfies

$$X_\mu Z_\mu = \mu I. \tag{10}$$

The set of all such pairs forms the central path (which turns out to be smooth).

Primal-dual interior point methods start with a feasible pair. The motivation is to approach so defined central path and at the same time decrease μ toward zero at each iteration. By dropping the quadratic term in equation for the next iterate, we get

$$XZ + \Delta \tilde{X} Z + X \Delta Z = \mu I \tag{11}$$

or equivalently

$$\Delta \tilde{X} = \mu Z^{-1} - X - X \Delta Z Z^{-1}. \tag{12}$$

To symmetrize, we apply the HRVW-KSH-M [4] search direction

$$\Delta X = \frac{\Delta \tilde{X} + \Delta \tilde{X}^T}{2}. \tag{13}$$

Together with maintaining the primal and dual feasibility

$$\begin{aligned} \forall i : \ &\langle A_i, \Delta X \rangle = 0, \\ &\Delta Z = - \sum \Delta y_i A_i \end{aligned} \tag{14}$$

this implies the (linear system of) normal equations in Δy

$$\langle A_i, \mu Z^{-1} - X + X \sum \Delta y_i A_i Z^{-1} \rangle = 0 \tag{15}$$

or shortly

$$K \Delta y = r \tag{16}$$

where K is the system matrix and the righthand vector r is defined coordinatewise by

$$r_i = -\langle A_i, \mu Z^{-1} - X \rangle.$$

The algorithm is starts with a primal-dual feasible initial pair X_0, Z_0 and terminates in polynomial time producing an almost optimal pair X_i, Z_i (with duality gap less than ε).

4 REDUCTIONS IN THE MATRIX MULTIPLICATION

In this section we recall results of Kanno, Katoh, Murota and Ohsaki [5] (for well-known facts see also the textbook [1]). First notice that, if matrices M and N are group H symmetric, i.e.,

$$\begin{aligned} P_\pi^T M P_\pi &= M, \\ P_\pi^T N P_\pi &= N \end{aligned} \qquad \forall H \in H,$$

then (since any permutation matrix satisfies $P_\pi P_\pi^T = I$) their product, sum, transposes and inverses are also group H symmetric

$$\begin{aligned} P_\pi^T M N P_\pi &= P_\pi^T M P_\pi P_\pi^T N P_\pi = MN, \\ P_\pi^T (M + N) P_\pi &= P_\pi^T M P_\pi + P_\pi^T N P_\pi = M + N, \\ P_\pi^T M^T P_\pi &= (P_\pi^T M P_\pi)^T = M^T. \\ P_\pi^T M^{-1} P_\pi &= (P_\pi^T M P_\pi)^{-1} = M^{-1}. \end{aligned}$$

Recall the central path (10) for the group H symmetric primal-dual pair (9)

$$Z_\mu X_\mu = \mu I. \tag{17}$$

Let $\pi \in H$. Since Z_μ and X_μ are feasible, so are $P_\pi^T Z_\mu P_\pi$ and $P_\pi^T X_\mu P_\pi$ as well. And the latter pair also solves (17). However, it is well-known that the solution of (17) is unique under mild assumptions, for example, it suffices that both primal and dual have strictly feasible points [6]. Therefore the solution Z_μ, X_μ of (17) is group H symmetric.

We will apply the HRVW-KSH-M search direction (13) to approximately follow the central path (17). The dual update computed in the (12) is group H symmetric by construction. Computing $\Delta \tilde{X}$ in (13) involves only inverses, products and sums of group H symmetric matrices. Therefore ΔZ in (14) is also group H symmetric.

Since we know that products, sums, transposes and inverses (and therefore the update ΔX, too) are group H symmetric, we need to compute only one element per orbit. In the case of strictly feasible group symmetric initial pair this considerably decreases the computational effort. At the same time this reduction does not seem to jeopardize numerical stability since even for SDP (2) and its feasible initial matrix X_0 which is not group symmetric numerical experiments [5] suggest that the computed iterates tend to (by approaching central path) become group symmetric up to machine error.

5 TECHNICAL DETAILS

Next we formalize the idea. Consider multiplying group symmetric matrices Z and X. The orbits of the action of the group H on $n \times n$ matrices can be defined by matrices

$$D_{[ij]} := \sum_{\pi \in H} P_\pi^T (e_i e_j^T) P_\pi \qquad (18)$$

since entries in this matrix are constant along the orbit and zero elsewhere. Notice also that the entry in ith row and jth column of $D_{[ij]}$ is a positive integer, and

$$D_{[ji]} = D_{[ij]}^T. \qquad (19)$$

So given an H group symmetric matrix $X = \sum d_{[ij]} D_{[ij]}$ its transpose

$$X^T = \sum d_{[ij]} D_{[ij]}^T = \sum d_{[ij]} D_{[ji]}$$

is also group H symmetric with just switched pairs of "coordinates" $d_{[ij]}$.

Denote all different so-obtained matrices (18) by $D_1, ..., D_r$, i.e., let there be r orbits. Obviously $\langle D_i, D_j \rangle = 0$, if $i \neq j$. There are vectors $x, p \in \mathbb{R}^r$ such that

$$X = \sum_{j=1}^r x_j D_j \qquad ZX = \sum_{k=1}^r p_k D_k \qquad (20)$$

resulting in smaller space complexity. Next see that $ZX = \sum_{j=1}^r x_j ZD_j$. So

$$p_k = \frac{\langle ZX, D_k \rangle}{\langle D_k, D_k \rangle} = \frac{1}{\langle D_k, D_k \rangle} \sum_{j=1}^r x_j \langle ZD_j, D_k \rangle \qquad (21)$$

implies that $p = Lx$ where

$$l_{kj} = \frac{\langle ZD_j, D_k \rangle}{\langle D_k, D_k \rangle} \qquad (22)$$

are entries in a (small) matrix $L \in \mathbb{R}^{r,r}$. We will denote this *pre-multiplication matrix* also by L_Z. In typical application we can compute these numbers in $\mathcal{O}(r^2)$ flops (l_{kj} involves computing a single entry in the product ZD_j) since we do not form matrices D_i explicitly. Analogous treatment of the product $\Delta ZX = \sum_{k=1}^r q_k D_k$ produces a matrix R_X such that $q = R_X \Delta z$. So (11) can be simplified to

$$L_Z x + R_X \Delta z + L_Z \Delta x' = \mu i \qquad \text{where} \quad I = \sum_{j=1}^r i_j D_j.$$

As trace inner product (1) can be stated as sum of diagonal elements of a matrix product, the same approach by which we have just "vectorized" matrix equation (11) also works for (14). Notice that there may not be more than r linearly independent (typically trivial) matrices A_i, Therefore the formation and solving of the normal equations (15) obtained from Newton's system (12)-(14) also involves only $r \times r$ and smaller matrices. This is the most important breakthrough as inverting the system matrix is typically the bottleneck in the interior point methods.

Up to this point we have recited the ideas from [5]. Observe that applying these reductions of dimension from $n \times n$ to $r \times r$-matrix products and inverses in computations like (12) is just an overoptimization as the next bottleneck becomes checking wheather the candidate for the next iterate is positive definite. This is done by computing Cholesky decomposition while the dimension of the involved matrices is $n \times n$, too. In the next two sections we adopt results from [2] to efficiently solve this problem.

6 DE KLERK-PASECHNIK-SCHRIJVER'S METHOD

Let us denote by $\sigma(A)$ the set of all eigenvalues of the matrix A.

Fact 6.1 $\sigma(Z) = \sigma(L_Z)$.

Proof. It is well-known that group symmetric matrices have multiple eigenvalues. Consider $Zx = \lambda x$. Then

$$P_\pi^T Z P_\pi x = \lambda x \Rightarrow Z P_\pi x = \lambda P_\pi x$$

shows that $P_\pi x$ is also an eigenvector. Define eigenvector

$$\bar{x} = \sum_{\pi \in H} P_\pi x,$$

and notice that $\rho \circ H = H \forall \rho \in H$ implies that $\bar{x}$ is also group symmetric in the sense that

$$P_\rho \bar{x} = \bar{x} \; \forall \rho \in H.$$

The matrix $\bar{X} = \bar{x}\bar{x}^T$ is group H symmetric. Therefore there exists $a \in \mathbb{R}^r$ such that $\bar{x}\bar{x}^T = \sum_{i=1}^r a_i D_i$. Now $Z\bar{X} = Z\bar{x}\bar{x}^T = \lambda\bar{x}\bar{x}^T$ is equivalent to

$$L_Z a = \lambda a.$$

This proves $\sigma(Z) \subseteq \sigma(L_Z)$. To prove the converse, start by any eigenvector, i.e. $L_Z a = \lambda a$. Define $\bar{X} := \sum_{i=1}^r a_i D_i$, and see that its columns are eigenvectors of Z. $\qquad \square$

This reduces the complexity of line search to $\mathcal{O}(r^3)$, too. So overall complexity of one step in IPM drops from $\mathcal{O}(n^6)$ where n is the dimension of matrix variables in (9) to $\mathcal{O}(r^3)$.

References

[1] E. de Klerk. (2002). *Aspects of Semidefinite Programming*. Kluwer Academic Publishers.

[2] E. de Klerk, D. V. Pasechnik, and A. Schrijver. (2007). Reduction of symmetric semidefinite programs using the regular ∗-representation. *Mathematical Programming*, 109(2-3):613–624.

[3] K. Gatermann and P. A. Parrilo. (2004). Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Appl. Algebra*, 192 (number 1-3):95–128.

[4] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz. (1996). An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361.

[5] Y. Kanno, M. Ohsaki, K. Murota, and N. Katoh. (2001). Group symmetry in interior-point methods for semidefinite program. *Optimization and Engineering*, 2 (number 3):293 − 320.

[6] M. Kojima, S. Shindoh, and S. Hara. (1997). Interior-point methods for the monotone semidefinite linear complementarity problems. *SIAM J. Optimization*, 7:86–125.

SERVER-SIDE NAVIGATION SERVICE BENCHMARKING TOOL

Martin Golasowski, Kateřina Slaninová, Jiří Ševčík, Vít Ptošek, David Vojtek

IT4Innovations

VŠB - Technical University of Ostrava, Ostrava, Czech Republic

martin.golasowski@vsb.cz, katerina.slaninova@vsb.cz, jiri.sevcik@vsb.cz,
vit.ptosek@vsb.cz, david.vojtek@vsb.cz

Abstract: Convenient access to a complex benchmarking and monitoring tool was the main motivation for implementation of the proposed benchmarking tool. This paper provides a brief summary of its design and implementation. Its main purpose is to test an experimental server-side navigation service which has to comply to a multi-criteria service level agreement. The service is deployed on a heterogenous high performance computing infrastructure which is monitored by the tool. Consistent test environment is provided by the tool for transparent analysis and optimization of the service performance.

Keywords: navigation, routing, benchmark, high performance computing

1   Introduction

There is a significant demand for new forms of vehicle routing algorithms. For example, Probabilistic time-dependent vehicle routing problem which could use precise route computation as an input was proposed in [1]. In the context of smart city, large number of routing requests can be issued in a very short time while there can be periods with very low or nonexistent traffic. Service which provides such functionality has to be flexible enough in order to satisfy the dynamic load while maintaining a given service level agreement (SLA).

These services are usually deployed on a heterogenous architecture consisting for example of both virtual machines and high performance computing (HPC) nodes. The SLA therefore consists of multiple criteria, such as response time or resources budget. These criteria can change over time. Maintaining the SLA over time is possible by using technologies from the ANTAREX project [10], such as the autotuning framework and a domain specific language.

In order to design and develop such a service a proper benchmarking and monitoring tool has to be provided. The goal of the tool is to provide users a way how to test the mentioned multi-criteria SLA against a given set of parameters of the service and its underlying technologies. The tool should provide a consistent and robust benchmarking environment which should be used in a transparent way for optimizing the service. Simultaneously, the tested service is provided as a black box and the tool only exposes only a subset of the service parameters which are relevant for testing. The tool can be used to obtain unified presentation of the test results, which is convenient for comparison of different parameter settings of the service.

The service provides navigation on a graph representation of the road network. It finds a path between a pair of two nodes or between a set of waypoints according to a selected criteria (speed, road/vehicle type, etc.). The performance metric of the service in our case is request throughput per second and average time per request. The input data set consists of a set of origin/destination pairs and parameters of the routing pipeline (PTDR, sample count, routing algorithm type, etc.). The service implements a subset of the HTTP communication protocol and uses simple JSON-based data structures.

1.1   Related work

There is an increased demand for architectures following the Anything as a Service (XaaS) paradigm [4]. Even though there are many robust and complicated protocols for providing such services; APIs based on top of the HTTP protocol are quite popular these days.

Any service provided over a computer network needs to be properly tested in order to determine possible performance bottlenecks in the architecture or possible errors in its output. Obtaining an energy efficient operating mode of the service is also a significant factor in testing, especially in big data centres [2, 3]. Demand for energy efficient adaption of this paradigm on powerful computing nodes is getting more significant with the overlap of cloud and traditional high-performance computing.

There are many benchmarking tools for HTTP-based services implemented for various platforms. They are usually implemented as a stand-alone terminal applications [5, 6]. Input specification is often realised by a template of the request or by a set of pre-defined requests to be sent. Some of the benchmarks provide more comprehensive statistical evaluation of the test, however, the output is often in plain text form, which is harder to parse and prone to errors. On the other hand, there are comprehensive tools, such as Locust [7], which provides all-in-one solution together with the analysis and visualisation front-end and result database. Such tools offer many features, which are unnecessary for our use-case and can be hard to alter and optimise. Therefore, we propose a custom light-weight solution which is designed for services running on heterogenous HPC infrastructure.

Next Section 2 is focused the benchmarking process including an example of the server-side navigation service benchmark used within the ANTAREX project [10]. The benchmark implementation is described in Section 3 and the asynchronous back-end which provides an execution and management of large number of small tasks follows in Section 3.2. Section 4 then concludes the paper.

2 Benchmarking Process

Benchmarking can be described as a process in which performance of a system is evaluated on a set of its parameters and input data. Performance of the system is then evaluated by relevant set of metrics (throughput, load, energy consumption, etc.). Subject of the testing is a relation between the parameter values and the selected metrics. Tested parameters can be for example I/O strategy, memory access pattern or various source code optimizations. The system is tested against a set of the parameter values while input data and run-time environment remain static.

Crucial property of a good benchmark is consistency. Individual tests share the same conditions through the entire test. It is often hard to assure completely consistent conditions for all the tests in real-world scenarios (for example in multi-task operating systems), therefore influence of the external factors should be kept to a sane minimum. For example, unnecessary services should be turned off as well as scheduling of maintenance tasks (backup, logging, etc.), in order to minimise performance fluctuation of the target hardware platform.

2.1 Server-side Navigation Benchmark

In our case, the subject for benchmarking is the server-side navigation service. Core of the service is the navigation pipeline which is deployed on a HPC cluster. This design allows flexible optimization of the pipeline according to various criteria to correspond with the current SLA.

Parameters of the benchmark are divided into three parts.

1. **Data sets** - type, size

2. **Static routing** - routing algorithm, route type

3. **Probabilistic time-dependent routing** - enabled, number of samples, *autotuning strategy*

The data set of origin-destination point pairs is sent to the service from a large number of concurrent processes. This approach allows us to simulate a large number of concurrent users

Figure 1: HPC Routing Benchmark Dataset

of the service. The service then logs each individual requests. A benchmark is defined by the time-frame between sending the first pair of the input data to handling reception of a response for the last pair of the input data. The metrics are computed from the service log within the time-frame. Total amount of the requests sent is controlled by initial size of the selected data set and its size parameter, which can be used to artificially inflate the number of sent requests and to create specialized use case (such as smart city).

2.2 Data Sets

The points in current version of the benchmark are selected from the region of the Czech Republic, see Figure 1. The region is divided to 77 units according to the LAU1 (Local Administrative Units Level 1) division, which corresponds to individual districts of the Czech Republic. Each district has its administrative center which usually corresponds to its biggest city. There are 72 district towns as some of districts share one district town. There are 10 randomly selected points, that correspond to nodes of a routing graph, inside each of the 72 district cities. The resulting point data set is stored in a PostgreSQL database with a spatial extension PostGIS as a basic data set. Source data for the basic dataset are obtained from the Open Street Map project (OSM)[11].

Since the source data for the routing graph are updated regularly, the IDs of the graph nodes (the selected points of the basic dataset), can change. Therefore, it is practical to generate derived data sets only by a database query stored as a procedure or materialized view. The data sets are versioned according to the current processed OSM data.

In the current version of the benchmark, there are three data sets, defined by their size and type of the location:
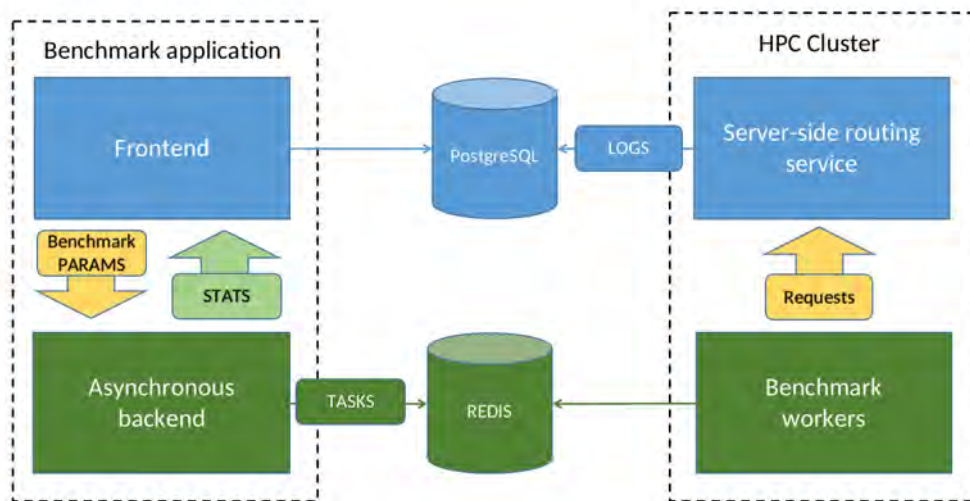
Figure 2: Overview of the application architecture

- **Smart city** - points within single LAU (city, 90 points)

- **Transit** - points between three main cities (Praha, Brno, Ostrava, 300 points)

- **All** - all 72 LAU1 regions (250 000 points)

3   Benchmark Application

The application design is roughly divided into several parts. The monitoring and benchmarking front-end is the first part while the asynchronous task and database back-end is the second part. Overview of the application architecture is in the Figure 2. Web applications are inherently synchronous and stateless which makes them unsuitable for our purposes. The user front-end should only serve as a tool for control and monitoring of the benchmarking process. Therefore, an asynchronous approach has to be implemented, such that the process execution is independent on the user requests to the application front-end.

3.1   Implementation

The frontend is implemented in a popular Python web framework Flask [8]. In the context of benchmarking, it provides an user interface for setup, execution and presentation of the benchmark results. The main page contains a form for submitting the benchmark task and a table of results of the past runs. Detail of a particular benchmark run can be accessed from the results table (Figure 3). The detail presented in Figure 4 contains information about parameters of the viewed run and its results. In this particular case, the presentation of the results is split into data about the run itself and several metrics computed from the service log. Content of the entire page is generated on the server side directly from the service logs. This approach allows fairly easy extension of the analysis with new types of metrics.

3.2   Backend

The asynchronous back-end of the application is implemented with the help of the Celery framework [9]. It provides a convenient way of execution and management of large number of small tasks asynchronously and in a distributed fashion. The client part of the framework is

Figure 3: Interface for parameter specification

Figure 4: Visualisation of the benchmark result

used by the front-end to create the benchmarking task and pass the parameters. The Celery then uses a type of in-memory concurrent storage (such as REDIS or RabbitMQ) to distribute the task across available worker processes. The worker processes communicate with the client through the in-memory storage, which is also used for storing the intermediate results. This mechanism is very useful in our case as it can be used to launch a large number of tasks which are processed by a pool of workers running on a HPC cluster. Celery takes care of the task scheduling itself and the benchmarking pipeline is implemented using the available task synchronization constructs.

The individual service requests are executed in a celery worker tasks. Each request is sent to the service in a blocking call to HTTP requests library and the result of the call is stored in the in-memory database. After all requests are sent, all results are passed to a finalization task, which computes aggregated statistics and stores them in the SQL database for presentation on the front-end.

The worker processes are deployed on a HPC cluster to ensure proper saturation of the service in order to maintain consistency of the benchmark.

4 Conclusion and future work

The authors presented benchmarking application which is used for transparent testing and further optimization of a server-side navigation service. The application implements an initial version of a consistent benchmarking environment. It can be further extended for testing other network-based services. Current version is used for testing of the autotuning framework and other tools developed within the H2020 project ANTAREX as well as for performing acceptance

tests of the navigation service. It will be also used for testing of various types of navigation algorithms for development of vehicle routing problem algorithm.

In the next version, we plan to implement another type of the test, which will bypass the network connection handling layer of the service and test only the routing pipeline running on the cluster.

5 Acknowledgements

References

[1] Režnar, T., Martinovič, J., Slaninová, K., Grakova, E., Vondrák, V.: Probabilistic time-dependent vehicle routing problem (2016). Central European Journal of Operations Research, pp. 1–16.

[2] Yichao Jin and Yonggang Wen and Qinghua Chen: Energy efficiency and server virtualization in data centers: An empirical investigation (2012). In 2012 Proceedings IEEE INFOCOM Workshops, pp. 133–138.

[3] K. Le and R. Bianchini and T. D. Nguyen and O. Bilgir and M. Martonosi: Capping the brown energy consumption of Internet services at low cost (2010). In International Conference on Green Computing, pp. 3–14.

[4] Y. Duan and G. Fu and N. Zhou and X. Sun and N. C. Narendra and B. Hu: Everything as a Service (XaaS) on the Cloud: Origins, Current and Future Trends (2015). In 2015 IEEE 8th International Conference on Cloud Computing, pp. 621–628.

[5] The Apache Benchmark. (2017) The Apache Software Foundation https://httpd.apache.org/docs/2.4/programs/ab.html [Accessed 20/05/2017].

[6] Siege (2012) Jeffrey Fulmer at al. https://www.joedog.org/siege-home/ [Accessed 20/05/2017].

[7] Locust (2016). Locust http://locust.io/ [Accessed 20/05/2017].

[8] Flask (2010). Armin Ronacher http://flask.pocoo.org/ [Accessed 20/05/2017].

[9] Celery - Distributed Task Queue (2017). Ask Solem http://docs.celeryproject.org/en/latest/index.html [Accessed 20/05/2017].

[10] Silvano, C., Agosta, G., Cherubin, S., Gadioli, D., Palermo, G., Bartolini, A., Bispo, J., et. al. (2016). The ANTAREX approach to autotuning and adaptivity for energy efficient HPC systems. In Proceedings of the ACM International Conference on Computing Frontiers (pp. 288–293). ACM.

[11] Haklay, M., Weber, P. (2008). Openstreetmap: User-generated street maps. IEEE Pervasive Computing, 7(4), pp. 12-18.

KNOWLEDGE TRANSFER ONTOLOGY OF PHOTOVOLTAIC ELECTRICITY PRODUCTION FORECASTING

Anja Goričan,
Amadeja Bratuša,
Drago Bokal,
Faculty of Natural Sciences and Mathematics, University of Maribor,
Koroška cesta 160, Maribor, Slovenia,
gorican.a@gmail.com, bratusa.amadeja@gmail.com, d@bokal.net

Abstract: We present the knowledge transfer ontology applicable when forecasting photovoltaic electricity production. The paper serves as an case study for the presentation on Knowledge transfer ontologies introduced at an earlier presentation, and develops a prototype ontology specialized in photovoltaic production. The ontology builds both on surveying existing bibliography on the topic, as well as on experiments with our own benchmark data, and is hence of independent interest to practitioners of photovoltaic production.

Keywords: photovoltaic electricity production, knowledge transfer ontology, forecasting.

1 INTRODUCTION

Electric system stakeholders need to know how much electricity will be produced and consumed in an area that they supervise, so that they can balance production and consumption of electricity. If there is not enough electricity produced, the stakeholders have to resort to primary, secondary or even tertiary reserve services of electricity system, which are expensive means of producing or curtailing electricity production or consumption when market imbalances occur. Because the production and consumption of electricity must be balanced, any purchases made in advance must accurately reflect future consumption as otherwise losses due to forced selling or expensive late purchases may happen. In such system, any participant in electricity market desires to have the most accurate forecast of electricity production as possible. Forecasting photovoltaic electricity production does not always return accurate results, because it depends on several parameters, such as type of photovoltaic panels, placement and angle of inclination of the panels, solar irradiation, temperature. When agregating over large sets of solar power plants, their microclimate plays a sifgnificant role. Forecasting also depends on the choice of forecasting method and forecasting interval. There are also other dimensions that we should consider and will be presented in Sections 3, 4, and 5. Approach to forecasting electrical consumption with different modelling techniques and use of taxonomy is presented in [4]. Use of taxonomy for electrical consumption was also described in [5], where a more detailed prediction using the ARIMA as modelling technique was presented and its comparison to other modelling techniques.

2 KNOWLEDGE TRANSFER ONTOLOGY

Ontologies are models that computer scientists use to express knowledge. The detailed definition of ontology can be found for example in [2]. Our ontology exists of three subontologies:

- ontology of use cases,
- ontology of methodologies,
- ontology of criteria.

Any ontology, whose knowledge base instances are problem sets for the methods that needed to be researched, is a use case ontology. The minimum requirement for use case ontology is that its concept set consists of identifiers for use cases that practitioners conformed in their domain. The identifiers from which concept set is formed can have internal structure. The instances of ontology establish the set of all possible use cases. The taxonomy is actually the hierarchy concept of this ontology. Furthermore, the ontology of use cases can be drawn-out from pure taxonomy to comprise problem instances. The ontology

of use cases can also establish an exchange standard for these instances, which allows practitioners to add data to research.

The ontology of methodologies consists of knowledge base of instances, which are implementations of methodologies used to solve the problems defined by the use cases. The ontology of methodologies needs at least methodologies, which solve problems, as its concept sets. Ontology can be extended with interfaces between implementations and data. This extension enables any new methodology to be tested on existing data or enables new data to be tested on existing methodologies. It can be done by respecting the extended interface ontology elements. We can apply any methodology to any use case, so we need to determine the Cartesian product of the methodology knowledge base's instance set and the use case knowledge base's instance set. It forms a formal space describing the experiments of applying all instances of methodologies to all instances of the use cases. More details are described in [3].

For specific use case instances, different methodologies perform better or worse, which can be evaluated using some criteria. Whenever selecting a methodology to a specific use case or use case to a specific methodology, we are faced with many different choices. Our ontology comprises one key information, which consist of the pairs of use cases and methodologies. They can be used to label papers that study the corresponding applications. The set of all used labels containing an use case would then correspond to all the methods that have been applied to that use case, so the practitioner would be able to select the most appropriate one among them. As papers compare the methods on the use cases (for instance according to accuracy of models or efficiency of computation), they define an implicit direct graph, whose vertices are used as labels and whose edges are directed from worse methodologies to better ones. Directed edges from worse to better methodologies can be labelled by using instances from ontology of criteria. These are different errors for comparing methodologies in photovoltaic electricity production and will be presented in Section 5.

3   PHOTOVOLTAIC PRODUCTION FORECASTING USE CASES

The instances of the ontology of use cases are data sets on which forecasting of photovoltaic electricity production is performed. If these data sets have the same internal structure, they belong to the same use case. These data sets can be described using several parameters, such as the **forecasting group**, **voltage level**, **maturity of forecasting**, **the delay between the last measurement and the first forecast**, **the service that needs forecasting**. We explicitly refer to these dimensions as use case parameters. They are usually constant within one application and within one research paper, but they tend to be different among different applications.

First use case parameter of the ontology of use cases is **forecasting group**, which describes the level of electricity infrastructure we want to forecast. Forecasting can be done for an individual solar power plant or all solar power plants that are connected to the same transformer station subgrid, the same distribution transformer station, or the same distribution network operator subgrid. It can also be done for a region, a country, a multinational region or even for all world.

The next use case parameter we consider is the **maturity of forecasting**. By that we determine for how many hours in advance we want to predict photovoltaic production. We can do ultra short term forecasting, which is forecasting from zero to twelve hours in advance or short term forecasting, which is forecasting from twelve to one hundred and sixty-eight hours (one week) in advance. Another option is medium term forecasting (seven to thirty-one days ahead). The last option is the long term forecasting, which is forecasting over one month in advance.

When forecasting, we need to consider the **delay between the last measurement and the first forecast** in electricity production data. It is called forecasting lag and is generated due to the process of transmitting and collecting data required by forecasting. There can be no lag between last measurement and first forecast or can be less than ten intervals (one interval lasts between two measurements; we call its length data resolution) or can be tens of intervals or finally there can be hundreds of intervals in electricity production data.

Furthermore, it is important to know for what purpose we forecast. Forecasting can be done for the purpose of balancing the production and the consumption or for the purpose of planning electricity production or for optimizing market or for research.

4   ELECTRICITY PRODUCTION FORECASTING METHODOLOGIES

We continue by describing the parameters defining the methodologies used for forecasting. First parameter is the list of time-series used as predictors in the process. The predictors are parameters that are related to electricity production or electricity production directly depends on them. **Model inputs** include time-series of weather data such as temperature, solar irradiation, and wind speed. All used data is a time-series, so it is suitable to add the hour of day and day period (daylight, dusk, night).

The next parameter that determines the methodology is **resolution of the input data**. It can be few minutes, hourly, daily, weekly, monthly, or yearly. If we have different resolutions of time series within the same use case instance, we must unify them either by interpolation (refining the coarser time series) or by discarding the intermediate data of more detailed time series so that we predict for the intervals of the coarsest time series.

According to available data, we also need to determine the **length of the training period**. We can choose training period to be one month long or one to two months long or to be two to six months long or to be six to twelve months long or one to two years long or two to four years long or four to six years long or six to eight years long or at last more than eight years long training period. It is also possible that the description of the methodology does not specify the length of the training period.

When setting up a model, we need to decide which **modelling technique** to use. Modelling techniques differ by time used to compute model and to return forecast. Most common modelling techniques for forecasting electricity production are: artificial neural networks, autoregressive models, seasonal autoregressive model, periodical autoregressive models, multiple linear regression, support vector machines, principal component analysis based methodologies and knowledge-base models.

Next methodology parameter we need to determine is **criterion optimized by the model fitting** that serves as a basis for verifying the performance of the model. For this parameter, we can choose mean absolute error or mean relative error or symmetrical relative mean error, unexplained variance, coefficient of determination R^2, AIC (Akaike information criterion), which is a measure of the relative quality of statistical models, BIC (Bayesian information criterion), which is also a measure for model selection, or the cost of errors.

Final two parameters defining the specifics of the methodology are **software implementation** and **hardware platform**. Both use cases influence efficiency of forecasting. Software can choose between R, KNIME, RapidMiner, Matlab and individual implementations. Most common hardware options are personal computer, the others are virtual computer, GPU, and high-performance computers.

5   CRITERIA COMPARING INSTANCES OF PHOTOVOLTAIC PRODUCTION

When comparing different modelling techniques, we need to describe criteria used to compare instances to each other. One parameter of the criteria comparing instances is **model comparison criterion**. This parameter can either be equal to methodology's criterion of model optimization, but they can also differ. Besides the options used there, we can compare methodologies using criteria used for model fitting, but there can also be speed of model learning, speed of forecasting, maturity of the technology, total cost of ownership and other criteria that are less scientific and more relevant to practitioners. Second parameter of comparing instances is **difference**, which measures difference between different criteria evaluation. It can be:

- high, which means that difference value is more than one percentage point,

- medium, which means that difference value is between 0.5 and 1 percentage point,

- low, which means that difference is between 0.1 and 0.5 percentage point,

- negligible, which means that difference is less than 0.1 percentage point.

6   EXPERIMENTAL RESULTS

Within a project "Napovedovanje proizvodnje sončnih elektrarn", which was funded through the grant "Po kreativni poti do praktičnega znanja", some analyses of papers on forecasting photovoltaic production has been done. All comparisons of methodologies and use cases that were found in papers were inserted into a taxonomy, which is described in Sections 3, 4 and 5, and it is presented in Figure 1.
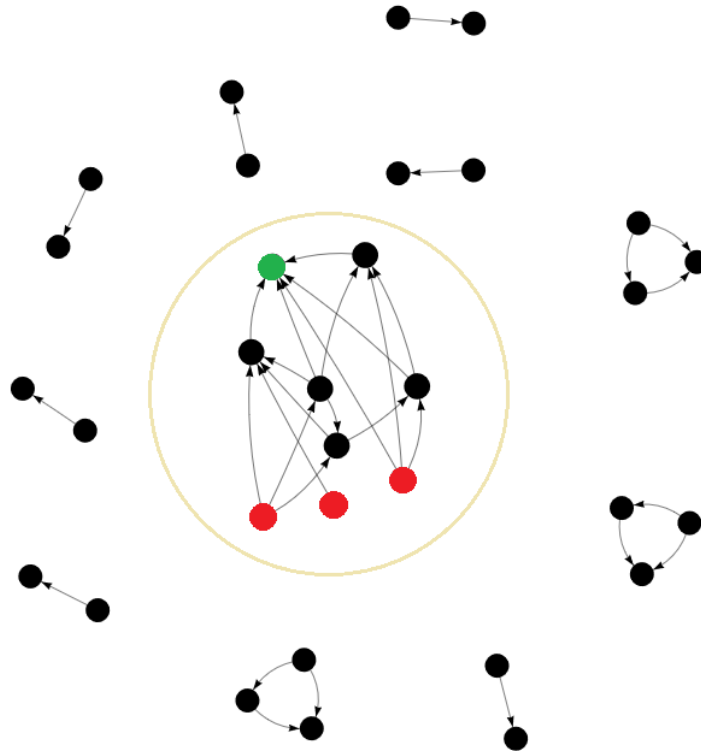
Figure 1: Figure of comparing different use cases and modelling techniques by comparison criteria.

The circled graph in Figure 1 is a comparison of two different methodologies, five model inputs all on the same use case. This comparison was done within a project, which was previously mentioned. Methodology parameters are present with next table.

Business case	Used
Forecasting quantity:	Solar electricity production
Forecasting group:	Regional
Forecasting voltage level:	Low
Forecasting interval:	Very short term forecasting
Lags between last measurement and first forecast:	No lags
Methodology	**Used**
Handling of special days:	None
Handling of weekends:	None
Length of training period:	1,95 years
Criterion optimized by the model fitting:	MAPE
Modelling technique:	MLR, SVM
Model inputs:	Production of the past day, Temperature, Solar irradiation, Physical forecasts of production, Real-time production (photovoltaic production with a delay of 15 minutes)
Resolution of input data:	30 minutes
Software implementation:	R
Hardware platform:	PC
Comparison criteria	**Used**
Model comparison criteria:	MAPE

The top-most dot with all-in arrows in Figure 1 represents the best combination of modelling technique and model inputs for the shown use case: supported vector machine (SVM) technique and model input parameters are a time-series of photovoltaic production of the past day, temperature, solar irradiation and physical forecasts of photovoltaic production. This combination has small MAPE error, which amounted to only 6.7427%.

Further, the three bottom-most dots with all-out arrows represents the worst modelling techniques and model inputs for the shown use case, as all edges are pointed away from them. Their parameters are represented in the next table.

	1. combination	2. combination	3. combination
Modeling technique	MLR	SVM	MLR
Model inputs	Real-time production, production of the past day, temperature, solar irradiation	Physical forecasts of production, production for the past day	Physical forecasts of production, production of the past day, temperature, solar irradiation
MAPE	45.3849%	31.1384%	13.674%

As shown in the table, these combinations have significantly bigger MAPE errors than combination whose MAPE error is only 6.7427%.

7 DISCUSSION AND CONCLUSION

We presented an ontology of use-cases and methodologies for photovoltaic electricity production that encompasses several studied papers as well as our own experiments with forecasting. This approach to represent published research data has several advantages for practitioners, most notably the ease of selecting the most suitable methodology for their specific use case. Researchers benefit from clarify of Pareto frontier of state-of-the-art methodologies that cannot be improved in a relevant criterion without worsening performance in another, as well as more easily accessible benchmark data adhering to standard format that can be defined with the ontology. Students benefit from the structure in which the knowledge is represented and made available to them, as well as from lowered bar for their own contribution to the meaningful work: an easy contribution comes from comprehending papers and contributing their content to the ontology, and a deeper contribution is to confirm the existing comparisons as an exercise in forecasting: both gentle introductions to doing their own research on applying existing methodologies on new use cases or even developing new methods. The psyhological benefits of such approach are discussed in [1]. Discussing our experience with the taxonomy, we observe each investigated paper studied its own use case or made no comparison to a standard benchmark, which resulted in many components of the comparison graph, which is shown in Figure 1. Most components compare only two methodologies, some three, which is interesting enough for a researcher, but not the level of detail a practical application would require. Hence we added a set of comparisons that shows the potential of these ontologies by comparing several modelling techniques with different sets of predictors. When forecasting, it is very likely for practitioners to have a unique use case, so they will have do their own structured analyses of forecasting according to similar use cases to see which modelling technique will perform better.

8 Acknowledgements

References

[1] Bokal D. (2017). Optimizing enjoyment of mathematics and OR education with introducing psychological concepts ow and grit using simulation-based model of emotional states of learning. Submitted to SOR 2017.

[2] Ehrig Marc. 2007. Ontology Alignment: Bridging the Semantic Gap. New York: Springer Science+Business Media.

[3] Goričan A. and Bokal D. (2017). Operations research as the bridge over technological valley of death. Submitted to SOR 2017.

[4] Šmigoc S. 2016. Primerjava pristopov k napovedovanju porabe električne energije. University of Maribor, Faculty of Natural Sciences and Mathematics. 57593. Maribor.

[5] Tajnik M. 2016. Integrirani avtoregresijski modeli s premikajočimi sredinami za napovedovanje porabe električne energije. University of Maribor, Faculty of Natural Sciences and Mathematics. 64888. Maribor.

[6] Tajnik M., Bokal D. (2017). Taksonomija primerjav tehnik za napovedovanje porabe električne energije. 13. konference CIGRE.

COMPARISON OF PARALLEL VERSIONS OF ALNS, ACO AND BRANCH AND CUT ALGORITHMS FOR VEHICLE ROUTING PROBLEM

Ekaterina Grakova, Radim Sojka, Jan Martinovič, Kateřina Slaninová, Jan Vargovský
IT4Innovations,
VŠB - Technical University of Ostrava, Ostrava, Czech Republic
ekaterina.grakova@vsb.cz, radim.sojka@vsb.cz, jan.martinovic@vsb.cz, katerina.slaninova@vsb.cz,
jan.vargovsky@vsb.cz

Abstract: Transportation companies use optimization algorithms to optimize their deliveries. Such daily planning requires large computing capacity. Fast progress of high performance computers allows to solve more complex problems and with better HPC effectivity, the cost of computations decreases. In this paper, we use our set of benchmarks for the experiments which show results of heuristic, metaheuristic and exact algorithm for solving Capacitated Vehicle Routing Problem. The experiments provide a comparison of algorithms in terms of the ratio between a cost of the computation (in sense of required time) and a quality of the reached solution. Although the exact algorithms have a limit in use in terms of size of instances, our results demonstrate that using exact method can be advantageous for solving smaller instances. The algorithms were run on the supercomputer Salomon.

Keywords: vehicle routing problem, algorithms, heuristic, capacitated, customers, vehicle.

1 INTRODUCTION

Daily planning for the dispatches is a common job for the transport companies. They use optimization algorithms to plan every day routes. The optimization algorithms can be divided into three groups: exact, heuristic, and metaheuristic algorithms. The algorithms are given by a formulation of the problem and a complexity of the task. Exact algorithms are highly computationally demanding especially for larger instances, but at the same time they give us an optimal solution to the problem. Heuristic and metaheuristic algorithms do not give us the optimal solution but they can solve computationally demanding tasks. Authors in [3, 5, 8] presented the comparison of the results between the optimization algorithms. It is needed high performance computing (HPC) to solve the large tasks, and to solve an optimal solution. In a supercomputer center, we can solve complicated optimization problems and use optimization algorithms to find the optimal solution. The algorithms were run on the supercomputer Salomon. The aim of this article is to compare heuristic, metaheuristic and exact approaches for solving Capacitated Vehicle Routing Problem (CVRP) for set of benchmarks[1] with less than hundred customers. For these benchmarks, the article tries to determine whether it can be beneficial to use a more computationally complex algorithm to get exact solution or to use a heuristic algorithm giving not necessarily optimal solution.

One of the well-known examples of combinatorial optimization is Traveling Salesman Problem (TSP). The point for the salesman is to visit as many places as possible and return to the point of the origin with a minimal expense and in the shortest time possible. This method becomes popular thanks to its simple formulation and wide usage. But sometimes it is quite complicated. In general, vehicle routing problem (VRP) is the TSP in which a group of vehicles starts in depot, services customers, and returns back to the depot. The objective is to plan the shortest way and a minimal number of vehicles used for the given task. There are some limitations for the Vehicle Routing Problem [11]. The structure of VRP problem is most of the time defined with oriented and not oriented graph, whose nodes represent the location of the

---

[1] Benchmarks used in this article can be found at https://code.it4i.cz/ADAS/CVRP_Benchmark.git.

customers and depots of vehicles, and arcs the routes between the nodes. Each arc has given a value, which in general represents the distance. Most common problem is considering the time windows in which the customer must be served. In reality, every vehicle has a limited capacity or must serve the customer at a given time. Consequently, the more limitations there are the more complicated the problem is.

2 CAPACITATED VEHICLE ROUTING PROBLEM

The distribution of goods from single depot is arranged by a certain number of vehicles with a given capacity. The capacity of each vehicle can be either the same or different. We also know the distance between the customers and the distance from the depot, and the needs of each individual customer. The goal of this problem is to create individual routes for the vehicles, in order not to overestimate the capacity of the given vehicle, and at the same time to serve all the customers with minimal expenses.

2.1 The description of the CVRP

We have a complete undirected graph $G = (V; E)$, where $V = (v_1, ..., v_n)$ represents a set of customers. The location of the customers is given by a node v_i, where $i = 1, ..., n$ and the location of the depot is represented by a node v_0. An arc between both nodes v_i and v_j is equal to the value of c_{ij} (travel cost), which is the distance the given vehicle must travel in order to move from a customer at the node v_i to a customer at the node v_j [12]. The amount that has to be delivered to the customer $i \in N$ is the customer's demand. The fleet K is assumed to be homogeneous, meaning that K vehicles are available at the depot, and all have the same capacity $Q > 0$ [1]. Every vehicle leaving the depot must come back to the depot. Each vehicle can make only one trip, and every customer must be served on this route. The objective of the problem is to suggest the shortest possible routes.

The mathematical formulation of the problem was introduced by Laporte, Nobert, and Desrochers [1]:

$$minimize\ c^T x, \tag{1}$$
$$x(\delta(i)) = 2 \ \forall\ i \ \in N, \tag{2}$$
$$x(\delta(0)) = 2|K| \ , \tag{3}$$
$$x(\delta(S)) \geq 2r(S) \ \forall\ S \ \subseteq N, S \neq \emptyset, \tag{4}$$
$$x_e \in \{0, 1, 2\} \ \forall\ e \in \ \delta(0), \ x_e \in \{0,1\} \ \forall\ e \in E \setminus \delta(0). \tag{5}$$

Where: S - customers subset $S \subseteq\ N$,
                 $r\ (S)$ - minimum number of vehicles needed to serve S,
$\delta(S)$ - cut-set defined by S.
The objective (1) is the minimization of the overall routing costs. Constraints (2) state that in a route, each customer vertex is connected to two other vertices which are its predecessor and successor. Similarly, constraints (3) ensure that exactly $|K|$ routes are constructed. Constraints (4) serve at the same time as capacity constraints and subtour elimination constraints [1].

3 ALGORITHMS

There are many algorithms to solve VRP. In the last few years authors started to divide the algorithms to exact algorithms, heuristic algorithms, and metaheuristic algorithms. The exact algorithms are typical mathematical methods researching the whole state space in order to find a general (global) optimum. If the number of customers increases the state space will grow

exponentially, and consequently, the problem will become more complicated [13]. Nevertheless, with the fewer number of the customers we can use these algorithms and so find the real optimum. The heuristic algorithms as oppose to the exact algorithms, search only through small part of the state space, overlook optimal solution. However, sometimes it can offer a suboptimal solution, and much faster. The most of the time, the heuristics were used for specifically modified problem or in combination with metaheuristic. The last group of the algorithms are metaheuristic algorithms, which are very popular, because they can quickly find a solution to the global problem. Metaheuristic algorithms are used for time demanding problems which have many limited conditions or have a large set of possible solutions.

In this paper, we used metaheuristic algorithm - Ant Colony Algorithm, heuristic algorithm - Adaptive Large Neighborhood Search and exact VRP algorithm - Branch and Cut (SYMPHONY library) [14, 15].

3.1 Ant Colony Optimization Algorithm

Ant Colony Optimization algorithm (ACO) belongs to a larger field of study called Swarm Intelligence (SI) and Bio-Inspired Computation [2,3]. The SI algorithms are inspired by real life behaviour of multi-agent systems organised for instance by ants, bats, bees or cuckoos. The ACO is a meta heuristic technique based on the observations of the behaviour of real ant colonies looking for food. An ant's eyesight is mediocre, some of them are even completely blind, nevertheless they are still able to find food. When ants are looking for food, they just go straight forward till they find some. To find a path back, they lay down a substance called pheromone. The amount of pheromone used in a path determines how good or how bad the path is. The more amount of pheromone that is laid, the more promising the path looks. This observation was used as an inspiration by Dorigo, Maniezzo and Colorni in 1991 [4,5]. What an ant does, is reducing the distance of the path from the anthill to food and vice versa. And that is exactly what we are trying to do in the optimization problems. As opposed to other algorithms, in this algorithm we have set the maximum and the minimum concentration pheromone [8]. In [7], Gutjahr suggested a graph based ant system. The solution is represented as a way in oriented graph. Authors in [8] suggested max-min ant system. Currently, there are many variations of the ant colony algorithms, which differ mainly in number and the leaved pheromone behind them. Pseudocode of the basic algorithm is presented in [4,5]. Consequently, the algorithm was modified for CVRP. Local Search can be executed in parallel by using C# threads. Table 2 describes what every parameter in code means and its equivalent name in theory, default value and short description.

Table 2: Configuration parameters description

Code name	Theory name	Brief description
AntCount	p	Number of ants
Iterations	it	Number of iterations
Alpha	α	Importance of pheromone
Beta	β	Importance of heuristic information
P	r	Evaporation rate
Q	Q	Constant to adjust laid pheromone
Min	τ_{min}	Minimum pheromone
Max	τ_{max}	Maximum pheromone
InitialPheromone	τ_0	Initial pheromone value
CandidateFraction	cf	Candidate fraction
Q0	q_0	Exploitation versus exploration

3.2 Adaptive Large Neighborhood Search

The first version of the Adaptive Large Neighbourhood Search (ALNS) algorithm was suggested by Pisinger and Ropke to solve various VRP [9]. It belongs to the family of very large neighbourhood search heuristic. In each iteration, a destroy neighbourhood is chosen to destroy the current solution, and a repaired neighbourhood is chosen to repair the solution. The new solution is accepted if it satisfies some criteria defined by the local search framework applied at the master level. They are exceptional, because they are searching for large neighbourhood partial solution and thanks to that they are choosing from many options and they find the quality results much faster. The authors in [9] implemented and adjusted the algorithm for the modification of the VRP problem with time window. The authors introduced better results than in original version in [11]. The authors in [6] proposed an adaptive large neighborhood search heuristic for the Two-Echelon Vehicle Routing Problem (2E-VRP) and the Location Routing Problem (LRP). 2E-VRP arises in two-level transportation systems such as those encountered in the context of city logistics. In such systems, freight arrives at a major terminal and is shipped through intermediate satellite facilities to the final customers [6]. Pseudocode of the basic algorithm is presented in [10].

3.3 Branch and Cut

For exact solution of CVRP, the VRP application included in SYMPHONY solver developed by Ted Ralph was used. The packages implement parallel branch and cut algorithm. Branch and Cut method modifies LP-based Branch and Bound method by using cutting planes to tighten the LP relaxation of original integer problem. Open-source solver SYMPHONY can be executed in parallel using distributed (PVM) or shared (OpenMP) memory parallelism. Detailed description of parallel implementation of Branch and Cut and SYMPHONY solver can be found in [14, 15]

4 EXPERIMENTAL RESULTS

Testing sets were created for testing the algorithms. The instances of this set contained 25 to 101 customers and 4 to 10 vehicles. Every customer had specific GPS location in Ostrava. Capacity of pick-up at the station was defined for each customer; each vehicle was determined by the same capacity. For each stop was defined the time delay at the station. Also, metric distances between the customers were created on the basis of real data. Table 3 shows the results from the exact, heuristic and metaheuristic algorithms.

   Experiments were performed on Salomon supercomputer where each node has two – twelve – core Intel Xeon processors and 128 GB RAM (2xIntel Xeon E5 – 2680 v3, 2.5 GHz, 12 cores). The exact method (Branch and Cut) results were computed by the SYMPHONY library [14, 15]. In the two options OVn80-k7 and OV-n101-k10 we could not find the optimal solutions. Therefore, in Table 3 there are introduced results of the lower bound, which indicates how could introduced result improve. Time of experiments with ACO algorithms was influenced by the way the experiments were executed. For turning on the ACO algorithm we used DotNet Core 1.1.0/Singularity 2.2 developer, because we were unable to install .NET CORE on CENToS 6.9. Singularity 2.2 dramatically increased the complexity of the result; therefore, the results are worse in that case. We used MONO 5.0.0 program for ALNS algorithm.

Table 3: Result from the heuristic algorithm a metaheuristic algorithm

Instance	Branch and Cut			ACO			ALNS		Gap [%]	
	Distance [m]		T [s]	Distance [m]		T [s]	Distance [m]	T [s]	ACO [%]	ALNS [%]
		Lower Bound								
OV-n25-k4	63093,8		0,1	$p = 16$, $it = 500$, $\alpha = 1$, $\beta = 2$	**63093,8**	69,2	64365,0	2,30	**0,00**	1,98
OV-n40-k6	82666,2		0,2	$p = 256$, $it = 300$, $\alpha = 3$, $\beta = 1$	83122,7	116,3	**83106,0**	44,60	0,54	0,53
OV-n50-k6	90411,7		0,3	$p = 128$, $it = 500$, $\alpha = 1$, $\beta = 2$	92440,8	148,9	94480,9	65,25	2,19	4,31
OV-n60-k6	96655,9		1,1	$p = 128$, $it = 500$, $\alpha = 1$, $\beta = 2$	**103562,3**	138,9	109402,9	46,10	0,07	11,65
OV-n71-k7	102882,6		57,2	$p = 64$, $it = 300$, $\alpha = 1$, $\beta = 3$	113174,7	70,72	104774,9	49,30	10,25	2,50
OV-n75-k8	109341,5		5,4	$p = 32$, $it = 300$, $\alpha = 1$, $\beta = 2$	**110024,3**	72,88	125992,0	69,44	0,62	13,21
OV-n80-k7	116156,3	114132,6	> 93427	$p = 128$, $it = 300$, $\alpha = 2$, $\beta = 3$	**123967,2**	123,8	124126,0	79,36	6,3	6,42
OV-n90-k8	122105,3		9,4	$p = 32$, $it = 500$, $\alpha = 1$, $\beta = 2$	**133706,5**	117,8	152016,0	39,68	8,6	19,68
OV-n101-k10	153265,9	143223,0	> 293084	$p = 128$, $it = 300$, $\alpha = 3$, $\beta = 1$	163318,4	120,3	158152,0	50,00	6,16	3,09

5 CONCLUSION

In this paper, we compared the results of exact, heuristic, and metaheuristic algorithm. With the introduced heuristic algorithm, we found the solution which is very close to the optimum. Percentage quality results in the heuristics and metaheuristic gaps are presented in Table 3. We can see from the results that runtime of the exact algorithm increases with the increasing size of the benchmark. The optimal solution for two instances was not achieved, as opposed to heuristic and metaheuristic algorithm where size of the tasks is not issue but the quality increase.

However, the computational time and cost of the exact solutions of most benchmarks are competitive with runtime for getting non-optimal solutions from heuristic and metaheuristic algorithms. These preliminary results indicate that using exact algorithms is beneficial for solving smaller instances because they obtain better solutions with similar time used by the heuristic and metaheuristic algorithm. In the future work, it is desirable to focus on more detailed comparison taking into account the price of core-hours used for the computation and CVRP solution. The experiments should be performed for more instances, especially for larger ones. Also, the comparison of the exact method with time-limit and (meta)heuristic approaches for larger benchmark will be carried out.

References

[1] Laporte, G., Nobert, Y., Desrochers, M. 1985. Optimal routing under capacity and distance restrictions. Operations Research, Vol(33): 1050-1073.

[2] Yang, X., Karamanoglu, M. 2013. 1 – Swarm intelligence and bio-inspired computation: An Overview. In: Swarm intelligence and bio-inspired computation (pp. 2-23). Oxford: Elsevier.

[3] Moghadam, B.F., Sagjadi, J.S., Seyedhosseini, S.M. 2010. Comparing mathematical and heuristic methods for robust vehicle routing problem. IJRRAS. Vol(2).

[4] Dorigo, M., Maniezzo, V., Colorni, A. 1991. The ant system: An autocatalytic optimizing proces. Technical repor. Vol(91-016).

[5] Fard, F.A., Setak, M. 2011. Comparison between two algorithms for multi-depot vehicle routing problem with inventory transfer between Depots in a Three-Echelon Supply Chain. Vol(28): 0975-8887.

[6] Hemmelmayr, V., C., Cordeau, J., Crainic, T., G. 2012. An adaptive large neighborhood search heuristic for two-echelon vehicle routing problem arising in city logistics. Vol(39): 3215-3228.

[7] Gutjahr, W. A. 2000. Graph-based ant systém and its convergence. Vol(16): 873-888.

[8] Toth, P., Vigo, D. 2014. Vehicle routing problem, problems, methods and applications. 2nd ed. MOS-SlAM series on optimization, SIAM.

[9] Pisinger, D., Ropke, S. 2010. A general heuristic for vehicle routing problems. Vol(34): 399-419.

[10] Režnar, T., Martinovič, J., Slaninová, K., Grakova, E., Vondrák, V. 2016. Probabilistic time-dependent vehicle routing problem. Springer-Verlag. 1-16.

[11] Dantzig, G., Ramser, R.: 1959. The truck dispatching problem. Management Science, Vol(6): 80-91.

[12] Contardo C, Martinelli R. 2014. A new exact algorithm for the multi-depot vehicle routing problem under capacity and route length constraints. Discrete Optim. 12:129–146.

[13] Christofides, N., Mingozzi, A., Toth, P. 1981. Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations. Mathematical programming, Vol(33): 1050-1-73.

[14] Ralphs, T.K. 2003. Parallel branch and cut for capacitated vehicle routing. Parallel Computing, Vol(29): 607-629.

[15] Coin-or/SYMPHONY: Version 5.6.16. 2017. Zenoob. https://zenodo.org/record/248734/export/hx#.WUvF-lEzWUl [17.01.2017].

AUTONOMOUS ON-LINE OUTLIER DETECTION FRAMEWORK FOR STREAMING SENSOR DATA

Klemen Kenda, Dunja Mladenić
Jožef Stefan Institute, Jamova ulica 39, 1000 Ljubljana, Slovenia
Jožef Stefan International Postgraduate School, Jamova ulica 39, 1000 Ljubljana

Abstract: This paper presents a real-world application for stream preprocessing, focusing on data cleaning of sensor data. Although stream mining and Big Data have been hot topics for a relatively long period of time, very little attention has been dedicated to data preprocessing in real-world scenarios. The described approach takes advantage of Kalman filters short-term prediction capabilities. The model can adapt to concept drift and data dynamics and is useful for detecting random additive outliers in a sensor data stream. The experimental evaluations show that in datasets with lower noise the proposed method gives better or comparable results to the commonly used method.

Keywords: on-line framework, real world system, data cleaning, data preprocessing

1   INTRODUCTION

Big Data is a term that is used for datasets that are too large in size and complexity to be handled with the current methodologies [5]. The meaning of this definition changes constantly with the development of technology and advances in computer science, however, translating the analysis into a streaming on-line process is always considered as a good approach to handle Big Data.

The field has received a lot of attention. Many stream modeling (regression, classification, clustering etc.) and evaluation methods have been developed, however, some data mining process phases as identified in CRISP-DM methodology [12] and depicted in Figure 1, have been left aside [7, 10]. Those are crucial for real-world applications.

Even in classical data mining task, where all the data is available beforehand, the researchers claim that data preparation takes up to 80% of the time. A lot of work is done manually. In stream mining scenario, there is no possibility for a constant human intervention, all the data preprocessing needs to be completely automated and autonomous.

Data cleaning represents the first step in data preprocessing. It represents a permanent challenge in data analytics. Not or badly performed can result in inaccurate predictions or unreliable decisions. The issue has

Figure 1: Cross-Industry Standard Process for Data Mining (CRISP-DM) [12].

been attacked recently both, by industry and academia, mostly to address the issues of scalability (Big Data), interfaces, new abstractions and statistical techniques [3].

Our framework is proposed to solve a data-cleaning sub-problem (detection of additive outliers on a dense sensor data stream) with a fast and easily scalable methodology.

1.1   Related Work

The field of time-series analysis has been lively for a number of decades. Kalman (Thiele (1880), Swerling (1958) and Stratonovich (1959, 1960) [15] before him) published his work on linear filtering already in 1960 [6]. Kalman stands out of the *crowd* due to the successful application of the equations to trajectory estimation in the NASA Apollo space program. Different applications have been reported since then and the field of time-series analysis has been *reinvented* in correspondence with advances in computer science and technology. In last years many applications were created for on-line streaming data analysis.

Outlier detection in time series using short term prediction models has been thoroughly discussed already in 1993 by Chen and Liu [2]. The paper identifies 5 different type of time series outliers: (1) Additive Outlier (AO), (2) Innovation Outlier (IO), (3) Level Shift (LS), (4) Temporary Change (TC) and (5) Seasona Level Shift (SLS). Authors propose usage of different models from ARIMA family (AR, MA, IMA, Seasonal IMA) for outlier detection, using its short-term prediction capabilities.

To the best of our knowledge the usage of Kalman filter for cleaning of streaming sensor data has firstly been proposed in [8]. This paper, on which we build this work, presents a methodology for data cleaning of sensor data using the Kalman filter. The Kalman filter is an on-line algorithm and as such is ideal for usage on the sensor data streams. The Kalman filter learns parameters of a user-specified underlying model which models the phenomena the sensor is measuring. Authors propose first and second degree polynomial models. Usage of the Kalman filter is proposed to predict the expected values of the measuring process in the near future. The very same idea has been proposed in [14], where it has been studied in depth in a wider context and extended. The authors coin the methodology as time series Kalman filter (TSKF).

Recently literature is examining potential Kalman filter extensions for data cleaning. For example Marczak et al. [11] study usability of augmented Kalman Filter (AKF).

2   ADDITIVE OUTLIER DETECTION USING KALMAN FILTER

Additive outlier is a point outlier which occurs at a given timestamp t_j and affects a single observation. In sensor data such outliers can be a consequence of a sudden change of ambiental conditions, communication glitch or some similar unexpected event. With sensor measurements we assume that they arrive much faster than the data changes.

We propose a method with short-term prediction, based on previous measurements. Short term prediction is compared to the new measurement and classified as an outlier if the difference exceeds a specified threshold. As proposed in [8] we introduce a safe guard to overcome a potential instability of the algorithm and enlarge the threshold in case that detected outlier is a false positive, which might be an indication of sudden concept drift in the data.

2.1   Kalman Filter

Kalman filter is an ideal algorithm to be applied for data cleaning in a streaming scenario. It is an on-line algorithm, that can produce short term predictions and even calculate covariance error matrix (used as a threshold for outlier classification). Algorithm assumes that our process can be described as Gauss-Markov process (see Figure 2), which has two properties:

(1) Every consequent internal state θ_{j+1} only depends on a prior internal state θ_j. Both states are connected through transition matrix Φ_j. (2) Each internal state θ_j can be observed through its observation x_j, which is connected to the internal state via observation matrix H_j and a subject of Gaussian noise.

In general, matrices H_j and Φ_j can change over time, but in our case they remain the same. Kalman filter equations are depicted in Figure 3.
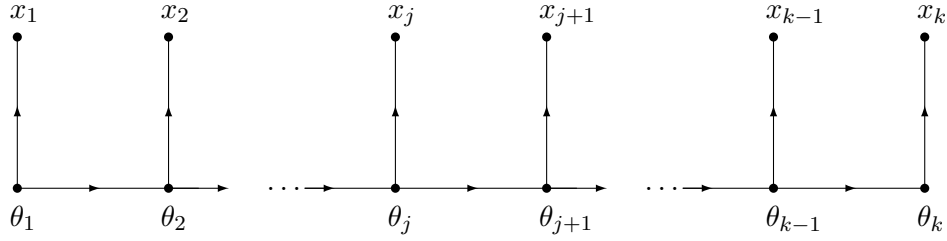
Figure 2: Diagram of a Gauss-Markov process. Arrows from internal state θ_j to another internal state θ_{j+1} depict transitions (transition equation) and arrows from internal state θ_j to observation x_j depict observation equation. Each state only depends on the previous.
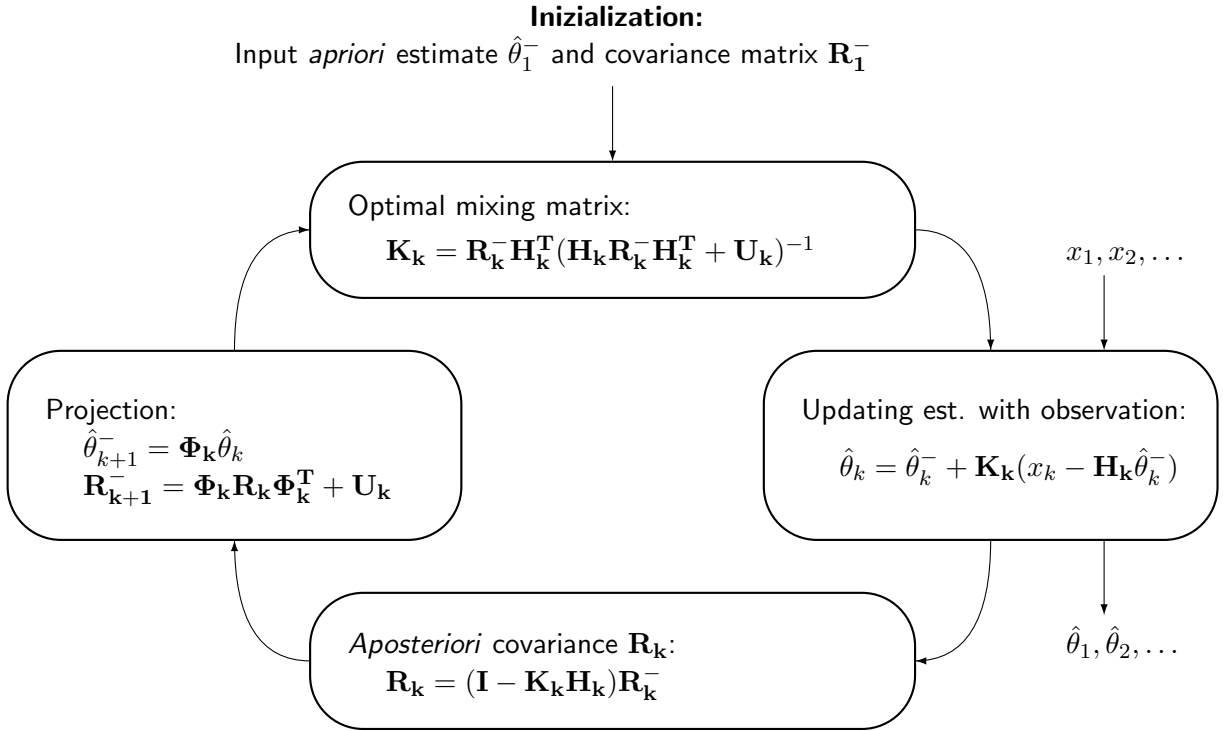
Figure 3: Kalman filter application cycle starts with initialization of *apriori* estimates for internal state θ_1^- and covariance matrix $\mathbf{R_1^-}$. With each new observation x_j the state and covariance matrix get updated. The next phase is dedicated to short-term one step ahead prediction (projection). Finally optimal new mixing matrix gets calculated (responsible for optimal updating of the projected state with an observation). $\mathbf{U_k}$ represents normal distribution variance noise matrix.

2.2 Parameter Learning

Initialization of Kalman filtering algorithm can be very demanding and there can be many free parameters involved, depending on the observation and transition matrix dimensions. Usage of expectation maximization (EM) algorithm [4, 14] can yield estimates for the initial internal state of the system and corresponding covariance matrices. Clean initial dataset is needed to obtain these parameters.

105

In our experiments with time series data the results from EM algorithm only have not provided good results (confidence into last state was exaggerated), therefore we propose an additional data-oriented approach. EM calculates estimates of the following parameters: apriori initial state θ_1^-, transition covariance $\mathbf{Q}$, observation covariance $\mathbf{R_k}$ and initial state covariance $\mathbf{R_1^-}$. We propose to do a grid search over multipliers of these matrices, optimizing F_1 score of the Kalman AO classification algorithm.

Grid search is time consuming, but it can find model configurations which result in much smoother model that follows the underlying processes in dynamics of the data better.

3 STREAMING SENSOR DATA PLATFORM WITH DATA CLEANING

We propose the usage of the filter at the lowest possible level in the preprocessing platform. The data-cleaning component should be implemented at the entry point of a particular data source to the preprocessing platform (see Figure 4). Clean data is then inserted into stream pre-processing engine, which is in charge of data enrichment and heterogeneous data fusion and finally this data is pushed into appropriate stream modeling methods. Cleaning at this level uses only autoregressive features. On a higher level, however, data-cleaning, which takes advantage of data fusion, could be used.
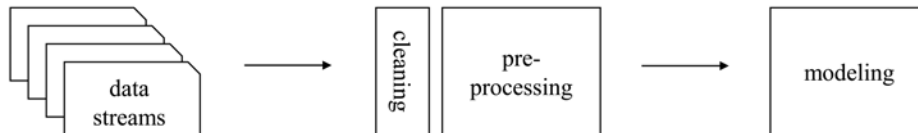
Figure 4: Position of data-cleaning (additive outlier detection) system within the stream mining analytical platform.

4 EVALUATION

The proposed methodology has been used for data cleaning in the following real-world set-ups: thermal power plant (Reggio nell' Emilia, Italy), public buildings (Turin, Italy and Athens, Greece) and public lighting (Miren, Slovenia). As clean data is not available in these scenarios, we tried to evaluate our methodology with publicly available datasets. Those are mainly focused on machine learning methods and are not useful for outlier detection in time series. Authors like Chen [2] create artificial datasets. Marczak [11] dedicates a section of the paper to modeling Markovian processes and obtaining artificial data from these simulations. Xu [14] uses artificial and augmented real-world data sets. No author, however, shares their datasets.

4.1 Artificial Data Set

We provide an artificial dataset, following the usual daily profile of a family of typical sensors. Each dataset introduces a different level of Gaussian noise $N(\mu = 0; \sigma)$ [9]. We have made the dataset publicly available at ResearchGate. Data points are a subject of noise, 1% of data points have been candidates for an AO. Additive factor has been uniformly sampled on the interval from 0 to $0.714 \cdot \max(f(t))$, where $\max(f(t))$ is the maximum value of the underlying dynamics function. Additive factors that were lower than $2 \times \sigma$ have been dismissed.

4.2 Results

Algorithm is illustrated in Figure 5. It shows the impact of Kalman filters short-term prediction and its variance on additive outlier detection. The measurement (depicted in dark blue) that

falls outside the interval around short term prediction (depicted in light blue) is considered an outlier.
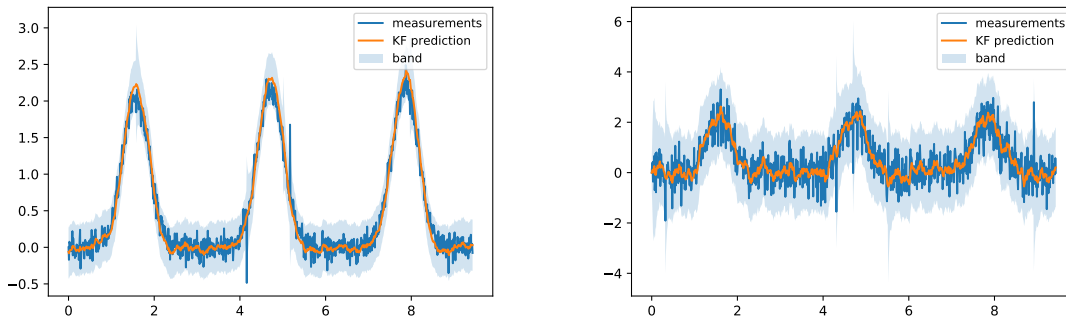
Figure 5: Illustration of algorithm results with 2 different datasets: left - little nois, $\sigma = 0.1$ (data set 1), right - more noise, $\sigma = 0.5$ (data set 5). Kalman filters short-term prediction is depicted in orange, measurements in dark blue. Any measurement outside of the light-blue band (defined by Kalman filter variance) is considered an outlier.

Experimental results are depicted in Table 1. Different data sets (from 1 to 9) introduce different Gaussian noise, which makes it more and more difficult to correctly classify outliers, which can be observed in decreasing values of precision, recall and F_1 in Table 1.

Table 1: Results of Kalman filter AO experiments on a set of artificial data.

Dataset	Noise σ	Kalman filter method			ARIMA method		
		Precision	Recall	F_1	Precision	Recall	F_1
1	0.036	0.866	0.967	**0.914**	0.624	0.874	0.728
2	0.071	0.776	0.983	0.867	0.940	0.829	0.881
3	0.107	0.737	0.872	0.799	0.906	0.750	0.821
4	0.143	0.681	0.946	0.792	0.944	0.740	0.830
5	0.179	0.695	0.592	0.640	0.902	0.643	0.751
6	0.213	0.455	0.873	0.598	0.896	0.520	0.658
7	0.250	0.587	0.373	0.456	0.790	0.448	0.571
8	0.286	0.435	0.779	0.558	0.816	0.461	0.589
9	0.321	0.353	0.545	0.428	0.741	0.336	0.462

We have compared our algorithm to a commonly used methodology from [2]. With datasets with higher noise (datasets from 2 to 9) the methodology (using ARIMA(1, 1, 1)) works better. Our algorithm was significantly better with datasets with lower noise. An obvious downside of the ARIMA methodology [2] is that it requires fitting of ARIMA model on the whole dataset, which makes it obsolete in Big Data scenario.

5   CONCLUSIONS

In this paper we have identified that efficient data pre-processing is needed in streaming data scenarios. We have focused on the first part of data pre-processing pipeline: data cleaning. We conducted a short research on the state-of-the-art in the field and proposed our own method based on Kalman filter. Method has been applied to 4 real world scenarios and quantitatively tested on an artificial data set. This is the first publicly available data set that can be used for comparison of different methods. We have compared our method to the ARIMA state of the art and got better results on the datasets with lower noise ratio and comparable results on the

datasets with higher noise ratio. The main advantage of our method is, that it can work with Big Data in a streaming scenario.

Acknowledgments

References

[1] Brown, R. G. and Hwang, P. Y. C., Introduction to random signal and applied Kalman filtering (1996). New York: John Wiley.

[2] Chen, C. and Liu, L., Joint Estimation of Model Parameters and Outlier Effects in Time Series (1993). Journal of the American Statistical Association, Vol. 88-421:284-297.

[3] Chu, X. et al. Data cleaning: Overview and emerging challenges (2016). In: Proceedings of the 2016 International Conference on Management of Data. ACM, p. 2201-2206.

[4] Dempster, A.P., Laird, N.M. and Rubin, N.M. Maximum likelihood from incomplete data via the EM algorithm (1977). Journal of Royal Statistical Society Series B, Vol 39:1-38.

[5] Fan, W. and Bifet, A., Mining big data: Current status, and forecast to the future (2013). SIGKDD Explor. Newsl., Vol. 13(2):1-4.

[6] Kalman, R. E., A new Approach to linear filtering and prediction problem (1960). Journal of basic engineering, Vol 82(1):34-45.

[7] Kandel, S. et al., Research directions in data wrangling: Visualizations and transformations for usable and credible data (2011). Information Visualization Journal, Vol 10:271-288.

[8] Kenda, K., Škrbec, J., and Škrjanc, M. (2013). Usage of Kalman Filter for Data Clearning of Sensor Data. SiKDD 2013, Ljubljana.

[9] Kenda, K. Artificial data-set for testing time-series additive outlier detecion methods (2017). ResearchGate (DOI: 10.13140/RG.2.2.23165.97760).

[10] Krempl, G. et al., Open challenges for data stream mining research (2014). ACM SIGKDD explorations newsletter, Vol 16(1):1-10.

[11] Marczak, M., Proietti, T., Grassi, S. A data-cleaning augmented Kalman filter for robust estimation of state space models (2017). Econometrics and Statistics, Vol(1-17).

[12] Shearer, C., The CRISP-DM model: the new blueprint for data mining (2000). Journal of data warehousing, Vol 5(4):13-22.

[13] Ting, J., Theodorou, E., Schaal, S. A Kalman Filter for Robust Outlier Detection (2007). Proceeedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, San Diego.

[14] Xu, S., Data Cleaning and Knowledge Discovery in Process Data (PhD thesis) (2015). Austin: The University of Texas.

[15] Wikipedia. Kalman filter. [Accessed: June 7th 2017.]

SETTING THE OPTIMAL PARKING PRICE USING QUEUING MODELS

Baruch Keren
Industrial Engineering and Management Department,
SCE - Shamoon College of Engineering
Bialik/Bazel St., Beer Sheva, Israel
E-mail: baruchke@sce.ac.il

Yossi Hadad
Industrial Engineering and Management Department,
SCE - Shamoon College of Engineering
Bialik/Bazel St., Beer Sheva, Israel
E-mail: yossi@sce.ac.il

Abstract: Parking spaces, especially in cities, are a rare resource. Given a choice, drivers prefer free parking, but free parking is not costless. The residents bear the parking costs through higher taxes and retail prices. The choice is actually between paying for parking directly or indirectly. Paying directly for parking is more justifiable and efficient. This paper proposes a model for managing parking demand by parking pricing. The optimal parking price is defined as the price for maximizing the revenue or for setting a given level of occupancy of the park slots. The paper uses a queuing model and the concept of price elasticity to calculate the best parking price.
Keywords: Queuing model, Parking policy, Parking pricing, Occupancy, Price-demand elasticity.

1 INTRODUCTION

Many theoretical and empirical papers analyzed the occupancy and the pricing of parking by concentrating on particular aspects of the issue (see the survey about the economics of parking done by [1]). Cars are parked 95% of the time and large amounts of land are used for parking [2]. Given a choice, drivers prefer free parking, but free parking is not choice is actually between paying for parking directly or indirectly [3]. Paying directly for parking is more justifiable and efficient. Drivers park free for 99 percent of automobile trips in the US. Shoup [2] claimed that cities should set the right price for curb parking. The results of underpriced parking are many drivers that cruise for park and increasing the traffic congestion in the street. Free parking shifts the parking cost from the transportation sector to everything else. Shoup [2] presented three basic recommendations: 1) Set the right price for parking. 2) Use the parking revenue for local public service. 3) Remove the minimum free public parking requirements for new building projects. Some cities adopted the policy of setting price to produce one or two open spaces on every block as performance pricing. This policy guarantees that the parking slots will be well used, but also remain readily available for drivers who want to park. Available parking decrease the cruising for parking, decrease the waste of fuel and time and decrease the air pollution.

This paper proposes a method for calculating and determining the optimal price for maximizing the revenue from a car park or the price for obtaining a given level of occupancy in a car park or in a city.

2 MODEL DEVELOPMENT

A car park has N parking slots. The drivers pay for their parking P\$ per hour. The average arrival rate to the park is λ cars per hour. The value of λ is depending on the parking price, P, denoted as $\lambda(P)$. The average parking time, $1/\mu$, is also depending on the parking

price, denoted $1/\mu(P)$. The assumptions are that the arrival rate of cars is Poisson with $\lambda(P)$ and the departure rate of the parked cars that leave the parking slots is exponential with $\mu(P)$ (a Markovian process). If a car arrives and the car park is full (all the N slots are occupied), the car leaves the park and not waits in a queue for a free slot. Hence, the maximum number of cars in the system is limited to N. Under these assumptions, the car park is a $M/M/N/N$ queuing system. The $M/M/N/N$ is a known que (see for example [4]), where the probability for x cars in the park, P_x, is given by the following Erlang's Loss function:

$$P_x = \frac{\dfrac{(\lambda(P)/\mu(P))^x}{x!}}{\displaystyle\sum_{x=0}^{N}\frac{(\lambda(P)/\mu(P))^x}{x!}} \tag{1}$$

By defining the average traffic intensity as $\rho(P)=\dfrac{\lambda(P)}{\mu(P)}$, the previous equation can written as follows:

$$P_x = \frac{\dfrac{(\rho(p))^x}{x!}}{\displaystyle\sum_{x=0}^{N}\frac{(\rho(p))^x}{x!}} \tag{2}$$

The last equations are valid only for $\lambda(P)>0$, $\mu(P)>0$ and $\rho(P)>0$.

2.1 MAXIMIZING THE REVENUE

The hourly price, P, multiplied by the expected number of the occupied slots, $E(X)$, is the expected hourly revenue from the park, $P \times E(X)$. It is clear that $E(X)$ is a function of the parking price P. Higher parking price decreases the average occupancy. Therefore, the aim is to set the parking price that maximizes the revenue:

$$P \times E(X) = P \sum_{X=0}^{N} X \frac{\dfrac{(\rho(P))^x}{x!}}{\displaystyle\sum_{x=0}^{N}\frac{(\rho(P))^x}{x!}} \tag{3}$$

The optimal price $P*$ that maximizes the hourly revenue is:

$$P* = \frac{\displaystyle\sum_{X=0}^{N}\frac{\rho(P)^X}{(X-1)!}}{\left(\frac{\displaystyle\sum_{x=0}^{N}\frac{\rho(P)^{X-1}}{(X-1)!}\rho'(P)\sum_{X=0}^{N}\frac{\rho(P)^X}{(X-1)!}}{\left(\displaystyle\sum_{x=0}^{N}\frac{\rho(P)^X}{X!}\right)} - \sum_{X=0}^{N}\frac{X\rho(P)^{X-1}\rho'(P)}{(X-1)!}\right)} \tag{4}$$

In order to compute P^*, the functions $\lambda(P)$ and $\mu(P)$ must be explicitly formulated. After that, the value of P^* can be calculated by numerical search methods, trial and error or by software tools such the Excel solver.

2.2 SETTING THE PRICE FOR OBTAINING A GIVEN LEVEL OF OCCUPANCY

As mentioned, decision-makers may prefer a policy of setting a price that leaves some open slots in the car park and not to maximize the revenue. This policy guarantees that the park will be well used, but also remain available for drivers who want to park. In other words, the aim is to set a price for obtaining a given level of occupancy, for example, occupancy of 90%. The average occupancy is $E(X)/N$. Therefore, the objective is to obtain $E(X)/N = K$, where K is the given level of preferred occupancy. The average occupancy is calculated as follows:

$$\frac{E(X)}{N} = \frac{1}{N} \sum_{X=0}^{N} X \frac{\frac{(\rho(P))^x}{x!}}{\sum_{x=0}^{N} \frac{(\rho(P))^x}{x!}} = K \tag{5}$$

The value of P for the required occupancy can be computed by numerical search methods, trial and error or by software tools such the Excel solver.

2.3 THE PRICE ELASTICITY

In order to find the optimal parking price, the price elasticity coefficients for the arrival rate and for the parking time should be measured. The price elasticity of demand [5] is defined as the percentage of change in the use of some commodity or service based on one percent change in its price, and it is expressed by an elasticity coefficient. The elasticity coefficient can be calculated in several ways. One method is the "Mid-point arc elasticity" that is frequently used in predicting road user behavior (for example, [6]).

It is proposed here to estimate the arrival rate to the park and the parking time by assuming a linear (or piecewise linear) relation between the price and the arrival rate and between the price and the parking time.

Denote:

P_1 – the current parking price per hour.

P_2 – the maximum parking price per hour.

λ_1 – the mean arrival rate (number of customers arriving per hour) where the current parking price per hour is P_1.

λ_2 – the mean arrival rate where the maximum parking price per hour is P_2.

Note that $P_1 < P_2$ and $\lambda_1 > \lambda_2$.

S_1 – the average staying time of a customer in park car (service time in hour) where the current parking price per hour is P_1.

S_2 – the average staying time of a customer in park car (service time in hour) where the current parking price per hour is P_2.

The assumption is that for $P_1 < P_2$, $S_1 > S_2$. When the parking price per hour is higher, the driver is motivated to leave the park sooner.

$\mu(P)$ – The average number of customers that leave one parking slot per hour (mean service rate) is $\mu(P) = \dfrac{1}{S(P)}$.

Note that when the slopes are varied with the price, the linear lines for the arrival rate and for the staying time can be extended to piecewise linear models.

3 NUMERICAL EXAMPLE

Consider a car park with $N = 100$ slots. The current price is 5.7 ILS (1 USD is about 3.8 ILS). For this price, the arrival rate per hour is 125 cars. The average staying time for each car is 2.75 hours. The estimation is that for price of 20 ILS, the arrival rate per hour would be 50 cars. The average staying time for each car in this price is 1.25 hours.

Assuming a linear relation between price and arrival rate and a linear relation between price and the parking time, one can calculate that $\lambda(P) = 154.89 - 5.2448P$ and $\mu(P) = \dfrac{1}{3.3479 - 0.1049P}$.

Figure 1 presents the revenue per hour as a function of the price per hour. The revenue function in this example (figure 1) seems to be a unimodal with one maximum value. The optimal price is $P^* = 17.22$ ILS (calculated by excel solver).
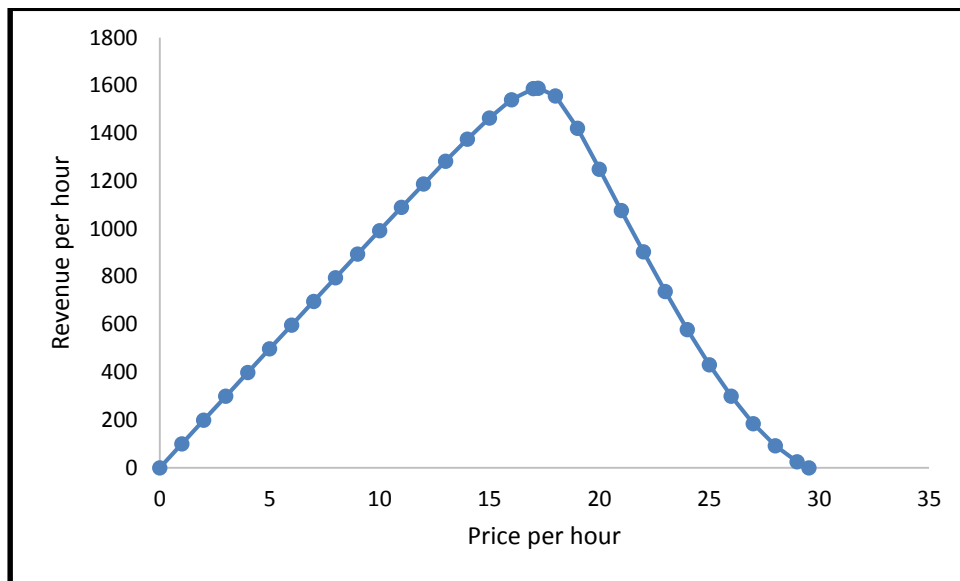
Figure 1: The revenue per hour as a function of the parking price

If the aim is to set a given value for the average occupancy (e.g., 90%) in order to preserve 10% of the parking slots free, one can set the appropriate price that can be calculated by equation (7) (17.58 ILS in this case). Figure 2 presents the occupancy of the park car as a function of the price per hour for this numerical example.
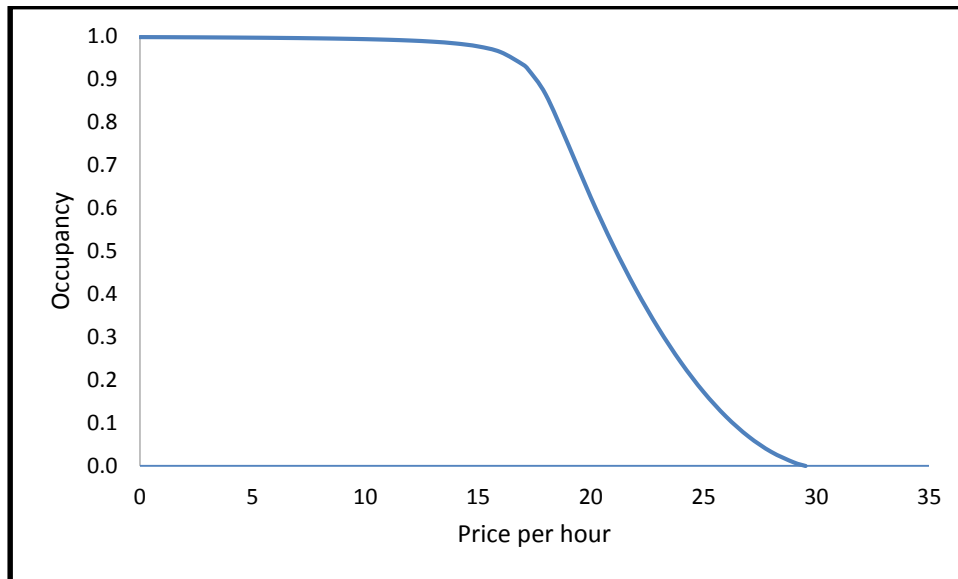
Figure 2: Occupancy as a function of the parking price

4 SUMMARY

This paper proposes a method for determining the optimal parking price via queuing model and the concept of price-demand elasticity. The model can be used by practitioners who want to set the right price for parking. The ideas of this research can be extended to more complicated queues and more sophisticated pricing schemes.

References

[1]   Inci, E. (2015). A review of the economics of parking. Economics of Transportation, 4(1), 50-63.
[2]   Shoup, D.C. (2005). The high cost of free parking. Chicago: Planners Press.
[3]   TDM Encyclopedia. (2017), Parking Pricing Direct Charges for Using Parking Facilities. http://www.vtpi.org/tdm/tdm26.htm, accessed 11/may/2017.
[4]   Gross, D., Shortle, J.F., Thompson J.M., Harris C.M. (2008). Fundamentals of queueing theory. New Jersey: Wiley & Sons.
[5]   TCRP (2005) 'Parking Prices and Fees: Traveller Response to Transportation System Changes'. Transit Cooperative Research Program Report 95, Chapter 13, Transportation Research Board, Washington D.C.
[6]   Simićević, J., Milosavljević, N., Maletić, G., Kaplanović, S. (2012). Defining parking price based on users' attitudes. Transport Policy, 23, 70-78.

A PARALLEL IMPLEMENTATION OF THE BOUNDARY POINT METHOD[1]

Mircea Simionica*, Janez Povh†,‡

* UniCredit Business Integrated Solutions, Financial Risks Factory Milan, Italy
†University of Ljubljana, Faculty of mechanical engineering, Slovenia
‡Institute of mathematics, physics and mechanics Ljubljana, Slovenia
email: mircea.simionica@gmail.com, janez.povh@fs.uni-lj.si

Abstract

The boundary point method has turned out to be a very efficient algorithm for special classes of semidefinite programming problems (SDP) with a large number of (nearly orthogonal) constraints. In this paper we present results of a project, where we have developed a C++ parallel version of this method. Our implementation outperforms normal (non-parallel) implementation of the method and scales very good for problems having block diagonal structure. The final product allows to solve problems that are beyond reach for the state-of-the-art methods.

Keywords: semidefinite programming, boundary point method, parallel algorithm, high performance computing

Math. Subj. Class. (2010): Primary 90C22, 68W10, Secondary 90C06

1 INTRODUCTION

Semidefinite programming (SDP) has experienced a big development in mathematical optimization since the 1990's. It is a non-trivial extension of linear programming and can be used to better model real-life problems. We can find theoretical and practical applications of SDP in real algebraic geometry [7], in combinatorial optimization [12], control theory [2], data science[3, 9], etc. – see e.g. [1] for nice overview of SDP applications.

Semidefinite programming is concerned with finding a positive semidefinite matrix that yields optimal value of a linear objective function and that satisfies a given number of linear equations. In mathematical terms a primal semidefinite program (PSDP) can be expressed as follows:

$$
\begin{array}{ll}
\max & \langle C, X \rangle \\
\text{such that} & A(X) = b \\
& X \succeq 0
\end{array}
\qquad \text{(PSDP)}
$$

and the dual problem, associated to (PSDP), is

$$
\begin{array}{ll}
\min & b^T y \\
\text{such that} & A^T(y) - C = Z \\
& Z \succeq 0
\end{array}
\qquad \text{(DSDP)}
$$

Here $\langle \cdot, \cdot \rangle$ denotes standard scalar product over the symmetric matrices, $A(X) = b$ denotes linear constraints in elements of matrix X and A^T is the adjoint operator to A. $X \succeq 0$ denotes that matrix X is positive semidefinite - symmetric with all eigenvalues non-negative. If we know that the matrix variable X possesses block diagonal structure, i.e. X has form

---

$$X = \begin{bmatrix} X_1 & 0 & \ldots & 0 \\ 0 & X_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & X_k \end{bmatrix}$$

then X is positive semidefinite if and only if each diagonal block is positive semidefinite and (PSDP) simplifies into

$$\begin{array}{ll} \max & \sum_i \langle C_i, X_i \rangle \\ \text{such that} & \sum_i A_i(X_i) = b \\ & X_i \succeq 0, \ \forall i \end{array} \tag{bPSDP}$$

2 IMPLEMENTATION

The Boundary point method (BPM) is a relatively new algorithm [8] that sets itself as an alternative to interior point methods. The algorithm operates on the dual formulation of the problem (DSDP). Note that the triple $(X; y, Z)$ is optimum for (PSDP) and (DSDP) if and only if it satisfies the following conditions:

$$A(x) = b, \ \ A^T(y) - C = Z, \ \ XZ = 0, \ \ X \succeq 0, \ \ Z \succeq 0.$$

BPM tries to solve these equations while respecting $X \succeq 0, \ Z \succeq 0$. More precisely, it maintains $XZ = 0, \ X \succeq 0, \ Z \succeq 0$ and tries to reach feasibility for primal and dual linear constraints. Lagrange multipliers are applied to the dual equations, yielding the following Augmented Lagrangian:

$$L_\sigma = b^T y + \langle X, Z + C - A^T(y) \rangle + \frac{\sigma}{2} \left\| Z + C - A^T(y) \right\|^2$$

The steps that define the complexity of the algorithm are solving a large system of linear equations and computing a spectral decomposition of a symmetric matrix in each iteration. The main idea of this method is visualised on Figure 1 and the pseudo-code of the algorithm is illustrated in Figure 2.
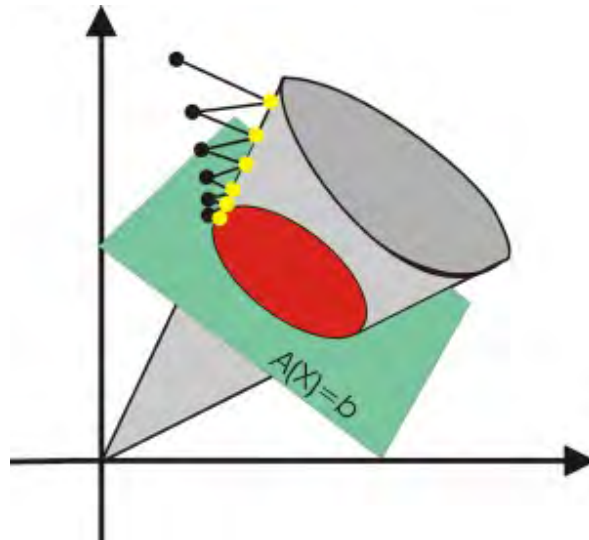
Figure 1: Idea of Boundary point method: we compute iterates outside the cone and project them to the cone. While running the method the iterates come closer and closer to the feasible hyperplane. Once they hit it they are actually an optimal solution.

<div align="center">Parallel Boundary Point Method to solve (bPSDP)</div>

INPUT: A, b, C

  Select $\sigma > 0, \{\varepsilon_k\} \to 0, \ \varepsilon > 0.$
  $k := 0; \ y^k := y; \ X^k := 0;$
  $W := A^T(y^k) - C - \frac{1}{\sigma}X^k; \ Z^k = W_+;$
  $\delta_{outer} := \|Z^k - A^T(y^k) + C\|;$
  while $\delta_{outer} > \varepsilon$ (Outer iteration for $k = 0, 1, \dots$)
      repeat until $\delta_{inner} \leq \sigma\varepsilon_k$ (Inner iteration: $(X^k, \ \sigma)$ held constant)
          Solve for y^k: $A(A^T(y^k)) = A(Z^k + C) + \frac{1}{\sigma}(A(X^k) - b);$
          for each block i
              $W^i := A^T(y^k) - C - \frac{1}{\sigma}(X^k)^i;$
              $(Z^k)^i := W^i_+;$
              $(V^k)^i := -\sigma W^i_-;$
              $\delta_{inner} := \sum_i \|A_i((V^k)^i) - b\|$
          end (for)
      end (repeat)
  $X^{k+1} := \mathrm{diag}((V^k)^1, (V^k)^2, \dots);$
  $k := k + 1;$
  $\delta_{outer} := \|Z^k - A^T(y^k) + C\|;$
  end (outer while)

Figure 2: Parallel boundary point method pseudo-code. The most straightforward part for parallelization is the central **for** loop which can be performed completely in parallel. Another rewarding part for parallelization is solving the system of linear equations before the **for** loop and computing the spectral decompositions inside the **for** loop.

The goal of this paper is to present a C++ parallel version of an existing MATLAB code, which was developed while the first author was a Summer of HPC fellow under supervision of the second author. The code is specifically prepared for semidefinite programs characterized by a block diagonal structure. This type of structure leads to a parallelization based on the number of diagonal blocks. The parallel aspect focuses on the inner iteration: once the sparse linear system has been solved, the remaining steps are repeated according to the number of blocks. They are also independent so they can be solved concurrently. The serial code loops through the blocks, whereas the parallel version solves each block independently, speeding up the computations. This allows to solve larger instances of semidefinite problems, defined by more diagonal blocks and with a bigger number of constraints, leading to new challenges for other methods.

The code is written in C++, MPI is used as message-passing system between processes and a manual is included, serving as documentation of the code. External linear algebra libraries are heavily used. Armadillo is utilized throughout the whole code, while the linear system is solved with Eigen's sparse solver. MATLAB (or Octave, freely available under the GNU license) is used for generating the input data. In the process Armadillo library turned out to be really helpful. Armadillo is an open-source, high quality linear algebra library useful for algorithm development. Its syntax is similar to MATLAB, so we saved time not implementing specific MATLAB features.

2.1 Few challenges

Some time was spent on a parallel implementation of the eigen decomposition using ScaLA-PACK. This was useful since it allowed us to better understand how the block-cyclic distribution of data among the processes works. We have soon realized that the sparse linear system was actually slowing the code down, so we have decided to change route. We encountered most of the difficulties working on this stage of the project, since we had issues linking SuperLU, a direct sparse solver, to Armadillo.

2.2 First improvements

We experienced first improvements once SuperLU was correctly linked, but the code was still underperforming with respect to MATLAB. We knew that Eigen was another good linear algebra library so we decided to use one of its sparse linear solvers. Specifically, we went for Eigen's *SparseCholesky* module. This class provides Cholesky factorizations, which allow for solving $AX = b$ linear systems. In particular, the matrix A is factored as $A = LL^T$, where L is lower triangular. The system $LL^T x = b$ is then solved with forward and back substitutions. This is good in terms of performance since the factorization is computed only once, outside of the main loop and not in each iteration. Choosing to include Eigen objects in our code allowed us to deal with multiple linear algebra libraries, pushing us to search and find out methods of importing/exporting the data between different classes in an efficient manner. In fact, Eigen objects can be mapped to those of Armadillo. This avoids copying a huge amount of data in every iteration.

2.3 Parallel implementation

Solving the sparse linear system using Eigen rather than Armadillo considerably pushed the solve time down. The C++ serial version was now aligned to MATLAB's performance. We decided it was the right moment to begin the parallelization phase. We focused on the for loop inside the inner iteration (2). Instead of computing several spectral decompositions in sequential, the calculations are now splitted among the processes and the current solution is updated locally.

2.4 Even more improvements

The parallel version looked promising. We performed some initial tests, alternating between increasing the number of blocks and the size of the matrix of coefficients. There was evidence of scalability with the number of blocks (hence with the number of processes). Naturally the code did not scale linearly because of communication overheads. The results were pointing in the right direction, but there was one more thing to do. Armadillo integrates with BLAS and LAPACK libraries. These are packages providing matrix operations and numerical linear algebra routines. Up to now Armadillo was linked against traditional BLAS and LAPACK packages. The biggest improvement in the code came in the moment we linked Armadillo against OpenBLAS. OpenBLAS is a multi-threaded, high performance replacement of BLAS. This allowed the C++ serial code to compete with MATLAB's performance and made possible bigger speedup ratios running the parallel version.

3 RESULTS

In this section we present preliminary results obtained by using parallel BPM to compute values of $\Psi_K(G)$, which is a lower bound for the chromatic number of graph G. Here we recall the formulation of $\Psi_K(G)$ from [6, 4]:

$$
\begin{aligned}
\Psi_K(G) \quad = \quad & \min \quad t \\
\text{s.t.} \quad & Y_{00}^0 = t, \\
& Y_{ii}^0 = 1 \ \forall i \in \{1,...,n\}; \\
& Y_{pq}^i = 0 \quad \text{if } \{i,p,q\} \text{ contains an edge;} \qquad (\Psi_{SDP}) \\
& Y_{pq}^i = Y_{p'q'}^{i'} \\
& \quad \text{if } (\{i,p,q\})\backslash\{0\} = (\{i',p',q'\})\backslash\{0\}; \\
& Y^0 - \sum_{i \in K} Y^i \succeq 0, \ Y^i \succeq 0 \ \ \forall i \in K.
\end{aligned}
$$

In Table 1 we report numerical results obtained by running serial and parallel version of BPM on graphs with 100, 200, 300, 400 and 500 nodes and with several cardinalities of K - note that the number of blocks is exactly $|K| + 1$.

Number of blocks	Matrix dimension	Number of constraints	Number of spectral decompositions	Degree of convergence	Time (sec.) MATLAB	Time (sec.) C++
6	500	378456	32100	0.02558	3980.93	840.70
8	200	82824	42800	0.01662	641.20	106.40
8	400	325592	42800	0.02141	2924.00	505.20
10	100	26080	53500	0.02938	183.14	25.76
10	300	228410	53500	0.01648	2155.16	291.20

Table 1: Time comparison between MATLAB and C++ parallelized version of BPM. We used HPC system, owned by University of Ljubljana, Faculty of mechanical engineering, which is composed of two systems: Intel Xeon X5670 HPC system with 1536 hyper-cores and E5-2680 V3 (1008 hyper-cores) DP cluster, with IB QDR interconnection, 164 TB of LUSTRE storage, 4.6 TB RAM and with 24 TFlop/s performance.

Figure 3 depicts the scaling factor - we can see that scaling is close to linear. This is due to the fact that the bottleneck of the BPM is the spectral decomposition and this is exactly the point where we applied parallelization. The final code is available under the GNU General Public License [11].

4 CONCLUSIONS

Semidefinite programming is a strong tool to approximately solve very hard problems from mathematical optimization: mixed integer linear and quadratic programming problems, a wide range of non-linear programming problems and also to detect positivity of commutative and non-commutative polynomials, an inspiring problem from real algebraic geometry. It is also a very powerful tool to computed advanced data models [3, 9]. Its main drawback compared to linear programming is the practical complexity - meeting semidefinite constraints is much harder compared to non-negativity constraints in linear programming. However, BPM, especially parallel version presented here, is an important step towards making SDP a method of choice also for medium size problems, especially if we run it on a high performance computer, like we did. However, scaling up the method to run efficiently on world largest HPC systems is still a challenge waiting to be tackled in the future.
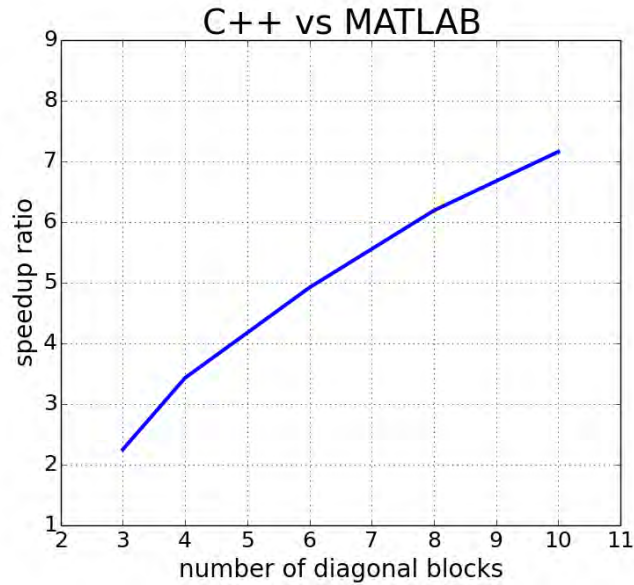
Figure 3: Speed-up ratios

References

[1] M. F. Anjos and J. B. Lasserre. Handbook on semidefinite, conic and polynomial optimization, International Series in Operations Research & Management Science, vol. 166, 2012.

[2] G. E. Dullerud and F. Paganini. *A course in robust control theory: a convex approach*, Vol. 36., Springer Science & Business Media, 2013.

[3] B. Hajek, W. Yihong and X. Jiaming Xu. Achieving exact cluster recovery threshold via semidefinite programming. IEEE Transactions on Information Theory 62.5:2788-2797, 2016

[4] J. Govorčin, N. Gvozdenović and J. Povh. New heuristics for the vertex coloring problem based on semidefinite programming. Central European Journal of Operations Research, 21(1):13–25, 2013

[5] G. Gaël and J. Benoît et al. *Eigen v3*. Retrieved from *http://eigen.tuxfamily.org*, 2010.

[6] N. Gvozdenović and M. Laurent. Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. SIAM Journal on Optimization, 19(2):592–615, 2008.

[7] J. B. Lasserre. Moments, positive polynomials and their applications. Volume 1 of Imperial College Press Optimization Series. Imperial College Press, London, 2010.

[8] J. Povh, F. Rendl and A. Wiegele. A boundary point method to solve semidefinite programs. Computing, 78(3):277-286, November 2006

[9] F. Ricci-Tersenghi, A. Javanmard and A. Montanari. Performance of a community detection algorithm based on semidefinite programming. Journal of Physics: Conference Series. Vol. 699. No. 1. IOP Publishing, 2016.

[10] C. Sanderson. Armadillo: an open source C++ linear algebra library for fast prototyping and computationally intensive experiments. *Technical Report*, NICTA, 2010.

[11] M. Simionica and J. Povh. Parallel Boundary Point Method - available on `https://github.com/mirceas08/bpm`.

[12] L. Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*, Vol. 27, American Mathematical Soc., 2016.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 3:
Logistics

THE DESIGN OF CHARGING INFRASTRUCTURE FOR ELECTRIC VEHICLES AND ITS PROPERTIES

Peter Czimmermann
University of Žilina, Žilina, Slovakia, peter.czimmermann@fri.uniza.sk
Michal Koháni
University of Žilina, Žilina, Slovakia, Michal.Kohani@fri.uniza.sk
Ľuboš Buzna
University of Žilina, Žilina, Slovakia, Lubos.Buzna@fri.uniza.sk

Abstract: We study methods to design a private charging infrastructure for a fleet of electric vehicles, for example taxicabs, small vans (used in the city logistics) or shared vehicles. In this contribution, we continue our previous work, where we suggested an MIP model of this location-scheduling optimization problem. We introduce an IP model and simplify it using flows in networks. We suggest an algorithm that finds a feasible solution. We also use the flow version to answer the question from our previous work: What is the cause of the gap in the number of feasible vehicles obtained from the optimization problem and evaluation procedure?

Keywords: electric vehicles, charging infrastructure, flows in networks

1   INTRODUCTION

In [2], we suggested and solved an MIP model for the location-scheduling optimization problem to design private charging infrastructure for electric vehicles. In the extended version [3] of the above mentioned paper, we modified this model and added a procedure which simulates the vehicle charging in specific cases, when drivers operate with incomplete information. The tests have been carried out using real data based on activities of taxicabs in Stockholm. We found out that there is a large gap in the number of feasible vehicles obtained from the solution of the mathematical model and the number of feasible vehicles that form an output of the evaluation procedure. The above mentioned model from the work [3] is in Figure 1.

Here, I is the set of candidate locations, where it is possible to place the charging infrastructure. T is the set of non-overlapping time intervals. The fleet is represented by the set of vehicles C. Vehicles have battery with capacity β (measured in kilometres) and s is the charging speed. R_c is the ordered sequence of parking events of vehicle $c \in C$ and N_{cr} is the list of intervals that overlap with the parking event $r \in R_c$. The fraction of the interval $t \in T$, when the vehicle $c \in C$ is parking, we denote by $a_{ct} \in \langle 0, 1 \rangle$. $B_{itc} = 1$ when the vehicle $c \in C$

Model formulation

$$\text{Minimize} \sum_{i \in I} s_i \tag{1}$$

$$\text{subject to} \sum_{c \in C} B_{itc} x_{ct} \leq s_i \qquad \text{for } i \in I, t \in T \tag{2}$$

$$d_{c0} \leq \alpha\beta \tag{3}$$

$$d_{cr} + \sum_{t \in N_{c,r}} a_{ct} x_{ct} s \leq \beta \qquad \text{for } c \in C, r \in R_c \cup \{r_c\} \tag{4}$$

$$d_{cr} \leq d_{c,r-1} - u_{cr} + \sum_{t \in N_{c,r-1}} a_{ct} x_{ct} s \qquad \text{for } c \in C, r \in R_c \cup \{r_c\} \tag{5}$$

Figure 1: Mathematical formulation of the location-scheduling optimization problem.

1542 vehicles				
ρ_{max}	M_{min}	Model	Strategy 1	Strategy 2
100	800	609	401	363
	150	1186	652	536
	100	1287	689	586
500	800	1102	592	589
	150	1442	653	603
	100	1475	709	695
1000	800	1347	739	692
	150	1499	785	701
	100	1510	793	706

Table 1: Numbers of feasible vehicles, obtained from the mathematical model and the charging simulation procedure.

parks at location $i \in I$ during the time interval $t \in T$, and $B_{itc} = 0$ otherwise. Vehicle $c \in C$ drives u_{cr} kilometres while driving from the parking event $r-1$ to the parking event r. We work with the following variables: $s_i \in Z_0^+$ represents the number of charging points placed at the location $i \in I$; $x_{ct} \in \{0,1\}$, where $x_{ct} = 1$ when vehicle $c \in C$ is charged during the time interval $t \in T$ and $x_{ct} = 0$ otherwise; and $d_{cr} \geq 0$ is the distance that the vehicle $c \in C$ is able to drive at the beginning of the parking event $r \in R_c$. In the objective function (1), we minimize the number of located charging points. Constraints (2) ensure that we cannot simultaneously use more charging points in each time interval than is the number of all existing charging points. Constraints (4) ensure that battery capacity is not exceeded and constraints (5) ensure the contiguity of charging and discharging of batteries.

In the charging procedure, we process events in ascending order with respect to their time of occurence. If we have a parking event, then we assign to the vehicle associated with this event a specified charging point according to the chosen strategy. Strategy 1 assigns to each vehicle a free charging point that allows charging to the maximum capacity. Strategy 2 assigns to each vehicle the first available charging point that is free.

The results of the comparison for one selected week can be seen in Table 1.

Where ρ_{max} is the maximum acceptable distance of a vehicle from a charging station, M_{min} is the least number of parking events in a location to be acceptable as a candidate for a charging station. In the remaining three columns, we have the numbers of feasible vehicles, which we obtain as results of optimisation (Model) and the charging evaluation procedure (Strategy 1, Strategy 2). More results and detailed description can be found in [3].

2 DISCRETE MODEL

In this section, we present a discrete version of Model 1. For technical reasons, we omit the constraints (3) and values a_{ct}. Variables d_{cr} should be from the set Z_0^+ of non-negative integers. The model is in Figure 2 and the example can be seen in Figure 3, where $t_1, \ldots, t_7$ are time intervals, i_1, i_2, i_3 are candidate locations and c_1, c_2 are vehicles.

3 SIMPLIFICATION

The discrete model from the previous section is a basis for the reduction that is presented here. We need to state several constraints to obtain a model that will help us to answer the question about the gap in the number of feasible vehicles in the columns in Table 1. The ordered set

Model 2 - discrete model formulation

$$\text{Minimize} \sum_{i \in I} s_i \tag{6}$$

$$\text{subject to} \sum_{c \in C} B_{itc} x_{ct} \leq s_i \qquad \text{for } i \in I, t \in T \tag{7}$$

$$d_{cr} + \sum_{t \in N_{c,r}} x_{ct} s \leq \beta \qquad \text{for } c \in C, r \in R_c \cup \{r_c\} \tag{8}$$

$$d_{cr} \leq d_{c,r-1} - u_{cr} + \sum_{t \in N_{c,r-1}} x_{ct} s \qquad \text{for } c \in C, r \in R_c \cup \{r_c\} \tag{9}$$

Figure 2: Discrete version of the location-scheduling optimization problem.
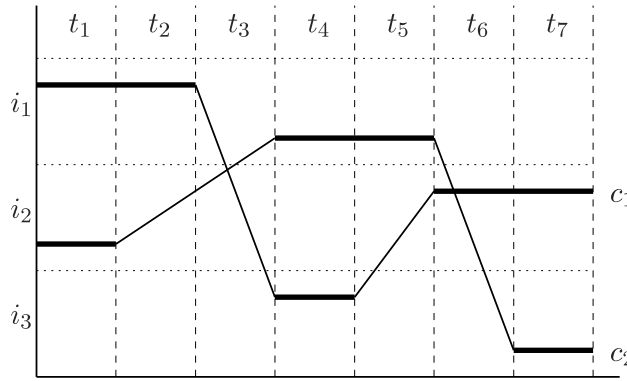
Figure 3: Diagram illustraiting parking events and movements of vehicles.

of all time intervals $T = \{1, 2, \ldots, n\}$ is partitioned into subsets T_{odd} of all time intervals on odd positions and T_{even} of all time intervals on even positions. We suppose that each parking event and journey of the vehicle takes one time interval. The parking events are in even time intervals and journeys are in odd time intervals. We also suppose that at most one charging point can be placed in each candidate location. These restriction can be expressed in Model 2 as follows: $s_i \in \{0, 1\}$, $s = 1$, $u_{c,r} \in \{0, 1\}$, $N_{c,r} = \{2r\}$, $B_{i,t,c}$ and $x_{c,t}$ can be equal to one only in cases $t \in T_{even}$. We call this simplification: Model 3, and we show that it is equivalent to a specific network flow model (explained below).

We construct a network G which represents the situation described above. The set of vertices of G is the union of the sets
$V_1 = \{v_{it} : i \in C, t \in T\}$, where v_{it} represents the vehicle i in time t,
$V_2 = \{u_{it} : i \in I, t \in T_{even}\} \cup \{w_{it} : i \in I, t \in T_{even}\}$, where u_{it} and w_{it} represent the candidate location i in time t,
$V_3 = \{s, z, w\}$, where s is the source, z is the sink and w an additional vertex of the network.

The set of edges is the union of the sets
$E_1 = \{(v_{i,t-1}, v_{i,t}) : i \in C, t \in T - \{1\}\} \cup \{(v_{i,n}, z) : i \in C\}$,
$E_2 = \{(u_{i,t-2}, u_{i,t}) : i \in I, t \in T_{even} - \{2\}\}$,
$E_3 = \{(u_{i,t}, w_{i,t}) : i \in I, t \in T_{even}\}$,
$E_4 = \{(w_{i,t}, v_{j,t}) : i \in I, t \in T_{even}, j \in C, B_{i,t,j} = 1\}$,
$E_5 = \{(v_{i,t}, w) : i \in C, t \in T_{odd}\}$,
$E_6 = \{(s, v_{i,1}) : i \in C\}$,
$E_7 = \{(s, u_{i,2}) : i \in I\}$,

E_1	$\langle 0, \beta \rangle$	E_6	$\langle \beta, \beta \rangle$
E_2	$\langle 0, |T_{even}| \rangle$	E_7	$\langle 0, |T_{even}| \rangle$
E_3	$\langle 0, 1 \rangle$	E_8	$\langle 0, |T_{even}| \rangle$
E_4	$\langle 0, 1 \rangle$	E_9	$\langle 0, \infty \rangle$
E_5	$\langle 1, 1 \rangle$		

Table 2: Lower and upper bounds for the flows on edges.

$E_8 = \{(u_{i,m}, z) : i \in I, m = \max(T_{even})\}$,
$E_9 = \{(w, z)\}$.
Every edge has assigned to it a lower and upper bound for the flow, which are represented by the interval $\langle l, u \rangle$. These bounds are given in Table 2.

The optimal solution of Model 3 is equivalent to the feasible flow in G with the minimum number of non-zero flows on the edges of the set E_7. If the flow on the edge $(s, u_{i,2}) \in E_7$ is non-zero, then we place a charging point in location i. Problems of finding a feasible flow in the network with lower bounds on edges and its minimisation are solvable in polynomial time [1].

We can use the following algorithm to obtain the solution with fewer edges from E_7 with non-zero flow.

Description of the algorithm. We construct a network G' (it is subnetwork of G):
Its vertex set is $V_1' \cup V_2' \cup V_3'$, where
$V_1' = \{v_{i,t} : i \in C, t \in T - \{1\}\} \cup \{z\}$,
$V_2' = \{u_{i,t} : i \in I, t \in T_{even}\} \cup \{w_{i,t} : i \in I, t \in T_{even}\}$.
The edge set is the union of the sets:
$E_1' = \{(v_{i,t-1}, v_{i,t}) : i \in C, t \in T - \{1, 2\}\} \cup \{(v_{i,n}, z) : i \in C\}$,
$E_2' = \{(u_{i,t-2}, u_{i,t}) : i \in I, t \in T_{even} - \{2\}\}$,
$E_3' = \{(u_{i,t}, w_{i,t}) : i \in I, t \in T_{even}\}$,
$E_4' = \{(w_{i,t}, v_{j,t}) : i \in I, t \in T_{even}, j \in C, B_{i,t,j} = 1\}$,
$E_5' = \{(u_{i,m}, z) : i \in I, m = \max(T_{even})\}$.
Lower and upper bounds of the edges are taken from G.

Algorithm.
Input is the network G with feasible flow x and subnetwork G'.
For each pair of vertices $u_{i,2}, u_{j,2} \in V_2$ such that $0 < x(s, u_{i,2}) \leq x(s, u_{j,2}) < |T_{even}|$ **do**:
$\longrightarrow$ **while** there is an augmenting path $P(j, i)$ with reserve r from $u_{j,2}$ to $u_{i,2}$ in G' **do**:
$\longrightarrow$ $\longrightarrow$ add the edges $(s, u_{j,2})$ and $(s, u_{i,2})$ to $P(j, i)$ to form a (non-oriented) cycle
$\longrightarrow$ $\longrightarrow$ $C = (s, u_{j,2}, \ldots, u_{i,2}, s)$,
$\longrightarrow$ $\longrightarrow$ change the flow x in C as follows:
$\longrightarrow$ $\longrightarrow$ **if** (u, v) is the forward edge in C, **then** $x(u, v) = x(u, v) + r$,
$\longrightarrow$ $\longrightarrow$ **if** (u, v) is the reverse edge in C, **then** $x(u, v) = x(u, v) - r$,
$\longrightarrow$ process another pair of vertices.

A polynomial algorithm for finding an augmenting path in network is presented in [4].
This approach does not guarantee that given solution is optimal. We aim to test the algorithm and study its properties in our future works.

Example 1.
Situation with two vehicles, two candidate locations and five time intervals can be seen in Figures 4 and 5.
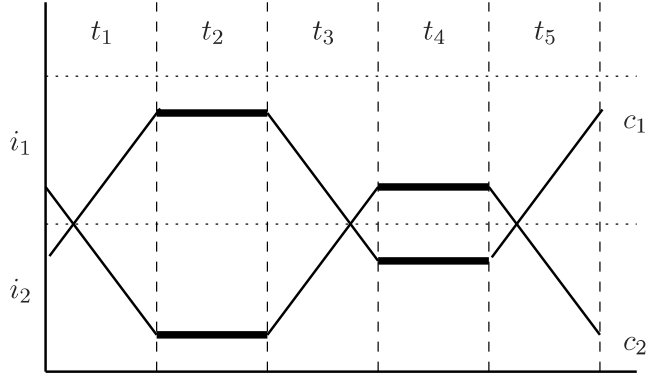
Figure 4: Diagram from Example 1.
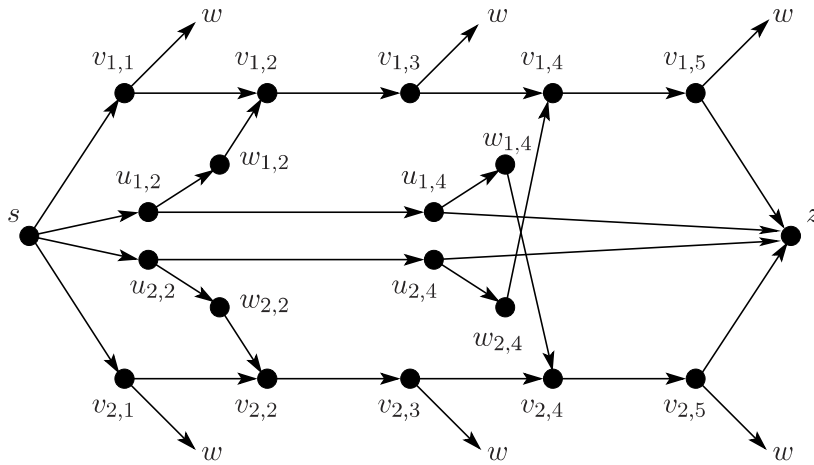
Figure 5: Network from Example 1, $|T_{even}| = 2$.

Question: What is the cause of the gap in the number of feasible vehicles obtained from the optimization problem and simulation procedure?

We suppose that the main reason is in the available information. When we use an optimization algorithm, we work with the whole network G. However, in the simulation procedure, at time $t \in T_{even}$, we only have information about edges from vertices $u_{i,t}$. These constraints do not allow us to find an optimal solution by the evaluation procedure. The open questions are: How do we design the charging infrastructure to be more succesful especially in cases where the constrained level of information is given? How can we spread the available information (from O-D matrices, probability models) to obtain a lower number of unfeasible vehicles?

4   CONCLUSIONS

The main idea of the work is the representation of the simpler version of Model 2 that works with flows in networks. In this approach, we obtain the optimal solution when we find the feasible flow with the minimum number of non-zero flows on the input edges.

We suppose that this approach will allow us to study the role of the available information in the charging process. For example, with this model, we are able to explain the gap between the results of optimization and the simulation procedure, which can be seen in Table 1. In optimization, we have all the information about the network. In the eveluation procedure, we have only the information on neighbourhoods of vertices that have actually been processed.

Our plan is to use this flow model to evaluate the dependency of the level of information about available charging points and suggest an appropriate system of assigning the jobs (journeys) and free charging points for vehicles.

Acknowledgements

References

[1] Dantzig, G.B., Fulkerson, D.R., (1954). Minimizing the number of tankers to meet a fixed schedule. Naval Res. Logist. Quart., 1, pp. 217–222

[2] Koháni, M., Czimmermann, P., Váňa, M., Cebecauer, M. and Buzna, Ľ. (2017). Designing Charging Infrastructure for a Fleet of Electric Vehicles Operating in Large Urban Areas. In: Proceedings of the 6th International Conference on Operations Research and Enterprise Systems - Volume 1:ICORES, pp. 360–368.

[3] Koháni, M., Czimmermann, P., Váňa, M., Cebecauer, M. and Buzna, Ľ. (2017). Location-scheduling optimization problem to design private charging infrastructure for electric vehicles. Operations Research and Enterprise Systems, ICORES 2017, Revised Selected Papers. Communications in Computer and Information Science, Springer. (Accepted for publication.)

[4] Nemhauser, G.L., Wolsey, L.A., (1988). Integer and Combinatorial Optimization. Wiley-Interscience series in discrete mathematics and optimization. ISBN 0-471-82819-X, pp. 65

An SD-based DST
for solving a volumes problem

Roman Gumzej, Bojan Rosi

Faculty of Logistics, University of Maribor, Mariborska cesta 7, SI-3000 Celje

e-mail: roman.gumzej@um.si, bojan.rosi@um.si

Abstract: Decision makers in supply chains are often faced with strategic decision-making challenges with only partial information on system structure and/or system operation. Negative consequences from making sub-optimal uninformed decisions can be avoided by utilizing decision support tools. They enable decision makers to make informed decisions based on rapidly changing data, not known in advance. By applying the holistic modelling approach to capture an extended supply chain at a strategic level of an enterprise within a decision support tool, the decision making problem may be efficiently addressed and solved. In this article system dynamics is applied to efficiently address a company's problem with shipping volumes.
Keywords: supply chain, management, system dynamics, decision support.

1 INTRODUCTION

Decision makers in supply chains (SC) are often faced with strategic decision-making challenges with only partial information on system structure and/or system operation. Hence, adopted decisions may have a negative impact on overall supply chain performance. These situations should be avoided and can be prevented by utilizing decision support tools (DST). DSTs enable business and/or organizational decision-making based on rapidly changing data, not known in advance. By applying the holistic modelling approach to capture the extended supply chain at a strategic level of an enterprise within a DST the decision making problem may be efficiently addressed and solved [1].

As a network of organizations, the SC is obliged to have clearly defined objectives. The main objective of the supply chain is the maximization of total profits derived across different echelons of the supply chain. The second most important objective, however, is the maximization of satisfaction of customers, expressed by the service level [2], also referred to as quality of service (QoS). While the majority of DST-based approaches deal with the first objective, the fulfilment of the second objective, in our opinion, is interrelated with it and as such should be considered in conjunction [3]. The set goals can only be attained by adequate supply chain configuration and proper use of management tools. The latter are addressed by supply chain management (SCM), being the process of planning, implementing and controlling the operations of the supply chain in an efficient way. SCM spans all movements and storage of raw materials, work-in-process inventory, and finished goods from the point of origin to the point of consumption [4].

SCM addresses a wide range of problems, which may be divided into three categories [5]:

1. supply chain infrastructure (network) design,

2. supply chain analysis and policy formulation,

3. supply chain planning and scheduling.

Within the process of modelling and solving the indicated problems, managers have to make decisions on different hierarchical levels. These decisions can be classified as strategic, tactical and operational, depending on their effects on the overall supply chain. Due to interdependence

among the three levels, SCM remains incapable of satisfactorily addressing many practical real-world problems at a strategic level. Let alone that these levels are difficult to tackle at any individual entity, due to inter-dependencies with other entities and their autonomous behaviour they are even more difficult to address. Hence, the majority of modelling approaches in DST provide invalid results [6]. For obtaining adequate solutions at the strategic level, decision-makers need comprehensive models to guide them in the decision-making process.

System dynamics (SD), being a simulation based modelling approach, which underpins DST and holistic modelling [7, 8], is widely used in logistics and supply chain management. SD-based models are mainly applicable in situations, where the type of flow elements can be unified (e.g. on transaction level) and the type and behaviour of their manipulation entities (supply chain nodes) does not change over time, and as such imposes no further restrictions to the simulation (e.g. [8, 9]).

In this article, an SD-based model of a multi-echelon supply chain, which may be used to analyse and fine tune the behaviour of supply chain nodes, is presented. First, the SC model and its transformation into the SD model are presented. Next, the SD model is presented with the demonstration of its behaviour and the way it acts as a DST system. Finally, the results of different simulation scenarios are analysed and a conclusion is drawn.

2 PROBLEM STATEMENT

A multinational company has 8 subsidiaries in several countries. 5 of them are production plants, while others are wholesalers of company's goods. There is a new product they wish to introduce to the market and due to its novelty they are only dealing with speculative numbers concerning sales volumes. Since the product is interesting, and the materials are easily obtainable within the EU, some of the company's subsidiaries could take over parts of the production volume in order to reduce shipping costs. On the other hand, existing storage facilities have limited capacity, so an appropriate supply chain strategy would be necessary to balance supply and demand. An appropriate DST tool is sought to balance the volumes and meet demands.

2.1 Supply chain structure

The company's headquarters and the largest production plant are situated in Slovenia. Its main production subsidiaries are in Germany, Poland, Hungary and Bosnia and Herzegovina. In addition to production plants, its wholesale companies are in Romania, Ukraine and Russia. All of them manage their own warehouses. The main production facility supplies its production subsidiaries with materials and they supply their own markets and the main company with finished products. The main production plant also supplies its subsidiaries and wholesale companies with finished products.

2.2 SD model

Our SD model (Figure 1) describes the company's SC and material flows between the entities. The primary production plant is denoted by SI for Slovenia. Its production subsidiaries, collectively denoted as SPs, are labelled DE for Germany, PL for Poland, HU for Hungary and BIH for Bosnia and Herzegovina. Its wholesale companies, collectively denoted as WHs, are labelled RU for Romania, UKR for Ukraine and RUS for Russia. Their nodes represent their warehouse facilities. Their incoming and outgoing flows correspond with their production and sales/distribution activities, respectively. To parameterise our model, production, shipping and consumption rate variables have been introduced at nodes' incoming and outgoing flows. Production rate ($prod\_rate\_i$ or $prodrate_i$) denotes the volume of production at node i. Shipping rate ($ship\_rate\_ij$ or $shiprate_{ij}$) denotes the proportion of node i's stock volume being shipped
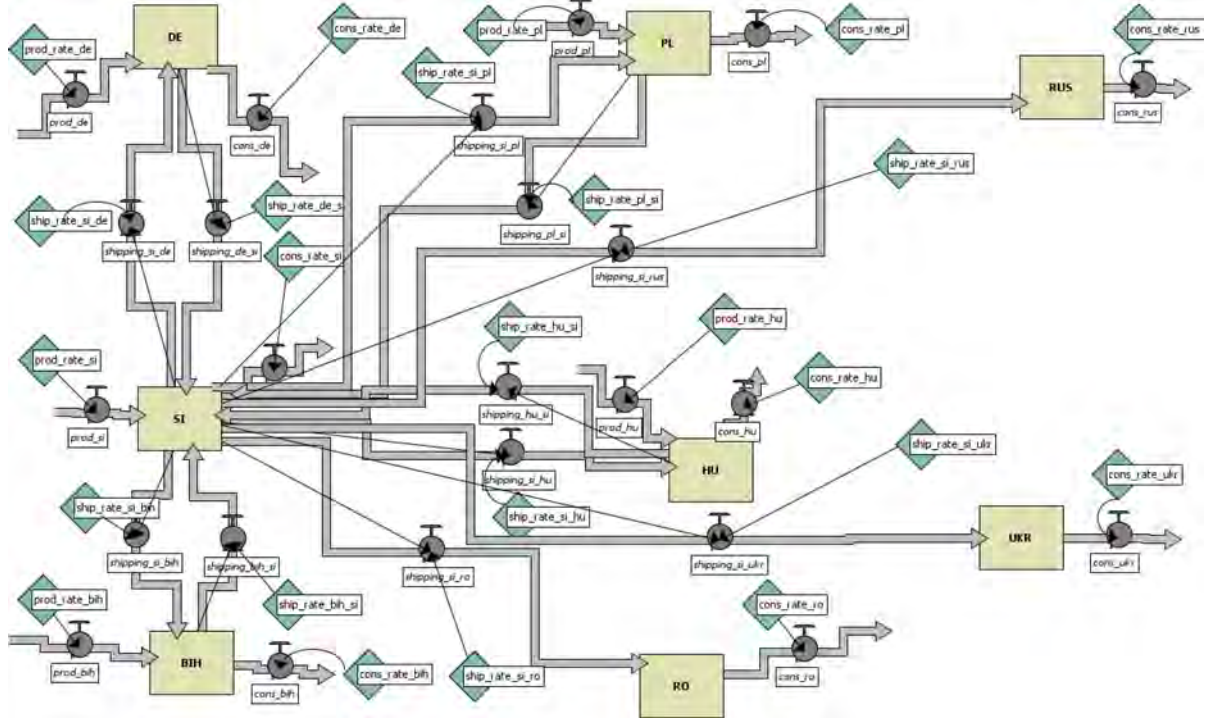
Figure 1: SD model layout

to node j. Consumption rate ($cons\_rate\_i$ or $consrate_i$) denotes the volume of consumption at node i.

To illustrate the combined effect of the corresponding parameters on the dynamics of our model, the following formulas represent the amounts of stock our primary production plant (1), subsidiary production plants (2) and wholesalers (3) at any given time instant t:

$$SI(t) = SI(0) + \int_0^t \ [prodrate_{si}(t) + \textstyle\sum_{sp} shiprate_{spsi}(t) - \sum_{sp} shiprate_{sisp}(t) \\ -x \cdot (consrate_{si}(t) + \textstyle\sum_{wh} shiprate_{siwh}(t))]dt$$

(1)

$$SP(t) = SP(0) + \int_0^t \ [prodrate_{sp}(t) + shiprate_{sisp}(t) \\ -x \cdot consrate_{sp}(t) - shiprate_{spsi}(t)]dt$$

(2)

$$WH(t) = WH(0) + \int_0^t [x \cdot shiprate_{siwh}(t) - consrate_{wh}(t)]dt$$

(3)

The rates of production, consumption and shipping transactions are considered on a common and regular daily basis ($dt = 1$ day), considering our supply chain policy. They occur simultaneously following the pre-determined schedule. x denotes the number of product components that form the final product. Since it is common practice to produce product parts at different plants, they may supply each other with component parts and provide finished products through their distribution channels.

The associated dashboard (Figure 2) serves as a DST system to manage production and storage volumes and balance them with the projected demands. The projected amounts of
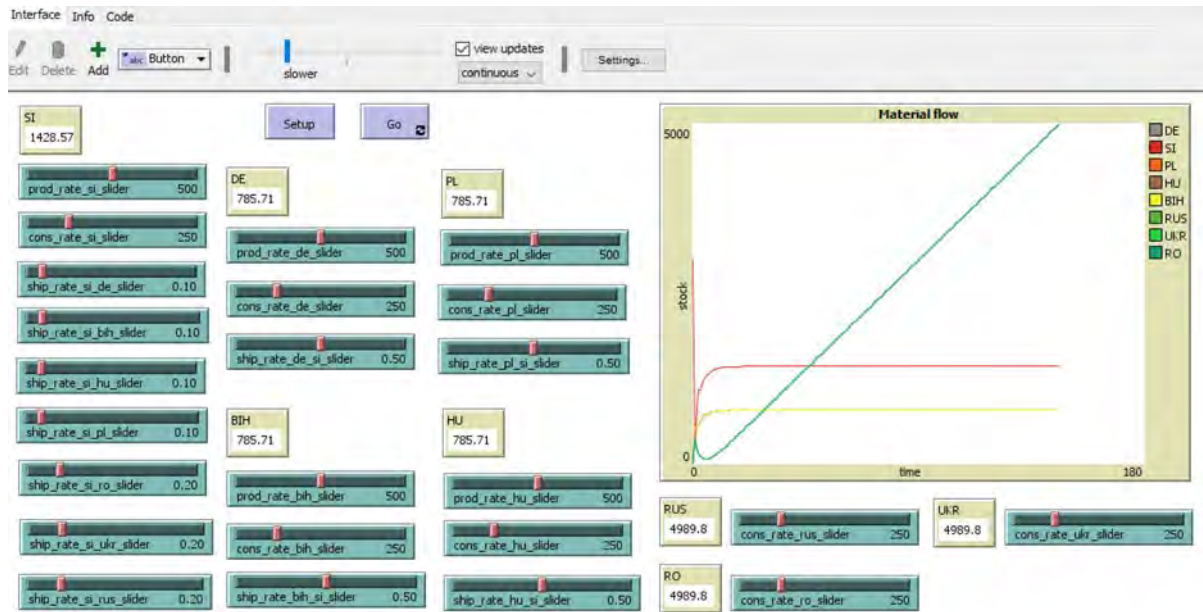
Figure 2: DST dashboard layout for our SD-model

products and product components, being produced, range from a few up to a thousand pieces per transaction. The amounts of products moving between facilities are given in percentages for each supply channel. The stock of production plants is considered inventory and may leave the production facilities through their wholesale channels. For each of the facilities the expected sales amounts are given as quantities of up to a thousand pieces.

2.3 Evaluation

Our DST system has been realised within the NetLogo [10] simulation environment. By utilising our DST system one can:

- Plan the layout of the SC.

- Manage the production and supply quantities.

- Determine the utilisation of distribution channels and the associated costs.

New nodes can be added or removed from our SC model, when needed. When integrated with the rest of the SC through distribution channels, they are automatically integrated with the existing material flows. Hence, planning new production or wholesale facilities is simplified. By parametrising the individual flows one can easily determine the production, consumption and distribution quantities. By adding dashboard items and linking them with the parameters of the SD-model one can utilise the DST tool to control them. By exporting the values of the produced graphs and monitor items from the DST dashboard, one can further analyse quantities and trends as well as determine the associated costs.

In our experiment we have concentrated ourselves on a time frame of one Year (365 days and the respective number of transactions). To evaluate our model we have tried three scenarios, which differ mainly in the number of component parts (x) of the finished product. Our intention was to foresee the load our supply chain could take and should bear considering the foreseen consumption values. To balance our stock values, the daily consumption has been set to 250 pieces at all sites for the entire experiment (see Figure 2).

Table 1: Scenario 1 - Stock volumes after one year

Facility	SI	DE	PL	HU	BIH	RUS	UKR	RO
Stock	1428	785	785	785	785	11239	11239	11239

Table 2: Scenario 2 - Stock volumes after one year

Facility	SI	DE	PL	HU	BIH	RUS	UKR	RO
Stock	769	487	487	487	487	19174	19174	19174

2.3.1 First scenario

Our first scenario is based on the given initial values of model's variables (see Figure 2) and represents the case in which our production sites are independent ($x = 1$) – they source locally to produce their products. After one year the stock volumes at our sites are presented in Table 1 and their dynamics in the graph of Figure 2.

2.3.2 Second scenario

Our second scenario represents the case in which our products are composed of 2 parts ($x = 2$) – the production sites must cooperate. After one year the stock volumes at our sites are given in Table 2 with similar dynamics as in the first scenario. We determined that, in order to ensure sustainability, we would have to increase production rates of our production sites to 750 pieces and increase shipment rates of subsidiary production sites to 0.67.

2.3.3 Third scenario

Our third scenario represents the case in which our products are composed of 3 parts ($x = 3$) – the production sites must cooperate. After one year the stock volumes at our sites are given in Table 3. Their dynamics oscillates around target values indicating instability, originating from frequent out-of-stock situations at the main production site. We determined that, in order to ensure sustainability, we would have to increase production rates of our production sites to the foreseen maximum production volume of 1000 pieces and increase shipment rates of subsidiary production sites to 0.75.

3 Conclusion

In this article, an SD-based model of a multi-echelon supply chain, which may be used to analyse and fine tune its behaviour, has been presented. According to available data on supply chain structure and the projected volumes for the new product, the quantities and costs can be effectively assessed. Since the model and the associated DST tool are flexible, any changes in supply chain structure and the volumes can easily be incorporated into the model and included in the analysis. Hence, we may conclude, that our SD model fulfils the demand for a simple to use and flexible DST tool, which may solve the current and future problems of supply chain management for our use case.

Considering the capacity of our supply chain, we have determined that we have reached it with three components. If there were more components, our production volumes would not

Table 3: Scenario 3 - Stock volumes after one year

Facility	SI	DE	PL	HU	BIH	RUS	UKR	RO
Stock	1315	368	368	368	368	47470	47470	47470

suffice. In order to prevent the initial out-of stock situation at wholesale nodes RUS, UKR and RO an initial stock quantity of 3000 units at node SI was determined as sufficient. After a month of production, our non-production sites continue to gain some reserves which stock-up. After this initial time period we could slightly increase the consumption at non-production sites for short time-frames. In case of larger and longer increases in demand quantities, however, their stock volumes would decrease very fast.

References

[1] Rosi, B. and Mulej, M. (2006). The dialectical network thinking - a new systems theory concerned with management. Kybernetes, 35(7/8): 1165-1178.

[2] Zhang, G., Shang, J. and Li W. (2011). Collaborative production planning of supply chain under price and demand uncertainty. European Journal of Operational Research, 215(3): 590-603.

[3] Gumzej, R. and Gajšek, B. (2011). Introducing Quality of Service Criteria into Supply Chain Management for Excellence. International Journal of Applied Logistics (IJAL), 2(1): 1-16.

[4] Melo, M.T., Nickel, S. and Saldanha-da-Gama F. (2009). Facility location and supply chain management - A review. European Journal of Operational Research, 169(2): 401-412.

[5] Shah N. (2005). Process industry supply chains: Advances and challenges. Computers & Chemical Engineering, 29(2005): 1225-1236.

[6] Georgiadis, P., Vlachos, D. and Iakovou, E. (2005). A system dynamics modeling framework for the strategic supply chain management of food chains. Journal of Food Engineering, 70(3): 351-364.

[7] Tako, A.A. and Robinson S. (2012). The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decision Support Systems, 50(4): 802-815.

[8] Intihar M.(2013). System Dynamics Modeling in Supply Chain Management. Eichler G, Gumzej R, editors. Proceedings of the 13th International Conference on Innovative Internet Community Systems and the International Workshop on Autonomous Systems, May/June, Hagen, Germany. Fortschr.-Ber. VDI 10(826): 241-252. Düsseldorf: VDI Verlag.

[9] Angerhofer B.J. and Angelides, M.C. (2000). System dynamics modelling in supply chain management: research review. Joines JA, Barton RR, Kang K, Fishwick PA editors. 2000 Winter Simulation Conference Proceedings, December 10-13, Orlando, FL, USA. IEEE, 1: 342-351.

[10] Wilensky U. NetLogo. (1999-2016). Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. https://ccl.northwestern.edu/netlogo/index.shtml [Accessed 18/04/2017].

COLLECTIVE FAIRNESS IN EMERGENCY SYSTEM DESIGNING

Jaroslav Janáček
University of Žilina, Faculty of Management and Informatics,
Univerzitná 1, 010 26 Žilina, Slovak Republic,
jaroslav.janacek@fri.uniza.sk

Lýdia Gábrišová
University of Žilina, Faculty of Management and Informatics,
Univerzitná 1, 010 26 Žilina, Slovak Republic,
lydia.gabrisova@fri.uniza.sk

Abstract: The paper deals with fair distribution of facilities in the set of service centres, which provide associated users' clusters with service. Transportation performance per one facility is considered as the basic notion of the models. According to min-max fairness scheme, objective of the first model is to minimize the maximal value of transportation performance among the service centers. The second scheme of fairness consists in minimization of dispersion of the values of transportation performance. Two models of fair deployment are presented and the results obtained by the associated methods based on branch-and-bound principle are compared. Furthermore, another method based on dynamic programming approach is presented as an alternative to the solving technique used for minimization of dispersion.

Keywords: collective fairness, emergency service system, facility distribution, transportation performance, min-max fairness, dispersion fairness

1 INTRODUCTION

The emergency service system services randomly emerging demands as accidents, fires and similar events. The users – population of a serviced region are spread over associated geographical area, where they are concentrated in a finite set of dwelling places as towns, villages and hamlets. Emergency service is provided from a finite number of service centers deployed in the serviced region, where each center is equipped with a different number of indivisible facilities. It is generally assumed that the average time of service accessibility is prevailing criterion of the system performance [1, 2, 4, 5, 8]. Thus each potential user is assigned to the nearest service center. It follows that population of user clusters associated with the individual service centers is only matter of time-distances among users' and center locations. This distance based serviced area partitioning generates user clusters with unequally distributed population. The limited capacity of assigned facilities together with the randomly coming demands may cause that the current demand cannot be served from the nearest service center due to its occupancy by previously occurred demands. The center occupancy incurs that the newly occurred demand must either wait or be serviced from some more distant service center. Both eventualities mean that the service accessibility gets worse, which is considered unfair by those system users, which belong to the dense populated districts, where the demand frequency is extremely high [3].

   In this paper, we concentrate our effort on reassigning the facilities to the individual service centers, to minimize the difference in transportation performance per one facility among the individual centers. First, we formulate two models of the fair facility allocation problem and compare the resulting facility deployments. The first of the models minimizes the maximal transportation performance per one facility in the set of the users´ clusters and the second one minimizes dispersion of transportation performance values per one facility. The both problems are formulated and solved as integer linear programming problems by branch-and-bound method. In addition, we suggest another solving tool based on dynamic programming approach for the second problem, which minimizes transportation performance

dispersion. All suggested approaches are tested and compared by solving a series of benchmarks derived from real medical emergency service systems performing in the Slovak Republic. The remainder of the paper is laid out as follows. Next section comprises the both models of the fair facility distribution problem and the third section contains an explanation of the dynamic programming approach. The fourth section is devoted to the comparison of the suggested methods and obtained findings are summarized in the fifth section.

2 THE FAIR FACILITY DISTRIBUTION PROBLEM

An emergency service system is represented by a set I_1 of w service center locations, by a set J of user locations, where volume b_j of demand is associated with the user location j. It is assumed that the shortest path from nearest service centre to serviced users location is used. Symbol t_{ij} denotes the time length of the shortest path from i to j in the transportation network. Let us denote $ass(j) \in I_1$ the center, which provides user j with service. Then the transport performance P_i of the user cluster associated with service center $i \in I_1$ can be computed according to (1).

$$P_i = \sum_{\substack{j \in J \\ ass(j)=i}} b_j t_{ass(j),j} \tag{1}$$

Let p denote the total number of facilities to be deployed. Obviously $p \geq w$. We assume that at least one facility must be located at each center i and thus only $v=p-w$ remaining facilities must be deployed. To describe the models of fair facility distribution problem, we introduce integer variables $y_i \in Z^+$ for each service center $i \in I_1$. The variable y_i models a decision on the number of additional facilities allocated at the center i. It follows $y_i \leq v$. Furthermore, we introduce an auxiliary variable $h \geq 0$ to model the upper bound of transportation performance per one facility. Then, a non-linear model of min-max fair facility distribution problem can be stated as follows.

$$\textit{Minimize } h \tag{2}$$

$$\textit{Subject to } P_i /(y_i +1) \leq h \textit{ for } i \in I_1 \tag{3}$$

$$\sum_{i \in I_1} y_i = v \tag{4}$$

$$y_i \in Z^+ \textit{ for } i \in I_1 \tag{5}$$

The problem minimizing dispersion of transportation performance per one facility can be formulated in the following way.

$$\textit{Minimize } \sum_{i \in I_1} (P_i)^2 /(y_i +1) \tag{6}$$

$$\textit{Subject to } (4) \textit{ and } (5)$$

The both models can be linearized by following adjustment, where we define a range $V=\{0, \ldots, v\}$ and introduce constants Q_{ik} and R_{ik} according to (7).

$$Q_{ik} = P_i /(k+1) \textit{ and } R_{ik} = (P_i)^2 /(k+1) \textit{ for } i \in I_1, k \in V \tag{7}$$

We introduce a family of binary variables x_{ik} for $i \in I_1$ and $k \in V$, where the variable takes the value of one if and only if $y_i=k$. After these preliminaries the model (2)-(5) can be reformulated as follows.

$$\textit{Minimize } h \tag{8}$$

$$\text{Subject to } \sum_{k \in V} Q_{ik} x_{ik} \le h \text{ for } i \in I_1 \qquad (9)$$

$$\sum_{k \in V} x_{ik} = 1 \qquad (10)$$

$$\sum_{i \in I_1} \sum_{k \in V} k x_{ik} = v \qquad (11)$$

$$x_{ik} \in \{0, 1\} \text{ for } i \in I_1, \ k \in V \qquad (12)$$

The model (6), (4), (5) obtains the following form.

$$\text{Minimize } \sum_{i \in I_1}{}' \sum_{k \in V}{}' R_{ik} x_{ik} \qquad (13)$$

$$\text{Subject to } (10) - (12)$$

3 DYNAMIC PROGRAMMING APPROACH

The min-sum form of the objective function (6) together with the simple form of the only one structural constraint (4) evoked us an idea to formulate the problem as a dynamic programming problem. Timeline of the associated dynamic process will be modelled by a discrete set of instants corresponding to an orders in the sequence of decisions on the facility allocation at the individual service centres. Thus the variable y_i models a decision on the number of additional facilities allocated at the center, which is processed at instant $i = 1, \ldots,$ w. The instant i will be used also as a subscript used in description of constants associated with the corresponding service center. The state s_i of the allocation process at the instant i is defined as the total number of facilities allocated by the decisions forgoing the decision y_i. This way, we can state the relation between s_i and s_{i+1} in the form of $s_{i+1} = s_i + y_i$. It is obvious that the set S of all state values is $\{0, \ldots, v\}$. The decision y_i contributes to the objective function value by the increment $f_i(s_i, y_i) = (P_i)^2 / (y_i + 1)$. We define initial values of Bellman's function B_w and control function Z_w at the instant w according to (14).

$$B_w(s_w) = (P_w)^2 / (v - s_w + 1) \text{ for } s_w \in S$$
$$Z_w(s_w) = v - s_w \text{ for } s_w \in S \qquad (14)$$

Then we can define the iterative process of Bellman's function computation at the instant i by (15).

$$B_i(s_i) = \min\{f_i(s_i, y_i) + B_{i+1}(s_i + y_i) : y_i \in \{0, \ldots, v - s_i\}\} \text{ for } s_i \in S$$
$$Z_i(s_i) = \arg\min\{f_i(s_i, y_i) + B_{i+1}(s_i + y_i) : y_i \in \{0, \ldots, v - s_i\}\} \text{ for } s_i \in S \qquad (15)$$

The iterative process is performed step by step for $i = w-1, \ldots, 1$. When the iterative process is finished, then $B_1(0)$ gives the optimal value of (6) and the optimal values of y_i can be obtained by the following recursive process. Set $s_1 = 0$ and $y_1 = Z_1(s_1)$, then set $s_{i+1} = s_i + y_i$ and $y_{i+1} = Z_{i+1}(s_{i+1})$ for $i = 1, \ldots, w-1$.

4 NUMERICAL EXPERIMENTS

The goal of these numerical experiments is to compare the two approaches to the fair facility distribution problem. The approaches are mentioned above as min-max and min-dispersion problems and they will be denoted as **mM** and **mD** respectively. The next goal is verification and comparison of the dynamic programming approach to the standard branch-and-bound approach to the min-dispersion problem.

To achieve the goals, we performed the series of numerical experiments. To solve the problems described in the previous sections, the optimization software FICO Xpress 7.9 (64bit, release 2015) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 4510U processor with the parameters: 2.00 GHz and 8 GB RAM. The dynamic programming procedure was programmed in language Mosel and it was run also under the FICO Xpress.

The used benchmarks were derived from the real emergency health care system, which has been originally implemented in whole region of the Slovak Republic. All cities and villages with corresponding population b_j were taken into account as users. The coefficients b_j were rounded to hundreds. This system covers demands of all communities - towns and villages spread over the region by given number $p=273$ of ambulance vehicles. In the benchmarks, the set of communities represents both the set J of users' locations, which cardinality is 2916. The original set I_1 of center locations has the cardinality $w= 208$. Matrix $\{t_{ij}\}$ of the time distances from individual center locations to individual users' locations was obtained from the associated Slovak road network.

Due to the lack of common benchmarks, the other problem instances used in our computational study were created in the following way. We chose a value of w from the range {188, 192, 196, 200, 204, 208, 212, 216, 220, 224, 228} and solved associated weighted p-median problem for the set of possible center locations equal to the set J. This optimization performed according to [6], [7] provided us with various lists I_1 of center locations and completed the set of benchmarks.

The results of the first portion of experiments are presented in Table 1, where each row corresponds to one benchmark denoted as follows. Symbol "***Orig***" stands for original service center deployment and symbols ***Der188***, …, ***Der228*** stand for the above-mentioned derived benchmarks, suffix of which corresponds to the number w of service centers.

Table 1: Results of numerical experiments comparing the approaches minimizing dispersion (***mD***) and maximal value of transportation performance per one facility (***mM***)

Benchmark	mD			mM			Difference		
	Disp	*Max*	*cT*[s]	*Disp*	*Max*	*cT*[s]	*DifD*[%]	*DifM*[%]	*Manh*
Orig	96268	1361	0.20	98349	1245	0.48	-2.1	9.3	6
Der188	31987	1041	0.24	34692	941	0.52	-7.8	10.6	14
Der192	33174	1041	0.23	36119	941	0.40	-8.2	10.6	16
Der196	35880	1041	0.22	40028	950	0.53	-10.4	9.6	18
Der200	39111	1041	0.22	43657	950	0.47	-10.4	9.6	20
Der204	39757	1037	0.21	44078	950	0.46	-9.8	9.2	16
Der208	42366	1041	0.20	45748	965	0.37	-7.4	7.9	14
Der212	45758	1045	0.21	48326	988	0.38	-5.3	5.8	12
Der216	48502	1045	0.19	50565	992	0.35	-4.1	5.3	8
Der220	51411	1043	0.19	52980	993	0.31	-3.0	5.0	4
Der224	55532	1089	0.17	56734	1043	0.27	-2.1	4.4	4
Der228	58353	1110	0.16	60407	1059	0.26	-3.4	4.8	6

The table is divided into three parts called *mD*, *mM* and *Differences*. The parts *mD* and *mM* are organized in the same way and contain three columns denoted by *Disp, Max* and *cT*, where column *Disp* contains values of dispersion of transportation performance per one facility for solutions of the given instances, column *Max* contains maximal value of transportation performance per one facility and *cT* contains computation times of the used approaches applied to the particular benchmarks. The third part also consist of three columns. They are denoted by *DifD, DifM* and *Manh*. The columns *DifD* and *DifM* contain differences between values of dispersions and maxima of the compared approaches respectively. The values are given in percentage, where results of *mM* approach are taken as a base (100%). The column *Manh* contains the Manhattan distance of the solutions (vectors *y*) obtained by the compared approaches.

The plotted data show that there is a significant difference between min-max and min-dispersion approaches. Nevertheless, both approaches represent some kind of collective fairness and none of them should be neglected. This feature can be demonstrated comparing the current facility distribution in the original set I_1 of service centres (**Orig**). The comparison can be performed for the instance **Orig** only, due to only the current facility distribution is at disposal. In Table 2, we present maximal values of transportation performance (*Max*) and dispersion *(Disp)* associated with the current state. This part of the table is denoted as **Current**. The parts **mD** and **mM** give the associated values of differences from the values obtained by application of above-mentioned approaches. The values *DifD* and *DifM* are given in percentage of improvement (decrease), where the values presented in the part **Current** serve as the base.

Table 2: Comparison of the current facility distribution *(Current)* and the distributions obtained by approaches minimizing dispersion (*mD*) and maximal value of transportation dispersion per one facility (*mM*)

	Current		mD			mM		
	Disp	Max	DifD[%]	DifM[%]	Manh	DifD[%]	DifM[%]	Manh
Orig	176337	2526	45.4	46.1	46	44.2	50.7	44

The second goal of our experiments was verification and comparison of the dynamic programming approach to the standard branch-and-bound approach to the *mD* problem. We performed the experiments with the same set of benchmarks presented in Table 1, and we obtained the completely same results for approach *mD* and dynamic programming approach. The only difference concerns computational time, where the dynamic programming approach proved to be at least two times quicker.

5 CONCLUSIONS

We suggested two models of collective fairness accompanied by exact solving methods, which enable to optimize facility distribution in a given set of service centers. We showed that the methods are able to solve real-world instances of the problem in very short computational time. These methods may be a significant part of a tool for complex emergency service system designing.

Future research will be aimed at usage of convexity of the separated nonlinearities in the objective function to speed up the computational process both of linear programming and dynamic programming approaches. Also generalization of the transportation performance model deserves attention in the next research.

References

[1] Brotcorne, L. and Laporte, G. and Semet, F. (2003). Ambulance location and relocation models. European Journal of Operational Research, 147: 451–463.

[2] Chanta, S. and Mayorga and M.E., McLay, L.A. (2011). Improving emergency service in rural areas: a bi-objective covering location model for EMS systems. Annals of Operations Research [online] DOI 10.1007/s10479-011-0972-6.

[3] Janáček, J. and Gábrišová, L. (2009) A two-phase method for the capacitated facility problem of compact customer sub-sets. Transport: Journal of Vilnius Gediminas Technical University and Lithuanian Academy of Sciences, 24(4): 274-282.

[4] Jánošíková, Ľ. (2007). Emergency medical service planning. Communications: Scientific Letters of the University of Zilina, 9(2): 64-68.

[5] Jánošíková, Ľ. and Žarnay, M. (2014). Location of emergency stations as the capacitated p-median problem. International scientific conference: Quantitative Methods in Economics-Multiple Criteria Decision Making XVII, Virt, Slovak Republic.

[6] Kvet, M. (2014). Computational study of radial approach to public service system design with generalized utility. Digital Technologies 2014: the 10th International IEEE conference, Zilina, Slovak Republic.

[7] Kvet, M. (2015). Exact and heuristic radial approach to fair public service system design. Information and Digital Technologies 2015: IEEE catalog number CFP15CDT-USB, Zilina: Slovak Republic.

[8] Marianov, V. and Serra, D. (2004). Location problems in the public sector, in facility location. Applications and theory (by Drezner Z (ed.) et al.), Berlin, Springer: 119-150.

OPTIMIZING THE LOCATION-ALLOCATION PROBLEM OF BIKE SHARING STATIONS: A CASE STUDY IN GAZIANTEP UNIVERSITY CAMPUS

Eren Özceylan
Department of Industrial Engineering, Gaziantep University, 27300, Gaziantep, Turkey
erenozceylan@gmail.com

Süleyman Mete
Department of Industrial Engineering, Munzur University, 62000, Tunceli, Turkey
suleyman489@gmail.com

Zeynel Abidin Çil
Department of Manufacturing Engineering, University of Batman, 72060, Batman, Turkey
cilzeynelabidin@gmail.com

Abstract: A growing number of cities are implementing bike-sharing programs to increase bicycle use. One of the key factors for the success of such programs is the location of bike stations in relation to potential demand. In this paper, potential locations of bike sharing stations in Gaziantep University campus are investigated to provide optimal utilization for users. To do so, first of all, 20 demand points and 20 potential bike stations are determined. Second, set covering mathematical model is handled to determine coverage capability of potential bike stations. Finally, P-center and P-median mathematical models are applied to set up potential stations and to assign demand points to the opened stations so that the total walking distance can be minimized. Computational experiments prove that our approach can find new potential bike stations which cover all demand points.

Keywords: Bike sharing station; Location-allocation models; Integer programming; Gaziantep.

1 INTRODUCTION

Bike sharing has evolved significantly since its inception in 1965, the first public-use bicycles, with the famous "White Bicycles" system in Amsterdam. It is proposed that 20,000 bicycles be painted white and distributed for pick-up and drop-off anywhere in the city center, free of charge. The next attempt at a bike sharing system occurred in La Rochelle, France, in 1993, which offered a free, but more regulated, program that allowed the public to check out bicycles for two hours. Extensive bike sharing system is Vélib' in Paris, consists of a network of 1800 stations (a station every 300 m), and more than 20,000 bicycles are always available.

In shortly, today, more than 600 cities around the globe have their own bike sharing systems, and more programs are starting every year. There are an estimated 822,000 bikes in operation around the world – so China has more bikes than all other countries combined. Detailed lists for the year of 2014 in terms of countries and cities are illustrated in Figure 1.
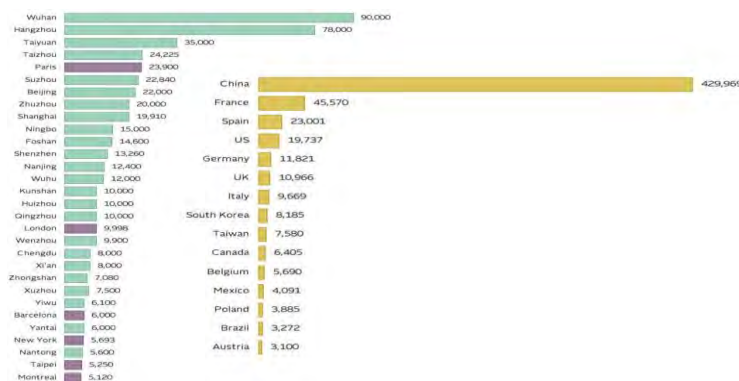
Figure 1: Number of bike share bikes per country and city (bikesharingworld.com)

In Turkey, traffic is at the top of the most important problems as in all developed or developing countries. That's why the use of bicycles has started to be used in many cities of Turkey. Only Istanbul, Kocaeli, Çanakkale, İzmir, Antalya, Konya and Kayseri out of 81 cities launched a bike sharing system in Turkey (Figure 2). In Turkey, even if using the bicycle is a sport activity, there is very small place in our lives. According to Turkish Statistical Institute, individuals have spent their free time distribution of sportive activities to evaluate when examined; 2.2% of them were riding bicycles. In order to removing this situation and solving the other problems, we need to promote using bicycles.

Figure 2: Cities in Turkey with bike sharing system (bikesharingworld.com)

Research on bike sharing systems has focused on site selection, bike dispatch, operations management, as well as user behaviour sand preferences [1, 2, 3, 4]. We focused on site selection of bike stations in this study. Because one of the keys to the success of bike sharing programs is the location of bike stations and their relation to trip demand. To gain user acceptance, the distance between stations and the origins and destinations of trips should be small, and the distance between the stations themselves should be appropriate for transport by bicycle. The bike sharing system can be usually set up to appropriate places by municipalities, private companies or universities.

In the case of bike sharing stations, the literature reports different approaches to tackling the location of the stations with facility location models. As an early work, Lin and Yang [3] addresses the strategic planning of public bicycle sharing systems with service level considerations. Later, a mixed-integer linear program performed through a heuristic that optimizes the location of shared bike stations is presented by Martinez et al. [5], assuming a fleet size and bicycle relocation calculation for a regular operating day. Besides mathematical modelling, a simulation–optimization method that optimizes the location of public bicycle stations is proposed by Romero et al. [6]. To consider spatial information, García-Palomares et al. [7] propose a GIS-based methodology. They estimate the potential trip demand and its spatial distribution and the location of the stations. Ghandehari et al. [8] present a study to find the best locations of bicycle stations through goal programming and multi-criteria decision making techniques. Liu et al. [9] propose an artificial neural network based prediction model for station demand and balance prediction, then an optimization problem aiming at maximizing station demand and minimizing the number of unbalanced stations is addressed and solved using a genetic algorithm. One of the current studies is proposed by Frade and Ribeiro [10]. They present an optimization method to design the bike sharing system such that it maximizes the demand covered and takes the available budget as a constraint. Finally, Wang et al. [11] build a spatial-temporal analysis model, then adopted the GIS to identify hot spot lacking-bikes and/or lacking-bike racks, and subsequently applied retail location theory to determine potential locations for rental stations.

Although there have been many studies which considers bike sharing systems as aforementioned above, there is still a gap to be filled on location and allocation of bike

stations. In view of this, set covering, P-median and P-center models are applied to locate and allocate bike stations in Gaziantep University campus. This paper contributes to the literature in several ways: (*i*) application of three popular location-allocation models hierarchically –to the best knowledge of the authors it is the first application on bike sharing site selection–, and (*ii*) providing a case study which is in a university campus firstly.

2 CASE STUDY

2.1 Study area and data

The campus of Gaziantep University is considered as the case study. Gaziantep University is located in Gaziantep and it has almost 40,000 population including students, administrative and academic staff on an area with 3,113,084 m^2 (Figure 3).

Figure 3: Gaziantep University campus

In study area, 20 demand points and 20 potential bike station sites are determined as point features and campus roads are vectorized as line features via geographic information system (GIS) with ESRI ArcGIS 10.2 software. While demand points are the places where students/personnel are located mostly, potential sites for bike stations are the places which are closed to demand points and available for possible infrastructure. Selected demand points and number of the student are shown in Table 1. In this table, number of student column shows the number of demanding students of each stop.

Table 1: Demand points.

Name of the building	No	Demand	Name of the building	No	Demand
Congress center	D1	275	Faculty of Art and Sciences	D11	300
Sport Center	D2	450	Vocational Schools	D12	175
Dormitory	D3	600	Medicine	D13	120
Cafeteria	D4	500	Theology	D14	100
Library	D5	550	Culture Center	D15	150
Market	D6	500	Department of Civil Engineering	D16	200
Techno city	D7	450	Conservatory of Turkish Music	D17	120
Department of Mechanical Engineering	D8	150	Department of Education	D18	200
Department of Electrical Engineering	D9	200	Student affairs	D19	200
Faculty of Economics	D10	150	Department of Food Engineering	D20	130

In this study, two kinds of GIS data, demand/station points as a point layer and roads as a line layer, are used. Road data is also used as network data set in GIS environment. For this reason, at first, university road map is collected as line data. Then, line-shape road layer is used to generate network between all points. Figure 4 shows the road network of Gaziantep

University and demand/station points. Distances between demand points and potential bike stations calculated by GIS and are available upon request.

2.2 Application of location-allocation models

In this section, the location-allocation models are used in this study to ensure the optimal distribution of the pharmacy warehouses to hospitals and pharmacies. Due to page limit, formulations of set covering, P-median and P-center problems are not given in this study. For details, reader can refer to the papers [12, 13 and 14]. It is noted that all runs were completed on a server with 1.8 GHz Intel Core processor and 4 GB of RAM. The computation time required to solve the model using the GAMS-CPLEX solver is less than 10 CPU seconds for all problems.

2.2.1 Results of set covering problem
The set covering model is primarily solved with five different coverage areas in the range of 300 to 700 meters. The results are illustrated in Figure 5.

Figure 4: Demand points (left-side) and potential station locations (right-side) in Gaziantep University campus

As can be seen from Figure 5, all demand points are reached by bike stations in all coverage areas. Only opened bike stations are shown in Figure 5. As expected, increasing the coverage area decreases the number of bike stations that need to be opened. For instance, while totally 11 bike stations are opened under 300m access distance, this number is reduced to three (stations of 6, 15 and 18) under 700m limit. Due to its location, 18th bike station is the only station which is preferred in all solutions.

Figure 5: Results of set covering model with different coverage areas

2.2.2 Results of P-median and P-center problems
After showing the effects of coverage areas on potential bike stations, we implement P-median and P-center models to assign potential bike stations to demand points so that the total walking distance is minimized. P-median model is implemented assuming the demands

are different as given in Table 1. We apply the P-median model for the demand points by setting 1 to 10 values for p. The results of P-median problem are given in Table 2. To make a fair comparison with P-center problem, total walked man-distance (objective function of P-median) and MaxL (objective function of P-center) values for each model are provided.

Table 2: Results of P-median problem with different *p* values.

	P=1	*P=2*	*P=3*	*P=4*	*P=5*
MaxL **Value (m)**	1313.50	1307.70	740.74	702.26	672.56
Walked Distance (man*km)	3423.61	2519.17	1959.43	1486.69	1216.49
Opened Stations	13	6-13	6-14-18	5-7-14-18	4-5-6-14-18
	P=6	*P=7*	*P=8*	*P=9*	*P=10*
MaxL **Value (m)**	546.18	546.18	546.18	403.22	403.22
Walked Distance (man*km)	1040.10	929.14	855.24	792.77	746.90
Opened Stations	4-5-6-13-15-18	4-5-6-11-13-15-18	2-5-6-7-11-13-15-18	2-5-6-7-10-11-13-15-18	2-5-6-7-10-11-12-13-15-18

According to Table 2, all P-median problems are solved optimally. As expected, increasing the number of *p* decreases the total distance between bike stations and demand points. Results in Table 2 show that increasing the number of available bike stations from 1 to 10, decreases the total travelled distance by 78.18%. Another outcome can be seen from Table 2 that 13th bike station is the closest station to all demand points.

In addition to considered set covering and P-median problems, P-center problem is also investigated. Table 3 presents the results of P-center problem for different *p* values.

Table 3: Results of P-center problem with different *p* values.

	P=1	*P=2*	*P=3*	*P=4*	*P=5*
MaxL **Value (m)**	1313.50	1014.64	740.74	622.68	458.84
Walked Distance (man*km)	3423.61	2868.80	2120.50	1629.40	1488.50
Opened Stations	13	8-18	6-14-18	3-5-13-18	4-5-8-15-18
	P=6	*P=7*	*P=8*	*P=9*	*P=10*
MaxL **Value (m)**	427.57	410.47	364.36	317.30	302.34
Walked Distance (man*km)	1107.10	1167.50	1033.40	847.21	800.76
Opened Stations	4-5-8-13-17-18	4-5-6-10-13-17-18	2-4-5-6-10-12-17-18	2-4-5-6-10-11-13-17-18	2-4-5-6-10-11-13-17-18-19

According to Table 3, increasing the number of bike stations to be opened decreases the longest distance between demand points and stations. While the longest distance between 13th bike station (*p*=1) and demand points is 1313.50m, it decreases to 302.34m with a 76.98% improvement in the situation of *p*=10. If two models are compared, it is clear that P-median model provides less walked distance than the results of P-center model in all situations.

Obtained results above may seem as a disadvantage for the campus planners or the decision-making authority, which might be looking for the site selection that minimizes distance. However, in our opinion, this is advantageous in that it provides the decision-maker many alternatives to choose from. For instance, if administration of campus wants to reach a bike station within approximately 400m, they have three options. According to the set covering problem, the first option is to open 8 bike stations which are 3rd, 5th, 8th, 10th, 14th, 17th, 18th and 19th stations. However, according to P-median problem, the second option is to open 9 bike stations which are 2nd, 5th, 6th, 7th, 10th, 11th, 13th, 15th and 18th stations. In this case, total travelled distance is 792.77 man-km. According to P-center problem, 7 bike stations which are 4th, 5th, 6th, 10th, 13th, 17th and 18th are enough for 400m limit as the last option. But at this time, total walked distance is increased to 1167.50 man-km. Instead of

deciding on a solution, we preferred to provide different alternative solutions for the campus planners.

3 CONLUSION

This study is designed to find the number and locations of the bike stations that need to establish in the Gaziantep University campus. To do so, three of the most common solutions in location-allocation models have been tested: set covering, P-median and P-center models. The P-median solution seems to be more advantageous in terms of accessibility because it generates a more uniform coverage. We believe that the methodology outlined here can provide university administrators with good insight into where bike-sharing stations should be located, and therefore it contributes significantly to the future planning of bike-sharing systems.

Interaction with public transport services, usage of electric bicycles within a framework of internet of things and the cost of the system should be considered in the future studies.

References
[1] Sallis, J.F., Frank, L.D., Saelens, B.E., Kraft, M.K., 2004. Active transportation and physical activity: opportunities for collaboration on transportation andpublic health research. Transportation Research Part A: Policy and Practice 38 (4), 249–268.
[2] Martens, K., 2007. Promoting bike-and-ride: the Dutch experience. Transportation Research Part A: Policy and Practice 41 (4), 326–338.
[3] Lin, J.R., Yang, T.H., 2011. Strategic design of public bicycle sharing systems with service level constraints. Transportation Research Part E: Logistics and Transportation Review47(2), 284–294.
[4] Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transportation Research Part A: Policy and Practice 46 (10), 1730–1740.
[5] Martinez, L.M., Caetano, L., Eiró, T., Cruz, F., 2012. An optimisation algorithm to establish the location of stations of a mixed fleet biking system: Anapplication to the City of Lisbon. Procedia - Social and Behavioral Sciences54, 513–524.
[6] Romero, J., Ibeas, A., Moura, J., 2012. A simulation–optimization approach to design efficient systems of bike-sharing.Procedia - Social and Behavioral Sciences 54, 646–655.
[7] García-Palomares, J.C, Gutiérrez, J, Latorre, M., 2012.Optimizing the location of stations in bike-sharing programs: A GIS approach.Applied Geography 35, 235–246.
[8] Ghandehari, M., Pouyandeh, V.H., Javadi, M.H.M., 2013, Locating of bicycle stations in the city of Isfahan using mathematicalprogramming and multi-criteria decision making techniques, International Journal of Academic Research in Accounting, Finance and Management Sciences 3 (4),18–26.
[9] Liu, J., Li, Q., Qu, M., Chen, W., 2015, Station site optimization in bike sharing systems, IEEE International Conference on Data Mining, 883–888.
[10] Frade, I., Ribeiro, A., 2015, Bike-sharing stations: A maximal covering location approach, Transportation Research Part A: Policy and Practice 82, 216–227.
[11] Wang, J., Tsai, C.H., Lin, P.C., 2016, Applying spatial-temporal analysis and retail location theory topubic bikes site selection in Taipei, Transportation Research Part A: Policy and Practice94, 45–61.
[12] Beasley, J.E., 1987. An algorithm for set covering problem. European Journal of Operational Research 31 (1), 85–93.
[13] Teixeira, J.C., Antunes, A.P., 2008. A hierarchical location model for public facility planning. European Journal of Operational Research 185 (1), 92–104.
[14] Narula, S.C., 1986. Minisum hierarchical location-allocation problems on a network: A survey. Annals of Operations Research 6 (8), 255–272.

THE FUTURE CUSTOMER DEMAND IN LOCATION-ROUTING PROBLEM

Engin Pekel

Yildiz Technical University, Faculty of Machine, Department of Industrial Engineering, A-622
34300, Turkey
E-mail: pekelc@hotmail.com

Selin Soner Kara

Yildiz Technical University, Faculty of Machine, Department of Industrial Engineering, A-631
34300, Turkey
E-mail: ssoner@yildiz.edu.tr

Abstract: This paper determines the factors that affect the customer demand by using a statistical test. After finding out the significance of the factors, genetic algorithm (GA) and ANN is hybridized and applied to predict the number of the customer demand for the 10 years later. The result of R^2 value presents the efficiency of both the factors and GA-ANN algorithm.

Keywords: Artificial neural network, Genetic algorithm, K-nearest neighborhood

1 INTRODUCTION

The most of the real life applications of the main problem consists of transferring a set of commodities between source–destination pairs efficiently in a network. In general, given vertices in the network are chosen to aggregate the traffic flow corresponding to several source–destination pairs. These vertices are called hubs and induce a back bone of transfers throughout the whole network. Hubs facilitate the transfers and optimize the costs, for this reason their localization in the network plays a central role.

The important steps in forming the distribution network are the placement of the location, such as warehouses, depots, and distribution centers and routing such a way that considers some given depot or vehicle capacity constraints, in order to satisfy the customer demands and to minimize routing costs, vehicle fixed costs, depot fixed and operating costs (Drexl 2013). The determination of depots and customers is subject to the demand. In general, most of the papers choose the demand as deterministic, however, it does not offer quite realistic results. In recent years, the performing of stochastic and fuzzy demand assumptions is taken a great attention in order to reach accurate and realistic results. Thus, we aim to predict the demand by using GA-ANN, which consists of a stochastic process, for the 10 years later, because, time is a key factor in finding out the best possible depots and routes, because, the condition of any constraint such as the customer demand and the transportation cost can change in time.

The main contributions of the paper are as follows: The performing of GA-ANN in LRP provides more reliable information about the future demands of the customers. Thus, the consideration of the future demands satisfies the customers and the management of the company.

The rest of the paper is organized as follows: Second section presents the information about the case study. Also, inputs and output of the model, which are performed in GA-ANN, are presented in this section. The third section presents the GA-ANN methodology that is applied to predict the demands for the 10 years later and it presents the best prediction parameters and tuning of GA-ANN with regard to R^2. The fourth section presents the results of GA-ANN. The last section is the conclusion that presents the overall findings.

2 CASE STUDY

The company is located in Istanbul, where the biggest urban settlement area in Turkey is. The company has two active depots (D1 and D2) and other possible depots are presented in Figure 1. At first, the effectiveness of the current locations and routes is investigated and the best possible depot(s) are searched to meet the growing demand for the 10 years later.
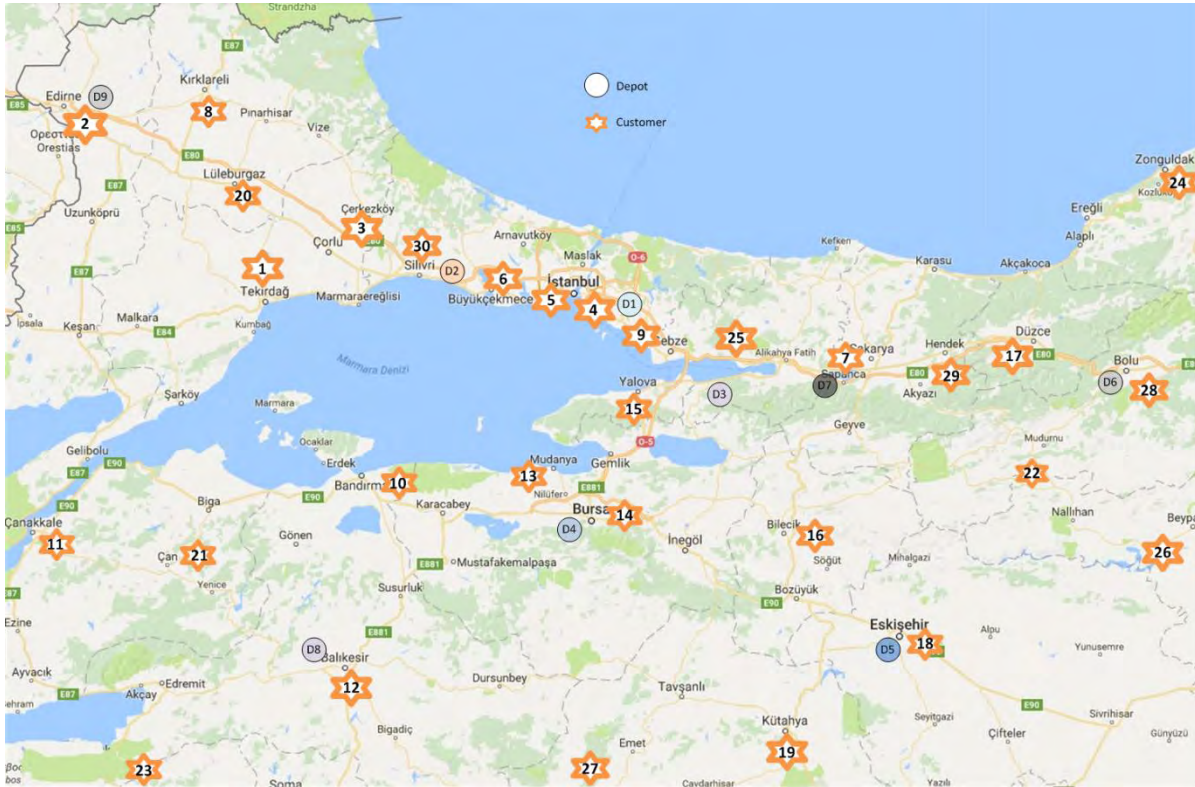
Figure 1. The location of depots and customers

The most of the data (unemployment rate (UR), customer confidence index (CCI), inflation, industrial production index (IPI), economic confidence index (ECI) and automobile number (AN)) are gathered from the Turkish Statistical Institute. Only the demand of the customers is obtained from the company. UR, CCI, inflation, IPI, ECI and AN are inputs and the demand of the customers is output for GA-ANN. The obtained data cover the months between 2012 and 2016 and thus offers 60 instances for training and testing. 20% of the instance is allocated for the testing and the rest is allocated for the training stage. Before the training stage of GA-ANN, each input is clustered with regard to the k-nearest neighbors (k-NN) to explore the main effects of the each input in the model.

K-NN is one of the oldest and simplest non-parametric classification and regression methods. In the k-NN algorithm, a class is allocated according to the most common class amongst its k-NN. The output is a class membership. The input consists of the k closest training examples. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k-NN (Chen et al. 2011). UR, CCI, inflation, IPI, ECI and AN are classified with regard to demand membership

3 METHODOLOGY

GA is initialized with an aggressive set of solution pools and then the solution pools develop through the stage of natural selection, where poor solutions run out and solutions, which have the highest quality, survive to reproduce. This stage is iterated until the optimal condition is satisfied. A hybrid GA-based neural network is basically a back propagation network, with the only exception being that the weight matrix was acquired from performing the genetic operations under optimal convergence conditions (Kadiyala et al. 2013; Pekel and Soner Kara 2017).

Pseudo code of GA-ANN is illustrated in Algorithm 1. Initial weights are set off randomly in the first iteration and the output of each hidden neuron and the error are computed. The updated weights are calculated with regard to GA by applying parent selection, reproduction and mutation. Then, the next iteration is carried out with regard to the updated weights.

Algorithm 1. GA-ANN

1 **Initialize** initial weights ($w(t = 0)$)
2 t←0;
3 **Compute** the output for every neuron
4 **Compute** the error at the output
5 **While** not_terminated() **do**
6 $w_p(t)$←Select_parents();
7 $w_r(t)$←Reproduction();
8 Mutate($w_p(t)$);
9 Evaluate($w_r(t)$);
10 $w_r(t)$←build_next_generation($w_r(t), w(t)$)
11 **End while**
12 t←t+1;

---

4 THE RESULTS OF GA-ANN

The proposed GA-ANN model is run with the combination of four different parameters, which are "population size", "crossover", "mutation" and "hidden neuron". Two values of population size are tested; 10 and 20. Ten different values for crossover (0.1, 0.2,…, 1.0), ten different values for mutation (0.1, 0.2,…, 1.0) and nineteen values for hidden neuron (2, 3,…, 20) are tested.

Various options such as population, fitness scaling, selection, reproduction, mutation, crossover and migration need to be specified as shown in Table 1. The parameters are tuned with respect to trial and error. Optimal condition of population size, crossover fraction, migration fraction and hidden neuron number are searched from 10 to 20, from 0.1 to 0.9, from 0.1 to 0.9 and from 2 to 19, respectively. Especially, when the number of population size and the hidden neuron is increased, computing time also raises.

Table 1: The optimal parameters of GA-ANN

	Options	Optimal conditions
Population	Creation Function	
	Population size	20
Fitness scaling		-
Selection		Stochastic uniform
Reproduction	Elite count	2
	Crossover fraction	0.80
Mutation		Gaussian mutation
Crossover		Heuristic crossover
Migration	Direction	Both
	Fraction	0.50
Algorithm settings	Initial penalty	10
	Penalty factor	100
Stopping criteria	Generations	1000
	Fitness limit	1.00e-8
	Stall generations	50
Hidden neuron number	Single hidden layer	16
CPU time (second)		568

The proposed GA-ANN have been run on a computer that has a 32-bit Windows 7 operating system, 2.4-GHz processor, and 16-GB memory. GA-ANN has been implemented in Matlab 7.12.

Table 2 shows the best 10 values of R^2 with regard to six parameters (there is no exclusion).

Table 2: The best 10 values of R^2 with regard to six parameters

Population Size	Crossover Rate	Mutation Rate	Hidden Neuron	R^2 for Training (Mean value)	R^2 for Testing (Mean value)	Standard Deviation for Training	Standard Deviation for Testing
20	0.8	0.5	16	0.9526	0.8011	0.0087	0.0117
20	0.5	0.7	15	0.9509	0.7927	0.0071	0.0063
20	0.8	0.1	19	0.9508	0.7209	0.0103	0.0094
20	0.7	0.4	16	0.9503	0.7978	0.0060	0.0078
20	0.6	1	12	0.9496	0.7585	0.0116	0.0148
20	0.7	0.1	19	0.9494	0.7589	0.0234	0.0219
20	0.5	0.1	16	0.9492	0.7826	0.0249	0.0170
20	0.7	0.7	15	0.9490	0.7083	0.0161	0.0247
20	0.5	0.1	17	0.9473	0.7480	0.0122	0.0129
20	0.7	0.2	13	0.9469	0.7730	0.0140	0.0144

CCI and ECI are excluded from GA-ANN model to see any improvement (based on R^2) is happening or not. Table 3 shows the best 10 values of R^2 with regard to the four parameters after two parameters (CCI and ECI) are excluded from the model. However, the exclusion of the two parameters does not offer a better R^2 values than all parameters included. R^2 values are found as the result of the combination of population size, crossover rate, mutation rate, and hidden neuron number.

Table 3: The best 10 values of R^2 with regard to four parameters

Population Size	Crossover Rate	Mutation Rate	Hidden Neuron	R^2 for Training (Mean value)	R^2 for Testing (Mean value)	Standard Deviation for Training	Standard Deviation for Testing
20	0.7	1	20	0.7417	0.6607	0.0126	0.0132
20	0.4	0.4	17	0.7411	0.7056	0.0098	0.0081
20	0.8	0.7	17	0.7400	0.6805	0.0040	0.0272
20	0.8	1	16	0.7357	0.6593	0.0086	0.0301
20	0.6	0.8	12	0.7320	0.6499	0.0092	0.0252
20	0.8	0.5	7	0.7319	0.7191	0.0198	0.0296
20	0.5	0.8	7	0.7245	0.6564	0.0120	0.0092
20	0.7	0.1	14	0.7219	0.6256	0.0144	0.0134
20	0.7	0.6	7	0.7213	0.6949	0.0138	0.0186
10	0.5	0.1	14	0.7200	0.7039	0.0129	0.0152

5 CONCLUSION

In this paper, At first, each input (factor) is clustered with regard to k-NN in order to generate a suitable parameter level. The factors that affect the customer demand are tested by using a statistical method. Later, the main effects of each factor are investigated to provide more accurate data for the prediction. Statistical test shows that CCI and ECI are not significant in the prediction model. However, the exclusion of the two parameters does not offer a better R^2 values than all parameters included. Therefore, GA and ANN is hybridized and applied to predict the customer demand by considering all factors for the 10 years later.

References

[1] Chen, H. L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S. J., & Liu, D. Y. 2011. A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. Knowledge-Based Systems, 24(8), 1348-1359.

[2] Drexl, M. 2013. Applications of the vehicle routing problem with trailers and transshipments. European Journal of Operational Research, 227(2), 275-283.

[3] Kadiyala, A., Kaur, D., & Kumar, A. 2013. Development of hybrid genetic-algorithm-based neural networks using regression trees for modeling air quality inside a public transportation bus. Journal of the Air & Waste Management Association 63(2), 205-218.

[4] Pekel, E., & Soner Kara, S. 2017. Passenger Flow Prediction Based on Newly Adopted Algorithms. Applied Artificial Intelligence, 31(1), 64-79.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 4:
MCDM – Software and Applications

MULTI-CRITERIA DEX MODELS: AN OVERVIEW AND ANALYSIS

Marko Bohanec
Jožef Stefan Institute, Department of Knowledge Technologies
Jamova cesta 39, SI-1000 Ljubljana, Slovenia
E-mail: marko.bohanec@ijs.si

Abstract: DEX (Decision EXpert) is a hierarchical, qualitative, rule-based, multi-criteria decision modelling method. Since its conception in 1979, it has been used in decision-support applications in various areas, including economy, finance, agriculture and tourism. In this study, we analysed 582 DEX models developed in the period 1979–2015, assessing statistical properties of the main components of DEX models: attribute hierarchies, attribute scales, and decision rules. We also studied the completeness, monotonicity, linearity and symmetricity of the underlying aggregation functions. The results are useful particularly for understanding the boundaries of the decision problems addressed by DEX, and for improving the methodology and the design of the supporting software.

Keywords: Multi-criteria decision model, qualitative model, hierarchical model, method DEX, applications.

1 INTRODUCTION

Multi Criteria Decision Modelling (MCDM) [9] is concerned with structuring and solving decision problems that involve multiple and possibly conflicting criteria. With the aim to support the decision maker, MCDM provides methods and means to obtain preferential information from the decision maker, to represent it in a form of a decision model, and to use the model to perform the intended decision-making tasks: choosing, ranking and/or sorting decision alternatives, and analysing and justifying the results.

In this paper, we focus on one MCDM method, called DEX, and its applications. DEX (Decision EXpert) was conceived in early 1980's [8] under the name DECMAK [1], combining the approach of hierarchical MCDM with rule-based expert systems and fuzzy sets. The name DEX was coined around 1990 [2] together with the method's implementation in a form of an expert system shell for decision support. DEX has three key characteristics:

1. It is *hierarchical*: a DEX model consists of hierarchically structured *attributes* (in MCDM, also called *criteria* or *performance variables*).
2. It is *qualitative*: all attributes are symbolic, taking values that are words rather than numbers, such as "bad", "medium", "excellent", "low", or "high".
3. It is *rule-based*: the hierarchical aggregation of values is defined with *decision rules*, acquired and represented in the form of *decision tables*.

Currently, the DEX method is implemented in a freely available software called DEXi [7]. DEXi supports an interactive construction of the decision model, and evaluation and analysis of alternatives [4]. The decision maker is aided in creating the model structure and defining decision rules. Following the principles of expert systems, DEXi can evaluate alternatives even in the case of incomplete input and preference data. DEXi is a third generation of DEX software; previous generations were called DECMAK [1] and DEX [2].

DEX has been used to support complex decision processes in various problem domains, including health care, project management, quality and risk assessment, environmental management, data mining, and many more. Literally thousands of DEX models have been developed worldwide and used to solve real-life decision problems [3].

The idea explored in this study is that we can learn from DEX models developed in the past. We can analyse their characteristics, such as the size and structure of the attribute hierarchies, types and scales of individual attributes, number and quality of decision rules. On

this basis, we may better understand the requirements of the decision-modelling process, and possibly develop better algorithms and tools in the future.

With these goals in mind, we have compiled a research database of DEX models. It contains 582 models developed in 140 decision-making projects conducted in the period 1979–2015. The collection is restricted only to DEX models that were available to the author of this study, who is also a decision analyst and DEX software developer; models developed elsewhere in the world were not included. Nevertheless we believe that the database is highly representative with respect to the addressed decision problems, decision makers involved, covered time period and observed model characteristics. All the included models are "real" in the sense that they were developed by real people with specific decision problems in mind.

In what follows we first present an overview of decision projects included in the database. Then, we describe the key components of the DEX models and their statistical properties: hierarchical structure of attributes, individual attributes and their scales, and decision rules.

2 PROJECTS

The database contains models that were developed in 140 decision-making projects in the period 1979–2015. Here, a "project" denotes a set of related DEX models. In a normal MCDM setting, only one model is expected to be developed per project, aimed at the evaluation of the corresponding decision alternatives. In reality, however, multiple models are often developed for various reasons, such as addressing different decision-making subtasks, different aspects of the problem, different classes of alternatives, or different decision makers.

In the database, the 140 projects actually contain 582 models in total (4.16 models per project). They address very different decision problems and consider the assessment of:
- *Computer technology*: software, hardware, IT tools, programming languages, data base management systems, decision support systems;
- *Projects*: investments, research and R&D projects, tenders;
- *Organisations*: public enterprises, banks, business partners;
- *Schools*: quality of schools, programmes and teachers, school admission, choosing sports for schoolchildren;
- *Management*: production, portfolio management, trade, personnel (employees, jobs, teams), privatization, motorway;
- *Production*: location of facilities, technology, logistics, suppliers, office operations, construction, electric energy production, sustainability;
- *Ecology and Environment*: dumpsite/deposit assessment and remediation, emissions, ecological impacts, soil quality, ecosystem, sustainable development, protected areas;
- *Medicine and Health Care*: risk assessment (breast cancer, diabetes, ski injuries), nursing, technical analysis, knowledge management, healthcare network;
- *Agriculture and Food Production*: economic and ecological effects of using genetically modified crops, crop protection, hop hybrids, garden quality;
- *Tourism*: nature trail, tourism farm facilities, mountain huts;
- *Services*: loans, housing loans, public portals, public services, leasing;
- *Other*: cars, hotels, electric motors, radars, game devices, awards, options, drug addiction, roof covering, data mining.

Among these, 52 projects (38%) are documented in publicly available conference and journal publications. Further 20 (14%) of projects are documented in internal reports. The remaining 67 (48%) of projects are not documented beyond the models themselves. For space limitations, we cannot include the published references here. Please see [3] for a review of some projects and publications. Some recent DEX projects were presented in [6] and [7].

3 MODELS

A DEX model consists of hierarchically structured attributes. Each attribute represents an aspect of considered decision alternatives that is interesting for their assessment in the given decision context. The hierarchy consists of *basic attributes* (terminal nodes), which represent model inputs, and *aggregate attributes* (internal nodes), which represent model outputs and provide assessments of the alternatives. The aggregate attributes depend on their (basic or aggregate) descendants in the hierarchy. The ultimate output of the model (overall assessment of the alternatives) is represented by one or more *root* attributes.

To illustrate these concepts, Figure 1 presents a small, but typical model for the evaluation of cars. This demo model is distributed together with the DEXi software and is also included in the database analysed in this paper. The model assesses cars using the seven basic attributes: BUY.PRICE (buying price), MAINT.PRICE (maintenance price), #PERS (number of persons), #DOORS (number of doors), LUGGAGE (luggage boot size), and SAFETY (car safety). The four aggregate attributes include COMFORT (depends on #PERS, #DOORS and LUGGAGE), TECH.CHAR (technical characteristics, based on COMFORT and SAFETY), PRICE (assessed from the buying and maintenance price), and the root attribute CAR (depends on PRICE and TECH.CHAR). Notice that the *depth* (number of levels excluding the root) of the model is three. The *width* (number of immediate attribute descendants) of CAR is two, and of COMFORT is three.

Attribute	Scale
CAR	**unacc**; acc; good; ***exc***
├─**PRICE**	**high**; medium; ***low***
│ ├─BUY.PRICE	**high**; medium; ***low***
│ └─MAINT.PRICE	**high**; medium; ***low***
└─**TECH.CHAR.**	**bad**; acc; good; ***exc***
├─**COMFORT**	**small**; medium; ***high***
│ ├─#PERS	**to_2**; 3-4; ***more***
│ ├─#DOORS	**2**; 3; 4; ***more***
│ └─LUGGAGE	**small**; medium; ***big***
└─SAFETY	**small**; medium; ***high***

Figure 1: Structure and scales of the car-assessment DEX model

The 582 models in the database are characterised as follows:

- *By category*: 175 (30%) models were developed in various research projects. 129 (22%) are "commercial", developed in decision-analysis projects for a paying customer. 213 (37%) models were developed as part of various educational activities, such as student assignments. The remaining 65 (11%) are various demonstrational models, such as the one in Figure 1.
- *By language*: 178 (30%) English, 388 (67%) Slovene, 11 (2%) Croatian, and 5 (1%) Spanish.
- *By software*: 53 (9%) of the models were developed by the first generation software DECMAK (mainly 1979–1991), 208 (36%) by the second generation DEX (1989–2001), and 321 (55%) by the third generation DEXi (since 2000).
- *By structure*: the majority of models (554, 95%) have a tree structure, while the remaining 28 (5%) employ a full hierarchy (i.e., a directed acyclic graph, where some attributes affect more than a single higher-level attribute).

Figure 2 (left) shows the histogram of model *sizes* (total number of attributes). Typical models have between 10 and 30 attributes, but there are also very large models containing more than 100 attributes. The average size is 27.8 attributes. The largest model in the database, which is aimed at the evaluation of cropping systems in agronomy, has 383 attributes. The average number of basic and aggregate attributes is 15.8 and 10.3,

respectively. The deepest model contains 10 levels, and the average depth is 3.5. The average width of a single node is 2.6, indicating that DEX models are somewhat "thin and deep". This is a consequence of a methodological recommendation to keep the number of immediate descendants low in order to avoid too large decision tables [4].
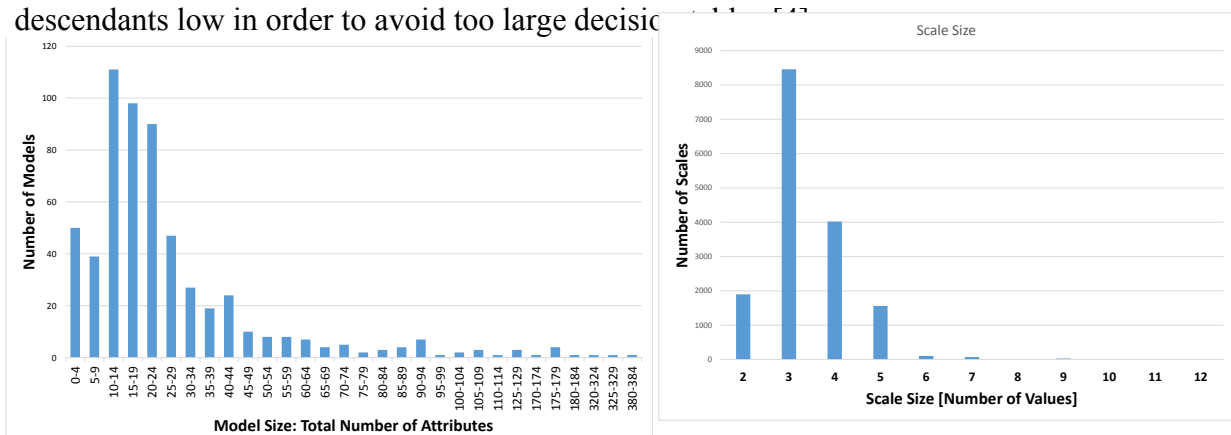
Figure 2: Histograms of model sizes (left) and scale sizes (right)

An additional analysis of models over time (not detailed here) indicated that the average size of models remained almost constant since 1979, however the size of the largest models, which are very few, grew from about 100 in 1980 to almost 400 in the recent years.

4 ATTRIBUTE SCALES

Attributes in DEX models are qualitative: their value scales are discrete and composed of words. Figure 1 shows the scales used in the car-assessment model. It can be seen that they contain only a small number of possible values, three or four in this case. The colours indicate that all the scales are also preferentially ordered from "bad" (printed in red) to "good" (green) values.

This is typical for most DEX models, even though the method permits using more values and having unordered scales. The database contains 16169 scales in total. Figure 2 (right) shows that the vast majority of scales contain 2 to 5 values. The average scale size is 3.4.

The definition of scale ordering has changed over different generations of software and is thus difficult to assess. Roughly, in all the models, 13138 (81%) scales are increasing, 838 (5%) decreasing, and 2192 (14%) unordered.

5 DECISION RULES

In DEX, the aggregation of basic attributes towards the roots of the hierarchy is governed by *decision rules*. These are typically formulated by the decision maker (with the support of software tools) and represented in terms of *decision tables*. Each aggregate attribute in the model has an associated decision table.

For example, Figure 3 shows a screenshot from DEXi in which a decision table for the attribute CAR is being developed. Notice that the table contains all the 12 possible combinations of the values of the descendant attributes PRICE and TECH.CHAR, and that the values of CAR for each combination are given in the rightmost column. Each column represents an *elementary decision rule*. The bold typeface indicates the values entered by the decision maker, while the normal typeface (only at rule number 10) indicates the values suggested by DEXi's decision-support algorithms [4].

Figure 3: Decision table for the assessment of CAR depending on PRICE and TECH.CHAR

In the following, we define some interesting properties of DEX decision tables and present the actual numbers assessed on the 6362 decision tables contained in the studied database.

- *Number of arguments*: defined as the number of conditional attributes in the decision table (there are 2 in Figure 3: PRICE and TECH.CHAR). The database contains decision tables that have 1 to 8 arguments, the average is 2.5.
- *Number of classes*: denoted $|Y|$, the number of values of the output attribute Y (in Figure 3, the output attribute CAR has 4 values). The observed range of classes is 2 to 11, the average is 3.7.
- *Size*: the total number of decision rules ($3 \times 4 = 12$ in Figure 3). In the database, the size varies greatly between 2 and 15625, but the average and median are only 39.3 and 16, respectively. This indicates that reasonably small tables are preferred by the decision makers and that the "combinatorial explosion" [4], possibly caused by too many attributes and attribute values, is generally kept under control well.
- *Definition*: the proportion of decision rules defined by the decision maker ($^{11}/_{12} =$ 91,67% in Figure 3). The database contains 5034 (79%) completely defined decision tables and 1328 (21%) incompletely defined ones.
- *Determination*: Similar to the above, but also taking into account the values suggested by the software. These suggestions raise the proportion of completely determined decision tables to 5889 (93%); only 473 (7%) decision tables still contain output values that are not fully determined (i.e., specified in terms of unknown or interval values).
- *Monotonicity*: Whether or not the decision table, interpreted as an aggregation function, is monotonically increasing with increased values of its arguments, taking into account the preferential ordering of attribute scales. In other words, whether or not do decision rules conform to the *principle of dominance* [4, 9]. As much as 5993 (93%) decision tables in the database are monotone, and only the remaining 429 (7%) are not. This indicates that the principle of dominance is indeed a powerful mechanism for ensuring the preferential consistency of decision tables, and that the latter has been managed really well.
- *Symmetricity*: Whether or not the decision table is fully symmetric with respect to its all arguments. In the database, 1663 (26%) decision tables are fully symmetric, and 4699 (74%) are not. Partial symmetricity is more abundant, but left out from this presentation.
- *Linearity*: Whether or not it is possible to fully approximate the decision table with a linear function. 1638 (26%) of decision tables are linear in this sense, and 4724 (74%) are not. The proportion of linear decision tables is thus similar to the proportion of symmetric ones.

6 CONCLUSION

The purpose of this study was to make an overview of MCDM DEX models developed since 1979, and to assess basic statistical properties of their primary components: attributes, scales and decision rules. Why is this important? First, it allows us to better understand the DEX models: their dimensions, properties, historical development, trends, possible errors, etc. Second, it facilitates the quality assessment of models and their components, such as the consistency of decision rules. Third, quality assessment may provide a solid basis for quality assurance, for instance, in developing better and more effective decision-support software. Last but not least, the developed database of DEX models provides a reach real-data resource for further research.

This study revealed the average and extreme dimensions of DEX models. An average model consists of roughly 28 attributes (16 of which are basic), 3.5 levels and 2.5 descendants per node. The largest models may span up to 400 attributes and 10 levels. Over time, extreme models were becoming larger, while the average models remained the same. An average scale contains 3.4 values and is ordered by increasing preference. An average decision table has 2.5 arguments, 3.7 classes and 40 decision rules (with the median of 16). Decision tables are mostly monotone (93%), non-symmetric (74%) and non-linear (74%). The overall completeness of decision tables is high (93%). It was also found that the quality of model components depends on implemented features of the supporting software; for instance, the improved handling of scale ordering in subsequent generations of software improved the overall completeness of decision rules. With this in mind, further development of the DEX method and supporting software should address an explicit handling of the symmetricity and linearity of decision tables as means to improve the knowledge acquisition process and the quality of acquired rules.

References

[1] Bohanec, M., Bratko, I., Rajkovič, V. 1983. An expert system for decision making. Processes and Tools for Decision Making (ed. H.G. Sol). North-Holland, 235–248.

[2] Bohanec, M., Rajkovič, V. 1990. DEX: An expert system shell for decision support, Sistemica 1(1): 145–157.

[3] Bohanec, M., Žnidaršič, M., Rajkovič, V., Bratko, I., Zupan, B. 2013. DEX methodology: Three decades of qualitative multi-attribute modeling. Informatica, 37(1), 49–54.

[4] Bohanec, M. 2015. DEXi: Program for Multi-Attribute Decision Making, User's Manual, Version 5.00. IJS Report DP-11897, Ljubljana: Jožef Stefan Institute.

[5] Bohanec, M., Trdin, N., Kontić, B. 2016. A qualitative multi-criteria modelling approach to the assessment of electric energy production technologies in Slovenia. Central European Journal of Operations Research, 1–15.

[6] Bohanec, M., Mileva Boshkoska, B., Prins, T. W., Kok, E. J. 2017. SIGMO: A decision support System for Identification of genetically modified food or feed products. Food Control, 71: 168–177.

[7] DEXi: A Program for Multi-Attribute Decision Making. http://kt.ijs.si/MarkoBohanec/dexi.html [Accessed 2017/06/15].

[8] Efstathiou, J., Rajkovič, V. 1979. Multiattribute decisionmaking using a fuzzy heuristic approach. IEEE Trans. on Systems, Man, and Cybernetics, SMC-9: 326–333.

[9] Greco, S., Ehrgott, M., Figueira, J. 2016. Multiple Criteria Decision Analysis: State of the Art Surveys. International Series in Operations Research & Management Science, Vol. 233. New York: Springer.

DECISION SUPPORT MODELLING FOR EFFICIENT IMPLEMENTATION OF ICT IN SCHOOLS

Borut Čampelj[1], Igor Karnet[1], Andrej Brodnik[2], Eva Jereb[1], Uroš Rajkovič[1]
[1] University of Maribor, Faculty of Organisational Sciences, Kidričeva 55, SI-4000 Kranj, Slovenia
[2] University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, SI-1000 Ljubljana, Slovenia
E-mail: borut.campelj@gov.si, igor.karnet@gmail.com, andrej.brodnik@fri.uni-lj.si, eva.jereb@fov.uni-mb.si, uros.rajkovic@fov.uni-mb.si

Abstract: The implementation of information communication technology (ICT) in schools has not always an impact to all employees' plans and activities. This paper presents a new hierarchical evaluation model (tree of indicator) of the level of implementation of ICT in school. The model offers simplicity of the evaluation process and provides transparency, because the user assigns only simple indicators (106) and then the model automatically computes the values of complex indicators (65) based on qualitative relations among indicators at lower level of the tree. The model enables non-professional ICT-users to organise and harmonise existing information and knowledge, and encourage them to identify priority goals and realizable changes.

Keywords: MCDM, DEX, qualitative model, decision rules, ICT in education, self-evaluation

1 INTRODUCTION

European schools are expected to provide conditions for further development and qualitative upgrade of activities through an open and innovative learning environment supported by information communication technology (ICT) with a co-operation of different stakeholders [3]. The review of the development in various countries showed the importance of different aspects of ICT: the analysis of the school vision, school leadership and organization of activities [13]. The European survey has shown the progress of achieved level of the school ICT-strategies and quality of support provided for schools [4 page 141, 14]. However, the school's ICT strategies do not have a big enough impact on all employees at a school and therefore the strategies are not fully integrated into the plans and activities at school or individual level [6, page 181].

The research question and contribution of this paper is a development of a new hierarchical multiple criteria decision model of evaluation of the implementation ICT in school using intelligent methodologies, based on the theoretical and practical background and our own experience. The idea is to support the sustainable process, in which schools find time, space and use efficient tools for comprehensive upgrade of implementation ICT in all school activities. The model should be a qualitative and motivational tool for transparent assessment of the current situation at school, which could lead to a higher level planning and fully introduction of ICT. Namely, this assessment procedure could increases the possibility that teachers, headmaster and others better understand the comprehensive implementation of ICT and unify the opinions about the current situation at school. Collected and organised data and ideas lead the school to identify realizable goals and concrete changes.

2 LITERATURE REVIEW AND METHODOLOGY

Information tools have been developed to help schools how to include innovations and opportunities that ICT brings and offers, but there are issues about sustainable use. The existing indicators of qualitative and efficient use of ICT in education could be grouped to school (leadership) and teacher level. At school level (leadership level), comprehensive evaluation information models of ICT at school have been developed, for example in the UK

[1], Future Classroom Toolkit from European Schoolnet [5] or OPECA model in Finland [10]. At teacher level, various national frameworks of digital competences of teachers have been developed, for example, in Slovenia "E-competent teacher framework" [8], Finish model OPECA [10]. These models consist mostly of general (complex) indicators at a school level (school vision, flexibility, human resources, infrastructure, budget etc.) or a teacher level (teacher training, learning scenarios, innovative learning activities, evaluation of lessons etc.). Mostly five-point scales have been defined for each complex indicator. These scales are qualitative (descriptive and not only quantitative) but in many cases too complex to allow users to uniquely identify adequate options or ensure the necessary transparency. Moreover, in the next evaluation procedure (after one/half year), the user could not easily identify the detailed progress of implementation of ICT in school because of the complexity.

Our goal was to use methodology to develop a model that requires the qualitative evaluation of simple indicators only and then combines the results of this evaluation to automatically set the values of complex indicators. Namely, it is hard to evaluate the complex indicators because of the complexity of the qualitative scales. On the other hand, simple indicators are easy to evaluate, as with simple indicators it is much easier to follow the progress and the evaluation based on the fact that simple indicators ensure transparency.

HMADM - Hierarchical Multi-Attribute Decision Model [11, 12] can be used for the qualitative analysis of the existing situation and for the interpretation of the results to support the upgrade of the existing situation. The HMADM model is based on the selection of the *indicators* of the interested profession. In order to reduce the complexity of the decision model with respect to the number of indicators and their interrelations, a hierarchical structure (tree) is employed. Indicators at the higher level are computed using utility functions, which take the indicators at lower levels as parameters. Based on the position of the indicator in the tree, we distinguish between the *simple indicators* - leaves (for example X_1, X_2, …) and the *complex indicators* (for example Y_1, Y_2, …). For each complex indicator there is a corresponding aggregated utility function, because the function domain depends on values of sub-indicators at lower levels, for example $Y_2 = F_4 (X_1, X_2, Y_1)$.

A DEX - Decision Expert Methodology [2] is developed on the base of HMADM. The indicators have discrete values and are usually represented by text rather than numbers at leaves (simple indicators). The utility functions at complex indicators are defined by decision rules and not only by mathematical function. The final domains are therefore presented in the tables and are points in a discrete multidimensional space.

3 QUALITATIVE MODEL AS A TREE OF INDICATORS

We used DEX methodology to define and hierarchical organize 171 indicators (65 complex and 106 simple) of implementation of ICT in school ("school informatisation"). Due to the number of indicator we do not look to all indicators in detail in this article.

3.1 Definition of indicators and their organisation as a tree

To develop the tree of indicators, we used a procedure of decomposition of complex indicators into more identifiable and independent sub-indicators. The most complex indicator (root of the tree) "Level of the school informatisation" is firstly divided into three less complex sub-indicators: "School and the environment", "Teachers and e-communities" and "Students and the environment". These three sub-indicators are further divided into less complex ones and the procedure is repeated until the sub-indicators are simple enough to be easily described and assigned. Only the sub-tree "School and the environment" with 24 complex and 36 simple indicators is presented in Figure 1. Although we developed also sub-

tree "Teachers and e-communities" with 20 complex and 34 simple indicators and sub-tree "Students and the environment" with 21 complex and 36 simple indicators.

The tree of indicators is better adapted to human cognitive abilities and comprehensive information processing [9] as the individual indicator is not fragmented into too many direct sub-indicators (not more than 3) and therefore the corresponding utility function is not too complex or unclear. We rather develop a tree that has a depth of 10 levels. This method is one of the major differences from other models. Furthermore, this approach supports schools to evaluate efficiently the situation and easily articulate the progress that needs to be made.

SCHOOL AND ENVIRONMENT **6** - Headmaster, **7** - First teacher, **8** - Second teacher

e-competent school 7,7,8

5,8,8 environment

vision and plan 7,8,9

5,7,7 cooperation

vision 7,8,9

plan 7,7,8

promotion 2,3,3
realisation 5,7,7

3,3,3

7,8,10 **3,3,4 teachers** distribution

periods

7,8,10 employees leadership 3,3,4 3,3,3 content

content 2,3,3

e-competences other services 3,4,4 3,2,2 7,8,10 levels

organisation 3,3,3

7,9,10 students 3,3,4

3,4,4 school

impact 2,4,4

content parents 2,4,4

3,3,4 individual

design 7,8,9 4,4,4 ICT-infrastructure

processes 3,4,4

organisation staff 3,3,3

7,9,9

upgrade 4,4,4
authors 3,2,2

strategic leadership 6,6,7

ICT team 3,3,3 plan 3,3,3

administration 2,2,3

staff train 3,3,3 organisation lessons 4,4,3

7,7,7 employment 3,3,3 7,7,8 other 2,2,4

equip 3,3,3

infrastructure use 3,3,3

7,7,6 standards 3,3,2

resources 6,6,6 plan 2,2,2

information system efficiency 2,2,2

services 5,5,5 accesibility 3,3,3

5,5,5 support plan 3,3,3

5,5,5 implement 2,2,2

7,7,7 plan 3,3,3
evaluation measure 3,3,3
impact 3,3,3

LEGEND:
Grey box: complex indicator (10-points scale)
White box: simple indicator (4-points descriptive scale)
Three numbers X, Y, Z at each box (indicator) are examples of evaluation made by Headmaster (X), First teacher (Y) and Second teacher (Z)

Figure 1: Sub-tree "School and the environment" (indicators and results of evaluation)

3.2 Domain of simple indicators

At each leaf of the tree (simple indicator), domain is set as descriptive 4-range scale, which ensures the qualitative and transparent evaluation. An example of the scale of the leaf "teachers" (in Figure 1 marked as bigger white rectangle with bold text) is in the Table 1.

Scale	Description
1	*There is no specific vision of e-competences of teachers, except in general that teachers need to develop e-skills*
2	*The vision contains mostly basic e-skills of teachers: word processing (preparation of lessons, tests, reports), presentations, communication (e-mail, web, video conferencing, ...) and educational technology (digital projector, interactive whiteboard ...)*
3	*The vision contains the major skills of the e-competent teacher framework ('Slovenian example'), personal competences and lifelong learning skills with the support of ICT*
4	*The vision includes all skills of the e-competent teacher framework, e-skills for personal development (profession, teamwork, personal growth) and the development of lifelong learning skills which could be encouraged by ICT including international recommendations (EU, OECD, UNESCO ...)*

There are three numbers beside each of the indicators in Figure 3. These present three examples of the evaluation in the testing phase. For example, beside the indicator "teachers" in Figure 3 are numbers 3, 3, 4 as one headmaster and one teacher assigned this indicator the value of 3 and another teacher assigned it the value of 4.

3.3 Domain of the complex indicators (utility functions)

Each utility function has as its domain discrete values from 1 to 10, because this is more visible for users in order to find the differences of evaluation results. Firstly, we used weights to define the values, because there are many complex indicators (eg. 24 complex indicators in Figure 1) and secondly, we adjusted some of the values of final domain. The weights are defined based on importance of a particular indicator, the knowledge of experts and testing results. This assures that there are no distortions of results for the complex indicators at higher levels and the utility functions define the qualitative relations between indicators from lower levels of the tree.

For example, let us consider the utility function of the top indicator in the sub-tree "School and Environment" (weights of other utility functions are not presented this paper). Table 2 defines the discrete utility function with the final domain 1 to 10. The utility function is defined by 100 possible options, because it has as the parameters the two direct sub-indicators "E-competent school" and "Environment" (both with final domain 1 to 10). Firstly, the weights of sub-indicators are set: "E-competent school" (60 %) and "Environment" (40 %) and then some values adjusted. For example, we underlined eight options in Table 2 (*No. 8, 9, 10, 42, 72, 91, 92 and 95*).

Table 2: Domain of the complex indicator "School and the environment" (utility function F)

	E-competent school (X_1)	Environment (X_2)	School and environment $F(X_1, X_2) =$ $= 0,6 * X_1 + 0,4 * X_2$		E-competent school (X_1)	Environment (X_2)	School and environment $F(X_1, X_2) =$ $= 0,6 * X_1 + 0,4 * X_2$
1	1	1	1	...	....	...	...
2	1	2	1	68	7	8	7
...	....	...	...	...	...	...	...
8	1	8	3	72	8	2	5
9	1	9	3	...	...	...	...
10	1	10	3	78	8	8	8
11	2	1	2	...	...	...	...
12	2	2	2	91	10	1	5
...	....	...	...	92	10	2	6
42	5	2	3	95	10	5	7
...	....	...	...	...	...	...	...
65	7	5	6	100	10	10	10

In the Table 2, three options (*No. 65, 68 and 78*) are indicated in bold as these are three examples of the evaluation in testing phase (Figure 1).

4 PILOTING AT SCHOOLS

The model was piloting at five basic schools in Slovenia also with 54 teachers, ICT coordinators and headmasters.

4.1 Piloting results

At each school, we performed a self-evaluation workshop (4 hours) in 3 steps with around teachers and headmaster at each school:
   a) Step 1: to determinate the objective situation of the school informatisation by using the tree of indicator such as a self-evaluation questionnaire (in pairs or triples), which encourages the systematic approach of collecting, analysing and assessing a situation and not missing anything important.
   b) Step 2: to reflect on a dialogue between participants and to harmonize the results and ideas obtained by using the self-evaluation questionnaire. They explained the results of the self-evaluation and gave feedback about the reasons for differences and finally, understood the others' opinions. Then a shift of the debate has been encouraged towards harmonization and common understanding of the current level of school informatisation.
   c) Step 3: to articulate common specific priorities and possible changes to future plans. All the strengths and weaknesses of the school informatisation, which were identified during the reflection, were used in the synthesis and determination of priority areas.

4.2 Responds from users about the model

The model supports a detailed objective determination of the situation of school informatisation. All participants (100%) answered that the tree of indicators (self-evaluation questionnaire) was a useful tool to assess the current situation of the school informatisation. Most of them (96 %) expanded the range of possibilities how to use ICT. All participants (100%) were sure that all areas of the school informatisation are effective and meaningfully designed and organized. The 4-level descriptive scales at the leaves of the tree were reasonable and simple understandable for 78% participants.

   The model encourages a harmonization of different views and supports common understandings and upgrade of the current situation of the school informatisation. All participants (100%) agreed that they harmonized different views and better understood the "comprehensive school informatisation". The self-evaluation questionnaire (tree of indicators) encouraged even half of the participants (50%) to change the priorities of the school informatisation. All participants answered that the final domain from 1 to 10 of the utility function at complex indicators provides sufficient dispersion of the results, which helps them to more easily identify the relevant specific differences and challenges for discussion between employees to identify priority goals and realizable changes.

CONCLUSION

The developed model enables users to organise efficiently existing information and knowledge, which supports schools in articulation of the progress of implementation of ICT in schools, which needs to be made. The model is defined throughout a tree of simple and

complex indicators where the simple reside in leaves and the complex ones are in the internal modes of the tree. The testing results indicate that it offers simplicity of the evaluation process and provides transparency. The model requires from the user to assign only the values at the leaves, which have a simple qualitative scales. The values of complex indicators are then automatically computed. Since only simple indicators are assigned by the user, the whole approach becomes very practical. It is much easier for the user to think about the simple indicators than the complex ones. The hierarchical organisation of the indicators and their interconnection based on qualitative methodology DEX implement the idea of how a multiple attribute model could become powerful enough to help decision makers to adopt the best solutions regarding future issues [7].

References

[1] Becta (2008). Schools - Leadership and Management, England, http://webarchive. nationalarchives.gov.uk/20110130111510/http:/schools.becta.org.uk/index.php?section=lv&catc ode=ss_lv_mis_im03&rid=14734 (17/6/2017).

[2] Bohanec, M., Rajkovič, V., Bratko, I., Zupan, B., Žnidaršič, M. (2013): DEX methodology: Three decades of qualitative multi-attribute modelling. *Informatica* 37, 49-54.

[3] European Commission (2013). Opening up Education: Innovative teaching and learning for all through new Technologies and Open Educational Resources; http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013DC0654&from=EN (17/6/2017)..

[4] European Commission (2013). Survey of Schools: ICT in Education, European Commission, Brussels; https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/KK-31-13-401-EN-N.pdf (17/6/2017).

[5] EUN (2014). Future Classroom Toolkit, iTec project, European Schoolnet, http://fcl.eun.org/sl/toolkit (17/6/2017).

[6] Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt E. (2014): Preparing for Life in a Digital Age, The IEA International Computer and Information Literacy Study - ICILS, International Report, International Association for the Evaluation of Educational Achievement (IEA).

[7] Hashemkhani Zolfani, S., Maknoon, R., Kazimieras Zavadskas, E. (2016). Multiple attribute decision making (MADM) based scenarios, *International Journal of Strategic Property Management,* 20(1), 101 – 111.

[8] Kreuh, N. (2012). (Ed) The way towards e-competency, E-Centre of the E-Education Project, No. 7/2012 (Bulletin), Ljubljana; http://www.sio.si/fileadmin/dokumenti/bilteni/E-solstvo_BILTEN_ANG_2012_screen.pdf (17/6/2017).

[9] Lindsay, Peter H. (1977). Human Information Processing: An Introduction to Psychology, Harcourt College Pub; 2nd edition.

[10] Sairanen. K, Vuorinen, M., Viteli, J. (2013). Collecting and Using Data to Develop Digital Learning Culture at School, University of Tampere, Finland; http://blogs.helsinki.fi/tepe-2013/files/2013/12/Sairanen_Vuorinen_Viteli_Collecting-and-Using-data-to-Develop-Digital-Learning-Culture-at-School.pdf (17/6/2017).

[11] Triantaphyllou, E. Multi-criteria Desicion Making Methods (2000). A Comparative Study, Boston: Kluwer Academic Press, 5-72.

[12] Turban, E., Aronson, J., Liang TP. (2004). Decision Support Systems and Intelligent Systems, 7th edn. New Jersey: Prentice Hall, 558 – 601.

[13] Vanderlinde, R., Aesaert, K., van Braak, J. (2015). Measuring ICT use and contributing conditions in primary schools: ICT use and contributing conditions, British Journal of Educational Technology (46/3).

[14] Wastiau, P., Blamire, R., Kearney, C., Quittre, V., Van de Gaer, E., Monseur, C. (2013). The Use of ICT in Education: a survey of schools in Europe, *European Journal of Education*, 48(1), 11- 27.

Experiences from Developing an Algorithm to Support Risk-Based Decisions for Offshore Installations

Gencer Erdogan[a], Atle Refsdal[a], Bjørn Nygård[b], Bernt Kvam Randeberg[c],
Ole Petter Rosland[b]

[a]SINTEF Digital, P.O. Box 124 Blindern, 0314 Oslo, Norway,
`{gencer.erdogan,atle.refsdal}@sintef.no`
[b]Statoil ASA, Forusbeen 50, 4035 Stavanger, Norway, `{bjnyg,olpr}@statoil.com`
[c]Oilfield Technology Group, Vassbotnen 1, 4313 Sandnes, Norway, `bernt.kvam.randeberg@otg.no`

Abstract: We present our experiences from developing a decision model to support risk-based decisions on offshore installations. The model was developed using the DEXi tool for multi-criteria decision modeling. We report on the method we employed, the efforts spent, and the evaluation of the resulting model, including feedback from domain experts representing the target group. In our view the results are promising, and we believe that the approach can be fruitful in a wider range of risk-based decision support scenarios.

Keywords: safety risk, operational risk, offshore, multi-criteria decision making

1 INTRODUCTION

During the autumn of 2016, we developed a computerized model to support decisions based on operational safety risk offshore. The model automatically provides a decision advice based on 28 input parameters, and was developed using DEXi [4], which is a tool for multi-criteria decision modeling. The choice of DEXi was made based on early experience two of the authors had from using DEXi in the domain of cyber-risk [7].

The contribution of this paper is a report on our experiences, the efforts spent on the model development, and an initial evaluation of the results. The aim is to help others who face related challenges to consider whether a similar approach may be suitable for them. We start by explaining the challenge and our success criteria.

During major maintenance projects on offshore installations, flotels are often used to accommodate the personnel. A flotel ("floating hotel") is a vessel providing sleeping quarters and other facilities. A gangway connects the flotel to the installation. The flotel needs to keep its position in a very limited area close to the installation. This can be done by means of Dynamic Positioning (DP), thruster assisted mooring or mooring systems. DP implies employing a computer-controlled system that allows the flotel to automatically keep its position by using its own thrusters. However, keeping the position is highly challenging due to the weather, waves, and other conditions offshore.

If conditions are unfavorable, the responsible offshore operatives need to decide whether to lift (disconnect) the gangway from the installation. If this is not done, there is a risk that an uncontrolled autolift (disconnection) occurs, causing harm to personnel and equipment. The decision is difficult because many different factors affect the risk. Moreover, lifting the gangway has high economic cost, as workers will be prevented from performing their tasks on the installation. Currently, the offshore operatives make use of paper-based Location Specific Operational Guidelines (LSOG), along with a number of other sources of information, e.g. the prevailing weather conditions and the weather forecast, to guide the decision.

To provide alternative decision support, ease the information handling and reduce dependency on the experience, competence and mental state of the individuals on duty at any given time, we envision a solution where advice is automatically generated based on a wider range of input parameters compared to the LSOG. This solution is illustrated in Figure 1. The Input Collector collects all the data for the input parameters, such as weather forecasts and sensor readings. The Decision Support Model aggregates these data to compute an advice. This
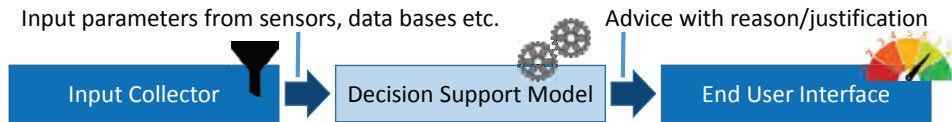
Figure 1: Vision for overall decision support solution.

advice is presented in the End User Interface, which should be tailored to the human offshore operatives making the decision.

The work presented in this paper concerns the Decision Support Model. We identified the following success criteria for the model:

C1: *The model should provide advice that correspond with expert expectations.*

C2: *The model should capture all aspects that are important for the assessment.*

C3: *The model should be comprehensible for domain experts.*

C4: *The expected benefit should justify the effort required to develop the model.*

The rest of this paper is structured as follows. First, in Section 2 we introduce the DEXi tool, before explaining the method used for the development in Section 3. Section 4 presents the decision support model, as well as expert feedback on the model. In Section 5 we discuss and evaluate the model with respect to criteria C1-C4. In Section 6, we present related work, before concluding in Section 7.

2   THE DEXi TOOL

DEXi [4] is a computer program for the development of multi-criteria decision models and the evaluation of options. Multi-criteria (also called multi-attribute) models are a class of models used for decision analysis that evaluate options according to several, possibly conflicting, goals or objectives. In this section, we introduce DEXi, focusing on the parts needed to understand the rest of this paper. For a detailed description, we refer to the DEXi User Manual [1].

A multi-attribute model decomposes a decision problem into a tree (or graph) structure. The overall problem is represented by the top attribute. All other attributes in the tree represent sub-problems, which are smaller and less complex than the overall problem. Each attribute is assigned a value. The set of values that an attribute can take is called the *scale* of the attribute. DEXi supports definition of discrete ordinal scales; typically, each step consists of a textual description. An example of an ordinal scale is {Unacceptable; Acceptable; Good; Excellent}.

Every attribute is either a basic attribute or an aggregate attribute. *Basic attributes* represent the inputs to the multi-attribute model. They have no child attributes. The value of a basic attribute is determined solely by the input to (or selected value for) the attribute.

Aggregate attributes are characterized by having child attributes (which may be basic or aggregate). The value of an aggregate attribute is a function of the values of its child attributes. This function is called the *utility function* of the attribute. The utility function of each aggregate attribute is defined by stating, for each possible combination of its child attribute values, what is the corresponding value of the aggregate attribute.

In summary, developing a DEXi model implies the following: (i) define the attributes and tree structure, (ii) define the scale for each attribute, and (iii) define the utility function for each aggregate attribute. For any given set of values for the basic attributes, the value assigned to the top attribute represents the overall aggregated evaluation.
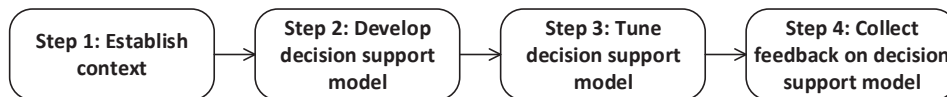
Figure 2: Overview of method.

3 METHOD FOR MODEL DEVELOPMENT

As illustrated in Figure 2, we developed the model in four steps. In the first step, we established the context by identifying the purpose and scope, as well as deciding which tool to use. In the second step, we developed the decision support model by carrying out points (i)–(iii) described in Section 2. This was primarily done during video meetings where the analysis leader shared his screen and edited the DEXi model based on input and comments from the domain experts, while the analysis secretary took notes about the reasoning and discussions. Some modifications and corrections where also done offline, through email interaction.

In the third step, we tuned the decision support model by first defining a set of six scenarios based on the following criteria: 1) All scenarios should be realistic, i.e. represent conditions that might actually occur. 2) The set should include scenarios that cover all the possible decision alternatives defined in the LSOG. After describing the scenarios textually, each of the identified scenarios was translated to an assignment of a value to each basic attribute, referred to as an *option* in the DEXi tool. This allowed us to compare the advice produced by the model for each scenario with the guidelines provided by the LSOG. In cases of mismatch, we updated the DEXi model. In the fourth step, we collected feedback on the model, with focus on model structure and outcome for the six scenarios defined in the preceding step.

As shown in Table 1, the above steps were carried out in 12 meetings held within a period of three months (from August 2016 to November 2016). All the authors took part in the model development. Of these, three are domain experts with technical experience within ship technology and marine systems in the petroleum industry, as well as software systems to support the petroleum industry with respect to risk-based decision-making. The remaining two (from SINTEF) served as analysis leader and secretary. The fourth step, i.e. feedback on the model, took place in meeting 12 (with preparations in meeting 11). The feedback was collected from three offshore operatives who represented the target group and who had not participated in developing the model or been involved in any other way before meeting 12. The feedback is explained further in Section 4. All meetings except meeting 12 were video meetings, while the 12th meeting was a combined physical and video meeting where one of the offshore operatives participated remotely from an offshore location.

Although the steps are presented chronologically, they were sometimes revisited to make updates and adjustments or to capture new factors that were brought forward by the domain experts. Roughly speaking, the first step took place in meeting 1 and meeting 2, the second step took place from meeting 3 to meeting 7, the third step took place from meeting 8 to meeting 10, and the fourth step took place in meeting 11 and meeting 12.

4 RESULTS FROM APPLYING THE METHOD

In this section, we first provide an overview of the decision support model, and then we present the feedback on the model.

The decision support model consists of 16 aggregate attributes including the top attribute and 28 basic attributes. It is beyond the scope of this paper to explain all details of the model. Instead we focus on a fragment.

Figures 3(a)–3(c) illustrate parts of the model, as shown in DEXi, starting from the top attribute and ending at three of the basic attributes. Aggregate attributes are labeled by a

Table 1: Overview of meetings. M=Meeting, D=Duration, S=Step in method.

M	Date	D	S	Activity
1	25.08.16	1.5 h	1	Establish context
2	16.09.16	2.5 h	1	Finalize context establishment, present DEXi and progress plan, and develop initial model structure
3	22.09.16	2 h	2	Continue developing model structure
4	06.10.16	2 h	2	Complete model structure, define scales for attributes and utility functions for aggregate attributes
5	13.10.16	3 h	2	Continue defining attribute scales and utility functions
6	25.10.16	3 h	2	Continue defining attribute scales and utility functions
7	27.10.16	2 h	2	Complete defining attribute scales and utility functions
8	02.11.16	2 h	3	Perform model tuning
9	11.11.16	2.5 h	3	Perform model tuning
10	24.11.16	2 h	3	Complete model tuning
11	28.11.16	1 h	4	Prepare feedback collection
12	30.11.16	6 h	4	Collect feedback from offshore operatives
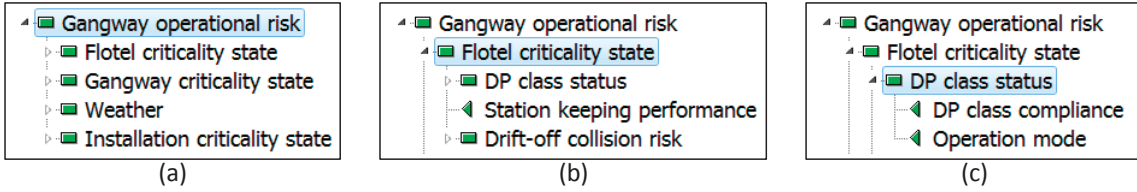
Figure 3: A small fragment of the decision support model.

small rectangle in front of their name, while basic attributes are represented by a triangle pointing horizontally to the left as illustrated in Figures 3(b) and 3(c). As can be seen in Figure 3(a), the top attribute is *Gangway operational risk*, which represents the operational risk for the gangway between the flotel and the installation.

The top attribute has the following child attributes: *Flotel criticality state*, *Gangway criticality state*, *Weather*, and *Installation criticality state*. In Figure 3(b) we have expanded the child attribute *Flotel criticality state*, which in turn has three child attributes: two aggregate attributes and one basic attribute. In Figure 3(c) we have expanded *DP class status*, which has two basic attributes as child attributes.

The value assigned to the top attribute represents the advice to the decision makers. Depending on the combinations of values assigned to the 28 different basic attributes, the top attribute is assigned one of following values {Abandon operation; Prepare to abandon operation; Advisory state; Normal state}. *Abandon operation* indicates that there are very strong reasons for an immediate disconnection of the gangway, for example since an autolift of the gangway may occur. *Prepare to abandon operation* indicates that there are strong reasons for disconnecting the gangway. Preparations for disconnection should be considered. *Advisory state* indicates an attentive state; the responsible offshore operatives need to hold an advisory meeting to assess if one or more states may be changed in order to improve the current or future operating conditions of the flotel and the gangway. Finally, *Normal state* indicates that the gangway may safely be (or remain) connected. Notice that these four values correspond directly to four risk levels, where *Abandon operation* corresponds to the highest risk and *Normal state* corresponds to the lowest risk.

Regarding feedback on model structure, the offshore operatives were asked to consider the

following three questions: Are there any important attributes that are omitted? Are there any attributes that should be removed? Are the attributes properly organized? Everyone agreed on the overall model structure. At the detailed level, we received three suggestions for additional attributes to be considered as descendants of one of the four existing attributes under *Gangway operational risk* shown in Figure 3(a). In addition, there was one suggestion for an attribute that could be removed, as it was judged to have little impact. Finally, there was one attribute for which a refinement of the scale was proposed in order allow a more fine-grained distinction between states.

With respect to feedback on outcomes for selected scenarios, the offshore operatives were asked whether they agreed with the advice produced by the model for the scenarios. They unanimously agreed for five of the six scenarios. For the sixth scenario, two expressed doubt or disagreement, even though the advice was consistent with the LSOG. The offshore operatives emphasized that the LSOG represents guidelines, rather than a set answer.

5   DISCUSSION AND EVALUATION

Based on our experience, we now discuss and evaluate the fulfillment of criteria C1-C4 defined in Section 1.

C1: In our context, expert expectations are represented by the opinions of the offshore operatives taking part in the evaluation in the final meeting, as well as the LSOG, which is based on expert knowledge. As explained in Section 3, we made sure that the advice produced by the model were consistent with the LSOG for the identified scenarios. DEXi proved to have the expressive power to achieve this without any problems. For the one scenario where the offshore operatives did not agree with the model, the disagreement was due to a discrepancy between the guidelines in the LSOG and the opinions of the offshore operatives. Hence, the contended scenario is actually an issue of resolving discrepancy between different experts. We consider our results promising, although a thorough evaluation of criterion **C1** requires a more extensive validation, preferably using more scenarios based on historical data, as well as involving more domain experts.

C2: The feedback on the model showed that the offshore operatives agreed with the overall structure and attributes. Incorporating their proposed modifications would not be a problem. Hence, we are confident that all the factors that the domain experts identified can be captured in the model. The aspects covered by the LSOG, which represent the current solution, is a proper subset of the aspects covered by the model.

However, one aspect not captured by the model is uncertainty. For example, input parameters, such as the weather forecast, are more or less uncertain. Even though the weather services provide an assessment of the uncertainty, this is ignored by the model. We considered including and aggregating uncertainty in the model, so that the advice offered as output could be accompanied by an aggregated assessment of the uncertainty. However, we saw no way to achieve this without significantly complicating the model, and the LSOG does not address the uncertainty of its guidelines. We therefore decided not to include uncertainty in the model.

While discussing **C2**, it is also interesting to touch on the issue of scalability. The most important aspect in this respect seems to be the size of the utility function for each attribute, i.e. number of possible combinations of values for its child attributes. This is determined by the number of child attributes and the granularity of their scales. The DEXi manual states that defining a utility function is quite hard for a size of 100 [1, p. 19]. In our model, the largest utility function, which belongs to the top attribute *Gangway operational risk*, has size 144. For this attribute, it was not acceptable to reduce the number of the child attributes, as the structure illustrated by Figure 3(a) was considered most appropriate. We found the size 144 to be manageable, due to functionality that DEXi offers for checking consistency of a utility function and automatically suggesting possible values for missing entries based on

already inserted entries. Still, we believe that utility functions larger than ca. 150 would be highly impractical.

C3: This criterion implies that the domain experts should be able to understand the algorithm by inspecting the model. This increases trust in the outputs from the model, and means that the model can also facilitate knowledge sharing and learning. None of the domain experts had any knowledge about DEXi before the process. Even so, after a brief introduction, they quickly grasped the DEXi concepts and were able to contribute to the model development. Throughout the process, the comments, suggestions and discussions demonstrated that all participants were able to understand the details of the evolving model. Thus, we avoided the misunderstandings and problems typically encountered when an executable algorithm is implemented in a language not understood by the domain experts. Basically, the DEXi model served as a combined specification and implementation of the assessment algorithm that was fully transparent for all participants.

C4: Our estimate indicates that the model development amounts to ca. 150 person hours in total. This includes ca. 100 hours spent on meetings 2 to 11. The estimate does not include meetings 1 and 12, as no model development or updates were done in these meetings. Ca. 50 hours was spent on work between meetings. Of the latter, ca. 16 hours was spent by the domain experts on checking the model and defining scenarios, while the remaining 34 hours was spent by the analysis leader and secretary on taking notes and correcting the model. We are not aware of other works reporting on the effort required to develop this type of model. However, for the model-based risk analysis method CORAS [5], the authors state that the expected effort required to carry out a CORAS analysis is typically from 150 to 300 hours. This gives at least an indication that the amount spent on developing our decision support model is reasonable.

Of course, a thorough evaluation of criterion **C4** would require that we quantify the benefit, as well as the cost. This is very hard, and we have not attempted to do so. Still, we believe that the benefit justifies the effort. First, the model produces consistent advice which may be a valuable supplement to a largely experience based decision making process. There is also a potential for reuse of (parts of) the model to support related decisions. Second, the process of developing the model collectively in a group creates learning and raises the awareness of those taking part. Third, the resulting model codifies and documents knowledge from all those taking part in the development, thus serving as a vehicle for knowledge transfer throughout the organization. While the first point was a central part of our motivation for initiating the work, and known in advance, the added benefit of the last two points became clear to us during the process.

6   RELATED WORK

DEXi is one of many approaches within the field of multi-criteria decision making (on which there is huge literature [8]), and has been tried out in a wide range of domains such as health care, finance, construction, cropping systems, waste treatment systems, medicine, tourism, banking, manufacturing of electric motors, and energy [3, 4]. To the best of our knowledge, DEXi has not been used to assess safety risks within offshore as reported in this paper. However, it has been applied to assess safety risks within highway traffic [6] and ski resorts [2].

The aforementioned two approaches are similar to our approach in the sense that they use DEXi models as the underlying algorithm to compute an advice based on relevant input data. In particular, the approach provided by Omerčević et al. [6] use DEXi models in a framework where input data is collected via sensors in the highway. This is in line with our envisioned automated solution illustrated in Figure 1. The details of the End User Interface and the Input Collector are beyond the scope of this paper and therefore not explained further. However, we are confident that our envisioned solution is feasible as we have in fact taken

part in implementing a similar approach in a framework for real-time cyber-risk assessment [7] developed by the WISER-project [9].

Unlike most of the existing publications on DEXi, we have focused on the overall approach, rather than the details of the model. In particular, we address the efforts spent to develop the model, the involvement of domain experts, and the comprehensibility of the model, as well as the quality of the final result. These aspects are important for others who consider a similar approach.

7   CONCLUSION

In this paper, we shared our experiences from using DEXi to develop support for risk-based decisions for offshore flotels. Our motivation was to make others who face related challenges aware of the possibilities, and help them to consider whether a similar approach is suitable for their needs. Space restrictions have prevented us from going deep into all the details of the process and resulting model. We have focused on the issues that we think are of general relevance. Based on our experience and overall evaluation, we consider our results quite promising, and believe that the approach can be fruitful for a wider range of risk-based decisions. In future projects, we hope to explore these possibilities further.

Acknowledgements

References

[1] M. Bohanec. DEXi: Program for Multi-Attribute Decision Making. User's Manual v 5.00 IJS DP-11897, DEXi, 2015.

[2] M. Bohanec and B. Delibašić. Data-Mining and Expert Models for Predicting Injury Risk in Ski Resorts. In Proc. 1st International Conference on Decision Support System Technology (ICDSST'15), pages 46–60. Springer, 2015.

[3] M. Bohanec, M. Žnidaršič, V. Rajkovič, I. Bratko, and B. Zupan. DEX Methodology: Three Decades of Qualitative Multi-Attribute Modeling. Informatica (Slovenia), 37(1):49–54, 2013.

[4] DEXi: A Program for Multi-Attribute Decision Making. http://kt.ijs.si/MarkoBohanec/dexi.html Accessed June 12, 2017.

[5] M. S. Lund, B. Solhaug, and K. Stølen. Model-Driven Risk Analysis: The CORAS Approach. Springer, 2011.

[6] D. Omerčević, M. Zupančič, M. Bohanec, and T. Kastelic. Intelligent response to highway traffic situations and road incidents. Proc. Transport Research Arena Europe 2008, pages 21–24, 2008.

[7] A. Refsdal, G. Erdogan, G. Aprile, S. Poidomani, R. Colciago, A Gonzalez, A. Alvarez, S. Gonzalez, C. H. Arce, P. Lombardi, and R. Mannella. D3.4 – Cyber risk modelling language and guidelines, preliminary version. Technical Report D3.4, WISER, 2017.

[8] M. Velasquez and P.T. Hester. An analysis of multi-criteria decision making methods. International Journal of Operations Research, 10(2):56–66, 2013.

[9] Wide-Impact cyber SEcurity Risk framework (WISER). https://www.cyberwiser.eu/ Accessed June 12, 2017.

USING BIPOLAR MIX IN THE PROCESS OF SELECTING PROJECTS APPLYING FOR CO-FINANCING FROM THE EUROPEAN UNION

Dorota Górecka

Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management,
Department of Econometrics and Statistics
Ul. Gagarina 13A, 87-100 Toruń, Poland
E-mail: dgorecka@umk.pl

Abstract: In this paper the issue of evaluation and selection of the projects applying for the support from the European Union funds using a novel tool called BIPOLAR MIX is discussed. The algorithm proposed employs the key notions of the classical BIPOLAR method introduced by Konarzewska-Gubała, and of the modified BIPOLAR method with stochastic dominance (SD) rules introduced by Górecka. Within this new MCDA approach evaluation matrix may include either deterministic or stochastic measurements of the performance of alternatives.

Keywords: decision analysis, MCDA, mixed information, BIPOLAR MIX, European Union funds, projects evaluation.

1 INTRODUCTION

The European Union Regional (Cohesion) Policy is a strategic investment policy targeting all EU regions and cities in order to support job creation, business competitiveness and sustainable development, boost their economic growth and improve people's quality of life. It is used to help less-developed EU regions to catch up and it plays an important role in the delivery of the Europe 2020 strategy. Funding for regional policy in the years 2014-2020 amounts to EUR 351.8 billion (almost a third of the total EU budget) [1]. It is extremely important to allocate these means in possibly the most effective way and that depends among other things on the proper selection of projects that are going to be co-financed. In order to help decision-makers in this challenging task the BIPOLAR MIX method can be applied. It is a novel MCDA technique based on the classical BIPOLAR method proposed by Konarzewska-Gubała (see [9, 10]) and on its alteration, namely the modified BIPOLAR method with stochastic dominance (SD) rules, proposed by Górecka (see [3, 4, 6]). It can definitely improve the appraisal and selection procedure since it accepts mixed evaluations of the alternatives (deterministic and stochastic ones), which is a desirable feature from the point of view of projects applying for co-financing from the EU. Obviously, the BIPOLAR MIX method can be used in the evaluation and ranking process of projects financed from private, state or municipal resources, and in other decision-making problems, for example risk management ones.

This paper consists of an introduction, three sections and conclusions. In the second section the motivation for developing the BIPOLAR MIX technique is given and in the third section the BIPOLAR MIX algorithm is described. Finally, the fourth section provides an illustrative example concerning the problem of ordering environmental infrastructure projects applying for co-financing from the European Union funds.

2 MOTIVATION OF THE BIPOLAR MIX METHOD

The motivation for developing the BIPOLAR MIX method stems from the characteristics of the analysed decision-making problem and the expectations of the decision-makers involved in the realisation of the EU Regional Policy, which are as follows [5]:
- the decision-making problem should be formulated as a problem of ordering a finite number of alternatives – it is indispensable to each beneficiary to be classified in the

ranking and know its own result (overall score), preferably a numerical one since otherwise it may be unconvincing for the applicants;

- there should be a possibility to employ both quantitative and qualitative criteria, and to use mixed data (deterministic and stochastic information);
- the problem is a group decision-making problem – experts engaged in the projects' appraisal individually and independently evaluate a finite number of competing projects and it is required to incorporate diverse individual views into a blended final decision;
- there is no room for the incomparability of the alternatives – ranking obtained should be complete as the explanation that a given project has not been selected for co-financing because of the incomparability with the others will not be accepted by the applicants;
- decision-makers are able to present the information about their preferences but they do not have much time for the interaction and cooperation with the analyst;
- the possibility of complete compensation occurrence should be removed – in the case of some criteria it may be hazardous and in the case of the others projects should fulfil the so called 'minimal quality';
- the possibility that a few projects will be classified on the same place in the ranking should be limited as it may create problems with dividing the funds;
- it is desired that the method enables to determine whether the highly ranked projects are really good or just better than the weak ones.

Taking into account all the above-mentioned properties of the decision-making problem analysed and its participants the BIPOLAR MIX method was designed. It is presented in the next section.

3 ALGORITHM OF THE BIPOLAR MIX METHOD

Let $F = \{f_1, f_2, ..., f_n\}$ be a set of n criteria examined (it is assumed that all criteria are maximized); $A = \{a_1, a_2, ..., a_m\}$ a set of m alternatives, and $R = \{r_1, r_2, ..., r_r\}$ a reference set consisted of two subsets: $D = \{d_1, d_2, ..., d_d\}$, $Z = \{z_1, z_2, ..., z_z\}$ – 'good' and 'bad' reference alternatives respectively, where $D \cup Z = R$, $D \cap Z = \emptyset$, $\forall_{d \in D} \forall_{z \in Z} \forall_{k=1,2,...,n} f_k(d) \geq f_k(z)$.

The BIPOLAR MIX procedure consists of the following steps:

<u>Step 1:</u>
A. Calculation of aggregated preference indices for each pair (a_i, r_j), where $a_i \in A$, $r_j \in R$:

$$c(a_i, r_j) = \sum_{k=1}^{n} w_k \varphi_k(a_i, r_j) \qquad (1)$$

where:

$$\varphi_k(a_i, r_j) = \begin{cases} 1, & if \quad F_k^i \; SD \; F_k^j \wedge \mu_k(a_i) - \mu_k(r_j) > p_k \\ -1, & if \quad F_k^j \; SD \; F_k^i \wedge \mu_k(r_j) - \mu_k(a_i) > p_k \\ \dfrac{\mu_k(a_i) - q_k - \mu_k(r_j)}{p_k - q_k}, & if \quad F_k^i \; SD \; F_k^j \wedge q_k < \mu_k(a_i) - \mu_k(r_j) \leq p_k \\ -\dfrac{\mu_k(r_j) - q_k - \mu_k(a_i)}{p_k - q_k}, & if \quad F_k^j \; SD \; F_k^i \wedge q_k < \mu_k(r_j) - \mu_k(a_i) \leq p_k \\ 0 \; otherwise \end{cases} \qquad (2)$$

or

$$\varphi_k(a_i, r_j) = \begin{cases} 1, & \text{if} \quad f_k(a_i) - f_k(r_j) > p_k \\ -1, & \text{if} \quad f_k(r_j) - f_k(a_i) > p_k \\ \dfrac{f_k(a_i) - q_k - f_k(r_j)}{p_k - q_k}, & \text{if} \quad q_k < f_k(a_i) - f_k(r_j) \le p_k \\ -\dfrac{f_k(r_j) - q_k - f_k(a_i)}{p_k - q_k}, & \text{if} \quad q_k < f_k(r_j) - f_k(a_i) \le p_k \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

depending on data, where:

- w_k - coefficient of importance for criterion f_k, $\sum_{k=1}^{n} w_k = 1$,

- F_k^i, F_k^j – distribution of the evaluations of alternative a_i and reference alternative r_j respectively with respect to criterion f_k,

- SD – stochastic dominance relation: FSD/SSD/AFSD/ASSD (see [12, 8, 11]) or OFSD/OSSD/OAFSD/OASSD (see [13, 3, 5, 7])[1],

- $\mu_k(a_i), \mu_k(r_j)$ – average performance (expected value of the evaluations' distribution) of a_i and r_j respectively on criterion f_k,

- $f_k(a_i), f_k(r_j)$ – performance a_i and r_j respectively on criterion f_k,

- q_k, p_k – indifference and preference threshold respectively for criterion f_k.

B. Calculation of credibility indices for each pair (a_i, r_j):

$$\omega(a_i, r_j) = \begin{cases} c(a_i, r_j), & \text{if} \quad c(a_i, r_j) > 0 \quad \text{and} \quad \forall k \in I^- \quad \mu_k(a_i) \ge v_k \\ c(a_i, r_j), & \text{if} \quad c(a_i, r_j) < 0 \quad \text{and} \quad \forall k \in I^+ \quad \mu_k(r_j) \ge v_k \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

where:
- $I^+(a_i, r_j) = \{k : \varphi_k(a_i, r_j) > 0\}$
- $I^-(a_i, r_j) = \{k : \varphi_k(a_i, r_j) < 0\}$
- v_k – veto threshold for criterion f_k.

Step 2:
A. Calculation of success indices for each alternative a_i:

$$d_{iS} = \frac{1}{d} \sum_{g=1}^{d} \omega(a_i, d_g), \ d_{iS} \in [-1,1]. \qquad (5)$$

Mono-sorting:
- category S1: alternatives a_i for which $d_{iS} > 0$ (type: *overgood*).

---

[1] In this paper it is assumed that decision-maker(s) is (are) risk-averse. If a decision-maker has also a decreasing absolute risk aversion, then the TSD rule (see [14]) should be additionally applied. If a decision-maker is risk-seeking, then FSD/SISD/TISD1/TISD2 rules (see [2, 15]) should be used.

- category S2: alternatives a_i for which $d_{iS} = 0$.
- category S3: alternatives a_i for which $d_{iS} < 0$ (type: *undergood*).

Mono-ranking: according to descending value of d_{iS}

B. Calculation of anti-failure indices for each alternative a_i:

$$d_{iN} = \frac{1}{z} \sum_{h=1}^{z} \omega(a_i, z_h) \, , \; d_{iN} \in [-1,1]. \tag{6}$$

Mono-sorting:
- category N1: alternatives a_i for which $d_{iN} > 0$ (type: *overbad*).
- category N2: alternatives a_i for which $d_{iN} = 0$.
- category N3: alternatives a_i for which $d_{iN} < 0$ (type: *underbad*).

Mono-ranking: according to descending value of d_{iN}.

C. Calculation of final scores for each alternative a_i:

$$d_{iSN} = \frac{d_{iS} + d_{iN}}{2}, \; d_{iSN} \in [-1,1]. \tag{7}$$

Bipolar-sorting:
- category B1: alternatives a_i for which $d_{iS} + d_{iN} > 0$ (type: *good*).
- category B2: alternatives a_i for which $d_{iS} + d_{iN} = 0$.
- category B3: alternatives a_i for which $d_{iS} + d_{iN} < 0$ (type: *bad*).

Bipolar-ranking: according to descending value of d_{iSN}.

4 APPLICATION OF THE BIPOLAR MIX METHOD FOR THE EUROPEAN PROJECT SELECTION PROCESS

The usefulness of the BIPOLAR MIX technique for the European project selection process will be illustrated by an example that concerns the applications for project co-financing by the European Regional Development Fund.

In the analysis 7 infrastructure projects were considered. They concern the surface water protection and include construction and modernization of wastewater and rainwater collection networks and wastewater treatment plants. They were evaluated using 12 criteria: 2 deterministic ones and 10 stochastic ones. Regarding the latter, 5 experts − specialists in the field of environmental protection infrastructure − scored them from 0 (the lowest evaluation) to 10 (the highest evaluation).

The model of preferences for the decision-making problem is presented in Table 1 while Table 2 provides the performance matrix for 7 projects taken into consideration in the case study and 4 reference ones (2 'good' and 2 'bad'). The results obtained by applying the BIPOLAR MIX method are given in Table 3.

Table 1: Model of preferences

f_k	Criterion	Min/max	Type of data	w_k	q_k	p_k	v_k
f_1	Total cost [PLN million]	min	deterministic	0.12	1	3	30
f_2	People using project [number]	max	deterministic	0.05	100	600	600
f_3	Efficiency [0-10; 5 experts]	max	stochastic	0.14	1	3	3
f_4	Influence on the environment [0-10; 5 experts]	max	stochastic	0.15	2	4	3
f_5	Influence on the employment [0-10; 5 experts]	max	stochastic	0.05	3	4	2
f_6	Influence on the inhabitants' health [0-10; 5 experts]	max	stochastic	0.14	3	5	2
f_7	Influence on the investment attractiveness [0-10; 5 experts]	max	stochastic	0.07	2	4	2
f_8	Influence on the tourist attractiveness [0-10; 5 experts]	max	stochastic	0.06	2	5	2
f_9	Validity of the technical solutions [0-10; 5 experts]	max	stochastic	0.08	1	3	2
f_{10}	Sustainability and institutional feasibility of the project [0-10; 5 experts]	max	stochastic	0.06	1	3	2
f_{11}	Complementarity with other projects [0-10; 5 experts]	max	stochastic	0.04	2	4	2
f_{12}	Comprehensiveness [0-10; 5 experts]	max	stochastic	0.04	2	4	2

Table 2: Performance matrix

f_k	a_1	a_2	a_3	a_4	a_5	a_6	a_7	d_1	d_2	z_1	z_2
	$f_k(a_i)$							$f_k(r_j)$			
f_1	8.42	31.55	9.24	9.25	5.93	20.00	26.01	5.00	3.00	20.00	30.00
f_2	2000	4784	9128	1550	582	3900	2782	4500	6000	750	1000
	$\mu_k(a_i)$							$\mu_k(r_j)$			
f_3	5.6	7.2	7.0	7.6	5.8	6.8	4.6	9.0	7.2	4.0	4.6
f_4	7.2	9.2	8.4	8.8	8.4	6.6	7.6	9.0	7.4	5.0	3.8
f_5	4.4	7.8	3.8	7.2	7.4	8.4	6.2	7.0	6.6	4.0	5.6
f_6	4.8	5.8	5.0	5.6	5.8	6.0	6.0	7.0	7.8	6.0	4.8
f_7	4.6	7.8	5.8	7.6	7.2	7.6	6.6	7.0	7.6	6.0	5.4
f_8	4.6	9.0	6.2	8.4	4.6	6.8	6.6	7.0	7.4	4.0	4.4
f_9	7.8	8.4	7.6	8.4	8.2	6.6	8.2	7.0	7.2	5.0	4.2
f_{10}	7.4	8.4	6.4	7.4	8.6	9.0	8.0	7.0	8.0	5.0	4.0
f_{11}	7.0	8.6	3.4	3.8	7.4	7.0	7.4	8.0	6.4	4.0	5.0
f_{12}	4.8	9.0	4.0	4.6	5.6	5.4	5.8	8.0	6.8	4.0	5.2

Table 3: Rankings of the projects obtained using the BIPOLAR MIX method

No.	Monorankings of the projects				Bipolar-ranking of the projects	
	a_i	Success indices d_{iS}	a_i	Anti-failure indices d_{iN}	a_i	Final score d_{iSN}
1	a_2	-0.142	a_4	0.634	a_4	0.210
2	a_3	-0.177	a_3	0.518	a_3	0.170
3	a_5	-0.195	a_6	0.396	a_6	0.089
4	a_4	-0.214	a_1	0.367	a_1	0.029
5	a_6	-0.218	a_7	0.306	a_7	0.006
6	a_7	-0.294	a_2	0.000	a_2	-0.071
7	a_1	-0.310	a_5	0.000	a_5	-0.097

According to the analysis conducted all the projects considered in the case study belong to the category S3 (*undergood*) and neither of them belongs to the category N3 (*underbad*). Taking into account final scores, into category B1 (so called 'good alternatives') 5 projects were classified, namely: a_1, a_3, a_4, a_6, and a_7. Project a_4 turned out to be the strongest and project a_3 was second-strongest. Projects a_2 and a_5, in turn, were classified into category B3 (so called

'bad alternatives') and should not be recommended for co-financing. This is due to the fact that project a_2 is very expensive (it costs PLN 31.55 million) and project a_5 is characterised by a very low number of people enjoying it (582), which was clearly caught by the BIPOLAR MIX method thanks to the veto procedure applied in this technique.

5 CONCLUSIONS

The BIPOLAR MIX method proposed in this paper is a functional technique that can enhance the European project appraisal procedure and improve the decision-making process. On the one hand it is not too simple (in order to limit the temptation of manipulating the results), and on the other hand it is not too complicated (in order to enable decision-maker(s) to understand how it works). Moreover, it allows us to use mixed information (deterministic and stochastic evaluations) and it eliminates the possibility of full compensation as well as the problem of the alternatives' incomparability. Besides, it enables ranking and sorting alternatives along with determining their quality, taking into account the reference system determined by the decision-maker(s). Finally, it allows us to obtain a numerical final score and it is not tedious and time-consuming for decision-makers.

References

[1] European Commission - Regional Policy – Inforegio. http://ec.europa.eu/regional_policy/en/ [Accessed 01/06/2017].

[2] Goovaerts, M.J., De Vylder, F., Haezendonck, J. 1984. Insurance Premiums: Theory and Applications. Amsterdam: North-Holland.

[3] Górecka, D. 2009. Wielokryterialne wspomaganie wyboru projektów europejskich. Toruń: TNOiK „Dom Organizatora".

[4] Górecka, D. 2010. Wykorzystanie metod wielokryterialnych w procesie oceny i wyboru wniosków o dofinansowanie realizacji projektu z funduszy Unii Europejskiej. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Nr. 108: 76-91.

[5] Górecka, D. 2011. On the choice of method in multi-criteria decision aiding process concerning European projects. In Trzaskalik, T., Wachowicz, T. (Eds.). Multiple Criteria Decision Making'10-11 (pp. 81-103). Katowice: Publisher of The University of Economics in Katowice.

[6] Górecka, D. 2014a. Metoda BIPOLAR z dominacjami stochastycznymi. In Trzaskalik, T. (Ed.). Wielokryterialne wspomaganie decyzji. Metody i zastosowania (pp. 149-152). Warszawa: PWE.

[7] Górecka, D. 2014b. Reguły wyboru oparte na relacji prawie dominacji stochastycznej dla kryteriów ocenianych w skali porządkowej. In Trzaskalik, T. (Ed.). Wielokryterialne wspomaganie decyzji. Metody i zastosowania (pp. 31-32). Warszawa: PWE.

[8] Hadar, J., Russell, W. 1969. Rules for Ordering Uncertain Prospects. American Economic Review, Vol. 59: 25-34.

[9] Konarzewska-Gubała, E. 1989. Bipolar: Multiple Criteria Decision Aid using the bipolar reference system. Paris: LAMSADE.

[10] Konarzewska-Gubała, E. 1991. Wspomaganie decyzji wielokryterialnych: system BIPOLAR. Wrocław: Wydawnictwo Uczelniane Akademii Ekonomicznej we Wrocławiu.

[11] Leshno, M., Levy, H. 2002. Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance. Management Science, Vol. 48: 1074-1085.

[12] Quirk, J.P., Saposnik, R. 1962. Admissibility and Measurable Utility Functions. Review of Economic Studies, Vol. 29: 140-146.

[13] Spector, Y., Leshno, M., Ben Horin, M. 1996. Stochastic dominance in an ordinal world. European Journal of Operational Research, Vol. 93: 620-627.

[14] Whitmore, G.A. 1970. Third-Degree Stochastic Dominance. American Economic Review, Vol. 60: 457-459.

[15] Zaras, K. 1989. Les dominances stochastique pour deux classes de function d'utilite: concaves et convexes. RAIRO/RO, Vol. 23: 57-65.

DECISION MAKING WITH THE ANALYTIC NETWORK PROCESS

Nikola Kadoić
Faculty of Organisation and Informatics
Pavlinska 2, Varaždin, Croatia
E-mail: nkadoic@foi.hr

Nina Begičević Ređep
Faculty of Organisation and Informatics
Pavlinska 2, Varaždin, Croatia
E-mail: nbegicev@foi.hr

Blaženka Divjak
Faculty of Organisation and Informatics
Pavlinska 2, Varaždin, Croatia
E-mail: bdivjak@foi.hr

Abstract: One of the most advanced and complex multi-criteria decision-making methods is the analytic network process (ANP). This method supports modelling dependencies and feedback between elements in the network. For this reason, the ANP is one of the most appropriate methods for making decisions in fields that are characterised by existing dependencies of higher-level elements on lower level elements. In addition to reviewing the ANP, this paper also studies some possible method upgrades that might decrease the complexity of the original ANP. We explore this by structuring the problem as a weighted graph and using the concept of compatibility between interdependent matrices in the ANP.

Keywords: analytic network process, dependencies, influences, feedback, structuring, weighted graph, interdependent matrices

1 INTRODUCTION

When we talk about multi-criteria decision making, many methods can be used. The most well-known multi-criteria decision-making method is the analytic hierarchy process (AHP). In that method, the decision-making problem is decomposed into a hierarchy. At the top of the hierarchy is the decision-making goal. The criteria are on the next level, which can be decomposed to the sub-criteria (and further decomposed to the lower levels). On the last level are the alternatives. By using pairwise comparisons (to be explained later in this paper), local priorities of alternatives as well as criteria weights are calculated. Then, it is possible to calculate global priorities of alternatives and make decisions.

In the decision-making problem field, if influences/dependencies exist between criteria, which the AHP does not consider, using the AHP might lead to a decision that is less than optimal. In those cases, using the analytic network process (ANP) is more appropriate. By using the ANP, we can model the dependencies and feedback between the decision-making elements, and calculate more precise weights of criteria, and local and global priorities of alternatives. In this paper, we will describe the ANP method, present its steps using a demonstrative example (Section 2), address some weaknesses of the method based on a literature review and our experience, and propose some upgrades to the ANP that might impact on eliminating the identified weaknesses (Section 3).

2 THE ANALYTIC NETWORK PROCESS (ANP)

The decision-making problems in the ANP are modelled as networks, not as hierarchies as with the AHP. The ANP is a generalisation of the AHP. Figure 1 presents the structural

differences between a linear hierarchy and a nonlinear network. The basic elements in the hierarchy and network are clusters (components; rectangles and ellipses in Figure 1), nodes (elements in clusters, not specified in Figure 1) and dependencies (arcs). The meaning of 'depend on' is the opposite of 'have an influence on'.
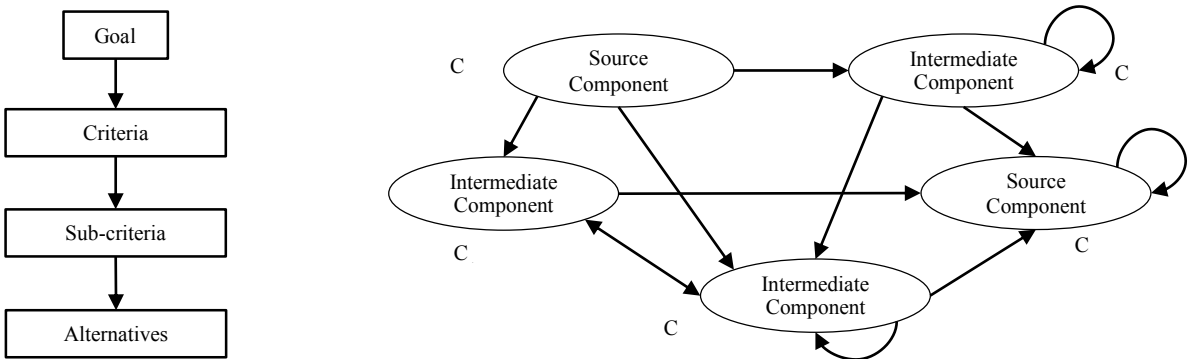
Figure 1: Structural difference between hierarchy and network (adapted from [1])

The left side of Figure 1 shows a linear network (hierarchy) in which elements from the lower level of the network have an influence on a higher level, e.g. criteria have an influence on the goal, which means that the goal depends on the criteria. On the right side of Figure 1, we have a network of clusters and some possible dependencies between them. In this case, it is possible for a cluster to depend on another cluster, but at the same time, it can influence the same one or even itself.

The steps of the ANP [2], [3] will be described through a simple example, *evaluation of scientists*.
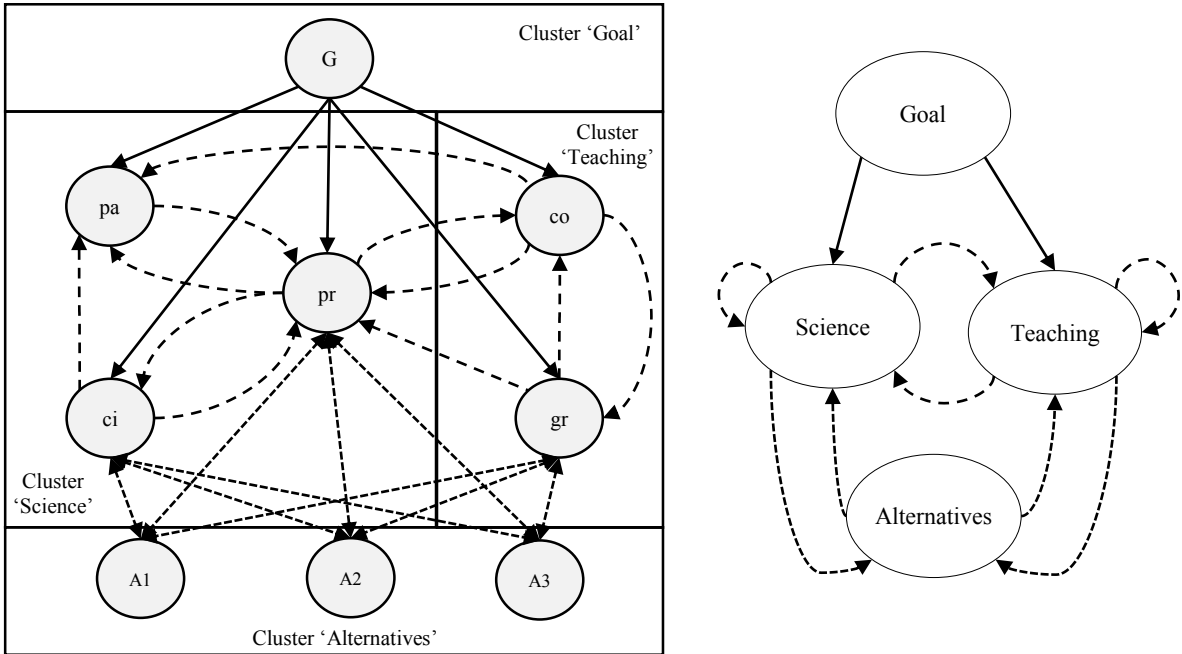
Figure 2: Structure of decision-making problem, 'evaluation of scientists'

Problem structuring. The goal of the decision making is to select the best scientists among three scientists. This means that we will have: (1) cluster *Goal* with node *G*; (2) clusters *Alternatives* with three nodes, alternatives *A1*, *A2* and *A3*; and (3) five nodes of criteria

(*papers – pa*, *citations – ci*, *projects – pr*, *courseware – co*, *grades from students – gr*) grouped into two clusters (*Science* – first three criteria and *Teaching* – last two criteria).

The problem structure is shown in Figure 2 (left). Solid arcs (arrows) are related to the dependencies of the goal on criteria. Dashed arcs are related to the dependencies between criteria. Dotted arcs are related to the dependencies of alternatives on criteria and dependencies of criteria on alternatives. Arcs between alternatives and criteria, *pa* and *co*, are not shown because of the complexity of the figure, but those dependencies also exist. The decision-making structure can be shown in a simpler way (with information lost), Figure 2 (right). The *Goal* is a source cluster that depends on *Science* and *Teaching* (black arcs). The other three clusters are intermediate components. The dashed arc between *Science* and *Teaching* is the result of the existence of at least one arc between at least one criteria from *Science* and at least one criteria in *Teaching*, e.g. *pr* depends on *co*. The dashed arc between *Teaching* and *Science* can be interpreted similarly. The loops in the clusters, *Science* and *Teaching*, are the result of the existence of at least one dependency of one criterion to another in the same cluster, e.g. *ci* depends on *pa*, or *gr* depends on *co*.

Pairwise comparisons on node level. Now, we should create the unweighted supermatrix. It is a square matrix of all nodes in the decision-making problem and contains local priorities. When making judgements in pairwise comparisons, we use Saaty's fundamental scale of absolute numbers [4], just as with the AHP. The scale has nine different intensities: 1 means that two elements in a pair are equally important with respect to the higher level element; 9 means extreme importance of one element over another. All real numbers between 1 and 9 can be used. Because of the axiom of reciprocity, we use reciprocal numbers of 1–9 as well [5]. When making the pairwise comparisons, we must take care of inconsistencies. If we say that element A is greater than element B, and B is greater than C, then, because of transitivity, A is greater than C. There is an inconsistency ratio, a measure that describes how inconsistent the decision maker was during the pairwise comparisons procedure. The allowed inconsistency ratios are all under 10%. See more in [3].

To fill the unweighted supermatrix (Table 1) with priorities, we have to make pairwise comparisons of nodes with respect to other nodes. The comparisons that have to be done are:

- Comparisons of the criteria with respect to the goal (see black arcs in Figure 2): comparisons of criteria in *Science* with respect to *G* (local priorities will be put into the supermatrix at rows *ci*, *pr*, *pa* and column *G*); comparisons of criteria in *Teaching* with respect to *G* (local priorities will be put into the supermatrix at rows *co*, *gr* and column *G*);

Calculations to get priorities for *co* and *gr* are as follows:

G	co	gr
co	1	0.5
gr	2	1
	3	1.5

0.33	0.33
0.67	0.67

AVG
0.33
0.67

Create a matrix of comparisons (coloured in grey). Put 1 on the diagonal (*co* is equally important as *co*). Make pairwise comparisons to fill the other cells. We ask the question, 'With respect to the goal, which criterion is more important, *co* or *gr*?' Let us say that the answer is that *gr* is more important than *co*, 2 on Saaty's scale. We put 2 on position (*gr*, *co*) and the reciprocal number (0.5) on position (*co*, *gr*) in a matrix of comparisons. Then, we sum the columns. We make the new matrix in which each value from the comparison matrix will be divided by the related column

sum. Then, we calculate the average of the rows of that second matrix. We fill the supermatrix with (*co, G*) = 0.33 and (*gr, G*) = 0.67;

- Comparisons of criteria with respect to other criteria – comparisons of the criteria that leave (influence) the same criterion from the same cluster with respect to it (see dashed arcs in Figure 2): *pa* and *pr* with respect to *ci*; *pa* and *ci* with respect to *pr*; *pa* and *pr* with respect to *co*. Priorities are put into the supermatrix depending on which criteria (rows) influence which criterion (column). When making a pairwise comparison between *pa* and *pr* with respect to *ci*, we try to answer the question, 'Which criterion, *pa* or *pr*, has a higher influence on criterion *ci*, and by how much?' When some criteria depend on only one criterion in the same cluster, we do not make a comparison, and write 1 in the related cell in the supermatrix, e.g. cell (*pr, pa*) = 1;
- Comparisons of alternatives with respect to each criterion (see dotted arcs from criteria to alternatives in Figure 2), the same as in the AHP. This will fill part of the supermatrix as follows: rows *A1*, *A2* and *A3*, columns *co, gr, pa, ci, pr*;
- Comparisons of criteria in each cluster with respect to each alternative (see the dotted arc from alternatives to criteria in Figure 2). This will fill part of the supermatrix as follows: rows *co, gr, pa, ci, pr*, columns *A1*, *A2* and *A3*. For example, let us say that alternative *A1* has a good value in terms of criterion *co*, and a very bad value in terms of *gr*. Let us use 4 on Saaty's scale to describe this importance; then, at column *A1*, in rows *co* and *gr*, we will write 0.2 and 0.8, respectively.

Pairwise comparisons on a cluster level. The goal of this step is to convert the unweighted matrix into the weighted supermatrix. For this, we have to do the following comparisons:

- Compare two clusters of criteria with respect to the *Goal*. For example, if we say that cluster *Science* is more important than *Teaching*, 3 in Saaty's scale, by using the same pairwise comparisons procedure as when comparing the nodes, we will get weights 0.25 (*Teaching*) and 0.75 (*Science*). This means that 0.25 will multiply *Teaching*'s criteria and 0.75 will multiply *Science*'s criteria in column *G*;
- Compare clusters *Teaching*, *Science* and *Alternatives* with respect to *Teaching*;
- Compare clusters *Teaching*, *Science* and *Alternatives* with respect to *Science*; and
- Compare clusters *Teaching* and *Science* with respect to *Alternatives*.

Table 1: Unweighted supermatrix

	G	co	gr	pa	ci	pr	A1	A2	A3
G	0	0	0	0	0	0	0	0	0
co	0.33	0	1	0	0	1	0.2	0.4	0.6
gr	0.67	1	0	0	0	0	0.8	0.6	0.4
pa	0.25	1	0	0	0.4	0.6	0.2	0.4	0.6
ci	0.25	0	0	0	0	0.4	0.2	0.5	0.3
pr	0.5	0	1	1	0.6	0	0.6	0.1	0.1
A1	0	0.5	0.5	0.2	0.4	0.6	0	0	0
A2	0	0.3	0.2	0.2	0.5	0.3	0	0	0
A3	0	0.2	0.3	0.6	0.1	0.1	0	0	0

Table 2: Weighted supermatrix

	G	co	gr	pa	ci	pr	A1	A2	A3
G	0	0	0	0	0	0	0	0	0
co	0.08	0	0.2	0	0	0.2	0.1	0.2	0.3
gr	0.17	0.2	0	0	0	0	0.4	0.3	0.2
pa	0.186	0.2	0	0	0.10	0.12	0.1	0.2	0.3
ci	0.186	0	0	0	0	0.08	0.1	0.25	0.15
pr	0.383	0	0.2	0.25	0.15	0	0.3	0.15	0.05
A1	0	0.3	0.3	0.15	0.30	0.36	0	0	0
A2	0	0.18	0.12	0.15	0.375	0.18	0	0	0
A3	0	0.12	0.18	0.45	0.075	0.06	0	0	0

Calculating the limit matrix. In this step, the weighted matrix is multiplied by itself as long as all of its columns become equal. This is how we get the final priorities. After this step, the **sensitivity analysis** is performed. Software called Superdecisions supports all the math behind the ANP (webpage: https://www.superdecisions.com/, Creative Decisions Foundation). Therefore, we did not go into detail about what this particular software supports.

3 PROPOSALS OF THE ANP UPGRADES

We focussed on the steps that users have repeatedly used, because according to our own experiences, many still do not completely understand the ANP. For example, the field of higher education is characterised by existing dependencies between criteria and feedback in decision-making problems. However, the literature review about which decision-making methods have been used in practice to solve problems showed that the AHP method was used most, and the ANP was rarely used [6]. The weaknesses of the ANP are related to the complexity of the method, the duration of implementation, and uncertainty in giving judgements, especially those on the cluster level [7]. When looking at the supermatrix, we conclude that the column of the goal and the rows of the alternatives are related to the AHP. This part is understood by the users. Some problems can appear in situations with a large number of alternatives, which can increase the duration of decision making. One solution to this is *ratings* as explained in [1], [4].

Our focus, in terms of proposing upgrades to decrease some ANP weaknesses, is related to calculating other parts of the supermatrix, not those that are related to the AHP. There are two parts of the supermatrix that can be calculated differently:
- The first part is related to the priorities of criteria with respect to criteria (because of dependencies between the criteria in the network); and
- The second part is related to the priorities of criteria with respect to alternatives (because of dependencies of alternatives on criteria).

The first upgrade requires a slightly different problem structuring procedure than the regular ANP. In the ANP, we model dependencies between criteria. Then, to make comparisons between criteria with respect to the criteria, we have to know the intensity of influences (dependencies) between criteria. We propose structuring the problem by using a weighted graph. So, to start, when we model dependencies between criteria, we also define the intensities of dependencies. Indeed, during the problem-structuring procedure, to draw an arc between two criteria, the decision maker should think deeply about the relationship between two criteria. During that process, (s)he is evaluating the influence of the dependency between two criteria. A similar problem-structuring procedure can be found in the decision-making trial and evolution laboratory (DEMATEL) method. In DEMATEL, instead of the dependencies, arcs represent influences between criteria. Thus, it is possible to structure the problem by using the approach from the DEMATEL and transform it from graphs with influences onto the graph of dependencies, keeping the intensities of dependencies the same. In DEMATEL, the intensities of the influences between criteria are measured on a scale of 5 degrees: 0 means no influence, and 4 means very high influence [8].

Now, when we have a weighted graph of dependencies between criteria, we can automatize the calculation of the limit matrix. One method of automatization is to apply the *normalisation by sum* and another method is to use a *matrix of transition*. A proposed matrix of transition is given in Table 3.

Table 3: Matrix of transition

Difference in terms of intensity of influence	Judgement on Saaty's scale
0 (1-1, 2-2, 3-3, 4-4)	1
1 (2-1, 3-2, 4-3)	2
2 (3-1, 4-2)	4
3 (4-1)	6

For example, on one hand, to calculate the local priorities of *pa* and *pr* in column *ci* with identified intensities of influences 2 and 3 for *pa->ci* and *pr->ci*, respectively, we apply the

normalisation by sum and get priorities 0.4 and 0.6. On the other hand, if we use a *matrix of transition* for the difference 1 in terms of intensity of influence, we get priorities 0.33 and 0.67. We tested this approach on several examples, and at times, one method showed closer results to the ANP results, and sometimes, the other. An average of both methods can also be an approach.

In this ANP upgrade, local priorities in terms of dependencies between criteria are calculated automatically. The main advantage of this approach is that the total implementation process takes less time than the regular ANP. Also, users do not have to make judgements and risk making wrong judgements because of their misunderstanding of pairwise comparisons of two criteria with respect to a third.

The second upgrade is related to applying the concept of compatibility between interdependent matrices in the ANP. By using this approach, the process of calculating the priorities of criteria with respect to alternatives can be shortened. An analysis of the compatibility between interdependent matrices in the ANP is explained in the paper [9].

For example, the original data about values of alternatives A1, A2 and A3 in terms of criteria *pa*, *ci*, *pr* are given in Table 4.

Table 4: Evaluation of scientists

Alternatives	pa	ci	pr
A1	5	40	6
A2	5	50	3
A3	15	10	4

When calculating the priorities of the alternatives with respect to criteria (as with the AHP), we made comparison tables as illustrated in Table 5. When we make comparisons of criteria per alternatives, we are taking into account the same data from Table 4 when we make comparisons of the alternatives with respect to the criteria. The concept of compatibility between interdependent matrices in the ANP can now be applied.

Table 5: Comparisons of alternatives with respect to criteria

pa	A1	A2	A3
A1	1	1	0.25
A2	1	1	0.25
A3	4	4	1

ci	A1	A2	A3
A1	1	0.5	3
A2	2	1	4
A3	0.33	0.25	1

pr	A1	A2	A3
A1	1	5	4
A2	0.2	1	0.5
A3	0.25	2	1

When we make comparisons of criteria with respect to, for example, *A1*, we can make any matrix of comparisons of criteria that is consistent at the local level, but inconsistent at the global level. An example is given in Table 6. The inconsistency ratio of the matrix is 0.00, but the comparisons are illogical. In Table 4, we see that *A1* has a low value in terms of criteria *pa*, and a high value in terms of criterion *pr*. That means that *pr* dominates over *pa*; however, in Table 6, we make the opposite judgement with an acceptable inconsistency ratio.

Table 6: Comparisons of criteria with respect to A1

A1	*pa*	*ci*	*pr*
pa	1	2	4
ci	0.5	1	2
pr	0.25	0,5	1

However, to use the concept of compatibility between interdependent matrices in the ANP, we need to do one comparison matrix, and others can be calculated automatically. An example is given in Table 7. Comparisons with respect to *A1* have to be manually input, while comparisons with respect to A2 and A3 are automatically calculated.

Let us say that *pr* dominates over *pa* with respect to *A1* – with 5 on Saaty's scale (according to the values in Table 4). Because *A1* and *A2* are equally important with respect to *pa*, and A1 dominates over *A2* with respect to *pr* with 5 (see Table 5), *pr* becomes equally important as *pa* with respect to *A2*. Similarly, we calculate other values.

Table 7: Comparisons of criteria with respect to alternatives

A1	pa	ci	pr
pa	1	0.33	0.2
ci	3	1	0.5
pr	5	2	1

A2	pa	ci	pr
pa	1	0.17	1
ci	6	1	5
pr	1	0.2	1

A3	pa	ci	pr
pa	1	0.25	3.2
ci	4	1	0.5
pr	0.31	2	1

4 CONCLUSION

In this paper, we gave an overview of the ANP method with a detailed illustration of the steps that we find crucial in the ANP, and which are often still not understood by users. Conducting the ANP is a time-consuming activity, and some steps are very challenging. Therefore, we proposed two upgrades of how to automatize some parts of the ANP to be less complex and more appropriate for users.

References
[1] T. L. Saaty and L. G. Vargas, *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks*. Springer; Softcover reprint of hardcover 1st ed. 2006 edition (December 28, 2009), 2006.

[2] T. L. Saaty and B. Cillo, *A Dictionary of Complex Decision Using the Analytic Network Process, The Encyclicon, Volume 2*, 2nd ed. Pittsburgh: RWS Publications, 2008.

[3] N. Begičević, "Višekriterijski modeli odlučivanja u strateškom plniranju uvođenja e-učenja," University of Zagreb, Faculty of organization and informatics, 2008.

[4] T. L. Saaty, "Decision making with the analytic hierarchy process," *Int. J. Services Sciences*, vol. 1, no. 1, pp. 83–98, 2008.

[5] P. T. Harker and L. G. Vargas, "The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process," *Management Science*, vol. 33, no. 11, pp. 1383–1403, Nov. 1987.

[6] N. Kadoić, N. Begičević Ređep, and B. Divjak, "E-learning decision making: methods and methodologies," in *Re-Imagining Learning Scenarios*, 2016, vol. CONFERENCE, no. June, p. 24.

[7] N. Kadoić, B. Divjak, and N. Begičević Ređep, "Effective Strategic Decision Making on Open and Distance Education Issues," in *Diversity Matters!*, 2017, pp. 224–234.

[8] J. Shao, M. Taisch, M. Ortega, and D. Elisa, "Application of the DEMATEL Method to Identify Relations among Barriers between Green Products and Consumers," *17th European Roundtable on Sustainable Consumption and Production - ERSCP 2014*, pp. 1029–1040, 2014.

[9] L. C. Leung, Y. V Hui, and M. Zheng, "Analysis of compatibility between interdependent matrices in ANP," *Journal of the Operational Research Society*, vol. 54, no. 7, pp. 758–768, Jul. 2003.

ASSESSING STAKEHOLDERS' INFLUENCE ON THE RESPONSIBILITY OF RESEARCH PROJECTS: APPLICATION OF ANALYTIC NETWORK PROCESS

Iván Ligardo-Herrera
Energy Engineering Institute (IIE), Universitat Politècnica de València
46022 Valencia, Spain
E-mail: ivliher@doctor.upv.es

Tomás Gómez-Navarro
Energy Engineering Institute (IIE), Universitat Politècnica de València
46022 Valencia, Spain
E-mail: tgomez@dpi.upv.es

Hannia Gonzalez-Urango
Ingenio (CSIC-UPV), Universitat Politècnica de València
46022 Valencia, Spain
E-mail: hangonur@doctor.upv.es

Abstract: In this paper we present a methodology to measure stakeholders´ influence within a project from the point of view of the responsibility. The methodology is based on a combination of some stakeholder analysis techniques and the multicriteria technique Analytic Network Process, which allows evaluating and ranking the influence among stakeholders using a responsible approach. The definition of influence is based on different criteria in the framework of a Responsible Research and Innovation that together describe an index which measures the influence of each stakeholder with respect to the responsibility of a research project. The main aim is to provide answers and guide towards how to evaluate the stakeholders of a research project in the framework of responsibility as a key aspect for its success.

1 INTRODUCTION

Responsibility has reached all disciplines including research and innovation (R & I) teams. Scientists should take into consideration their unwanted research process and outcomes impacts, and be responsible for them. With the intention of fostering responsible research, no matter whether it is basic or applied, publicly or privately funded, the European Commission has been promoting a cross-cutting issue named "Responsible Research and Innovation (RRI)". The most widely used definition of RRI could be the one given by Von Schomberg [9] (p. 9): *'(RRI) is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products'*. In this sense how responsibility has been framed has varied over time and place [10]. In terms of science the questions would be about the aims and the consequences of any research, or innovation activity [4].

Several works under the auspices of the European Commission have found that RRI involves a dialogue between stakeholders during the whole research and innovation process in order to better align both the process and its outcomes. Six key areas for that dialogue were identified: Public Engagement; Gender Equality; Science Education; Open Access; Ethics; and Governance [5]. Recently two more areas have been added, sustainability (environmental); and Social Justice [11].

Therefore when a research group decides to do responsible research, it is necessary to orient it towards what the stakeholders need or towards those aspects that do not harm them. Corporate Social Responsibility (CSR) and Project Management theories, highlight the relevance of a detailed analysis of stakeholders and their impact [3]. The answer to the question of how to evaluate the stakeholders of a project in the framework of responsibility has to focus on the integration of stakeholder management in their activities. Stakeholder management includes the process required to identify stakeholders, analyze stakeholders' expectation and their impact on the project [2]. Identifying stakeholders is closely related to the analysis of their influence and potential impact on the success of the project [1].

In the framework of a research project, stakeholders to consider should be those influential in the project [7]. This influence can be determined considering the following two aspects: the first one related to the success of the project. Here stakeholders are studied because they have different implications for the success of the project. This is the way in which they have traditionally managed by research groups and it is possible to find several works related [2, 6, 7]. The second point is related to the responsibility of the project, meaning the way in which the project answers for those possible consequences of research or innovation activity. Stakeholders can contribute at different levels, to making the project more responsible to society.

The responsability of a project should be considered as a criteria to evaluate the sucess of a project. This success could be defined by aspects such as legal permits, concessions, usability, media support and responsibility, among others. A global assessment of the influence of stakeholders, should include all of those aspects. This paper proposes a development based only on how to evaluate the stakeholders of a project in the framework of responsibility as a key aspect for its success.

The stakeholder management traditionally used criteria like: power, legality and urgency. This work propouse an assessment appling a MCDM technique called ANP to response to these two researh questions: i) How stakeholder's influence is useful for project's responsibility? and ii) How to evaluate project stakeholders within the framework of responsibility as a key aspect of project success? To solve these questions we propose a dialogue with stakeholders to identify key points that researchers should consider.

2 METHODOLOGY

Fig. 1 shows a general schematic that describes the overall flow of the proposal used in this research based on the ANP procedure [8], stakeholder theory [1, 2, 7] and the framework of RRI [4, 10]. The following section explains how the methodology was applied to the project.
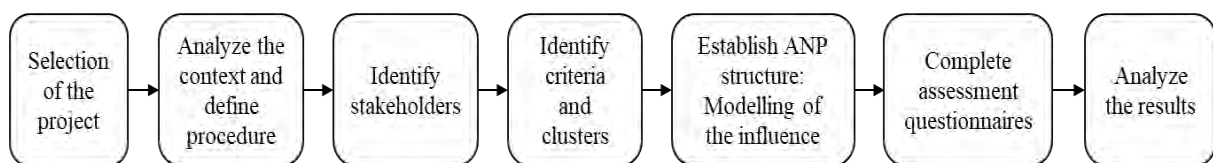
Figure 1 : Methodology proposed

3 MODELLING AND ASSESSING RESPONSIBILITY

The model has been applied to an ongoing project. The project's aims is develop a real-time recommendation system containing dynamic content based on the context of the user in mobility and their social networks to reduce the human interaction with the mobile device and improve the user's experience. In this case the project is developed for improving the

tourist experience and support local businesses for a city. It is developed by a multidisciplinary team from University, Local Tourist office and private sector.

Identification of stakeholders

In order to identify key stakeholders a network analysis based on the snowball method was applied. A first group was defined according to some documents, previous experience and actual members of the project consortium. A snowball procedure based on the information given by the first group to further identify more stakeholders was then carried out. Seven stakeholders were identified:
- S1. Users: Tourists, visitors or residents.
- S2. Business: Anyone who offers an activity of leisure or entertainment in the city. For example: restaurants, museums, hotels, mobility and transportation, concerts, events, exhibitions, etc.
- S3. Local administration.
- S4. Developers of digital content.
- S5. Neighborhood associations: They are directly affected or benefited by tourism.
- S6. NGO's: Interested in assessing the social impacts of tourism.
- S7. Financial support: They guarantee the economic viability of the project.

Identification of criteria and clusters

In order to identifying the rest of the network elements and their relationships, criteria and clusters were determined. Criteria which could evaluate the influence of stakeholders in the responsibility of the project were identified. It was necessary to make sure that these criteria could be grouped, that they were relevant, not redundant and related to the RRI approach. The final list of 16 criteria grouped in three clusters was defined on the basis of a bibliographic review and with the assistance of some of the experts and members of the project team. The clusters are:
- Cluster 1. RRI Areas: Group criteria aimed at assessing the knowledge that one stakeholder specifically possesses related to RRI concepts. In general this is a weak point, since there is a general lack of knowledge on the topic, which implies the need to inform the stakeholders from the most basic concepts of responsibility. The Criteria of this cluster are the eight key areas of the RRI.
- Cluster 2. Diffusion: Refers to the ability of one stakeholder to spread the project and to generate debates and relationships related to the project environment. Also to generate o identify relevant aspects. This cluster has four criteria.
- Cluster 3. Codetermination: Refers to the willingness and capability of one stakeholder to provide the project with resources. Four criteria compound this cluster.

Modelling the influence assessment with the ANP model

In this section, we will have to follow all the steps proposed by Saaty [8] for the ANP method. After the identification of the model elements, influences among them were determined using a relationship matrix with the help of the experts and the project team. The proposed model is illustrated by the network shown in Figure 2. The bidirectional arrows indicate influences between clusters in both directions.
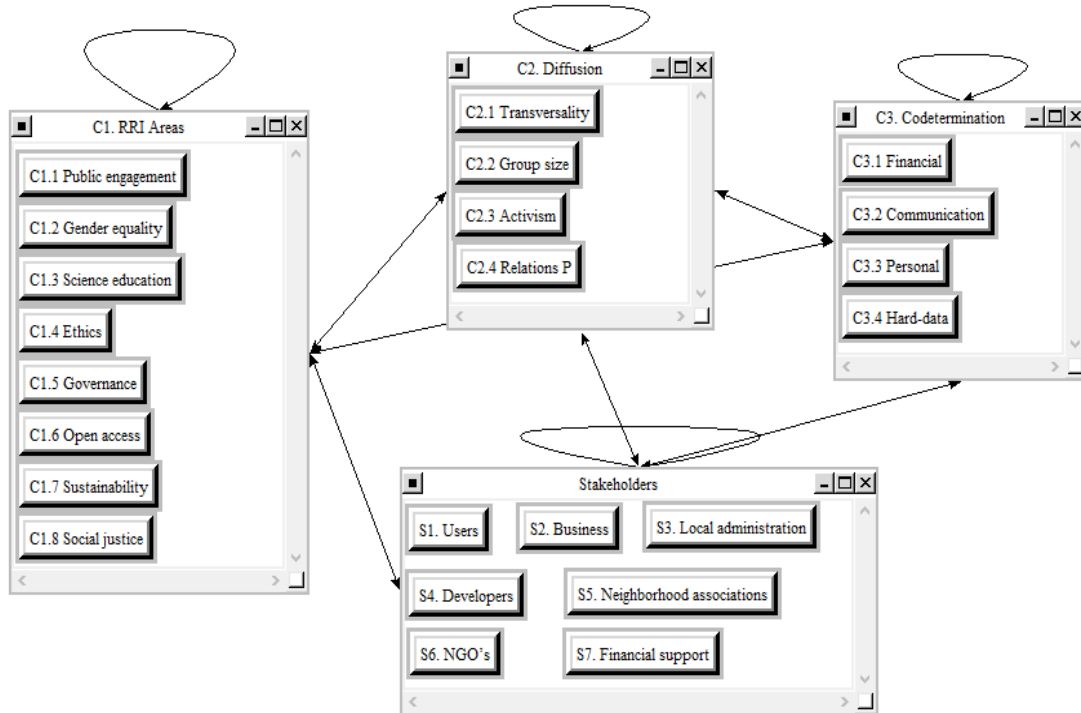
Figure 2: ANP network model of the case study.

To determine the weights of the criteria and stakeholders of the model a questionnaire was designed in order to assess to what extent each element has some influence on others elements to which it is related. For the assessment process two expert were selected. Both are researcher at the Ingenio Institute (CSIC-UPV) have been considered. This is a recognized Institute with experience in social analysis.

All the calculations were performed using the Superdecision© v.2.0.8. software. Once experts have finished all pairwise comparisons, judgement aggregation was performed using the Geometric Mean in order to obtain a global judgement [8].

Analysis of results
The final limit matrix shows the priority obtained for each criterion, a non-dimensional value that can be considered as its relative importance. Results show (
Figure 3) that altogether, the clusters have similar values C1. RRI Areas (0.345), C2. Diffusion (0.332) and C3.Codetermination (0.324).

Results for each criteria show that the most important one is C3.2 Communication (0,134) followed by C2.1 Transversality (0,102), and C2.4 Relations with the project (0,097). Other strong group are C1.1 Public engagement (0,068), C3.1 Financial (0,080), C2.3 Activism (0,068), C2.2 Group size (0,064), C3.3 Personal (0,062), C1.8 Social justice (0,059) And C1.5 Governance (0,059). The least valued are C1.4 Ethics (0,025) and C1.2 Gender equality (0,018).
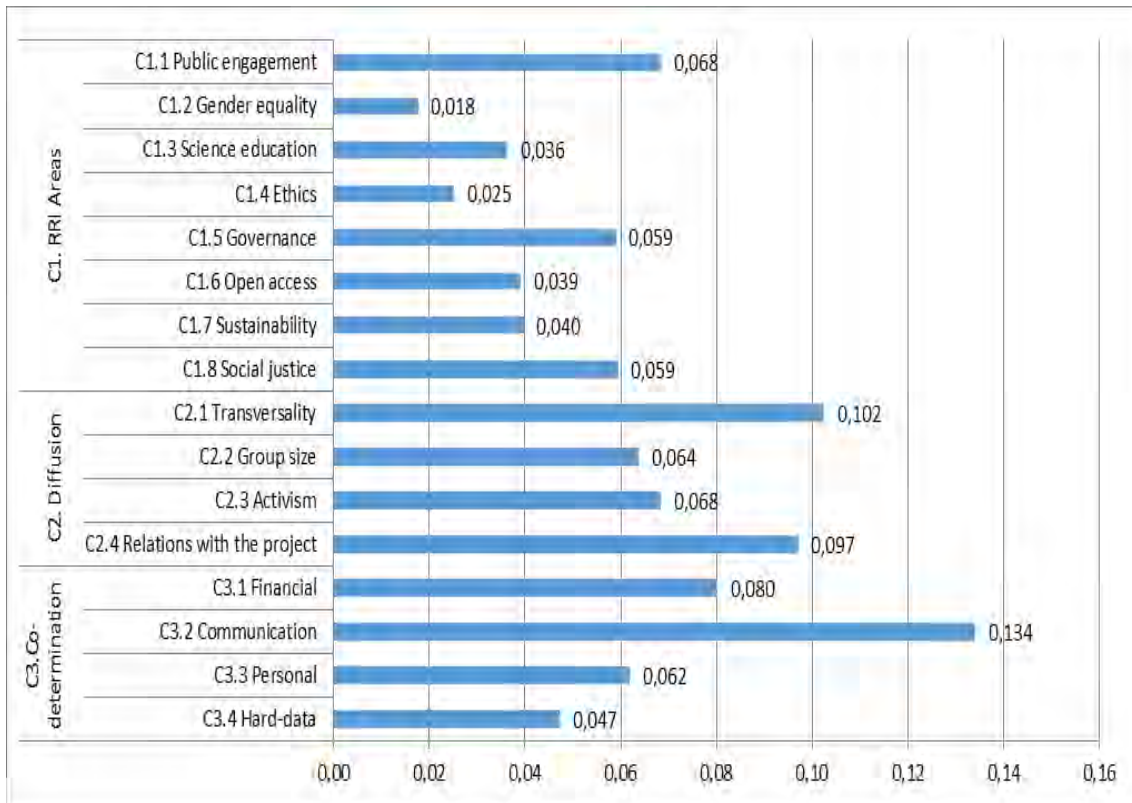
Figure 3: Results for the criteria.

An index for each stakeholders with regard to all considered criteria have been obtained. We called it the Preference Index, so the higher the index value, the more influential the stakeholder is. According to the result the most influential stakeholders is S2. Business, followed by S3. Local administration, S1. Users and S4. Developers.
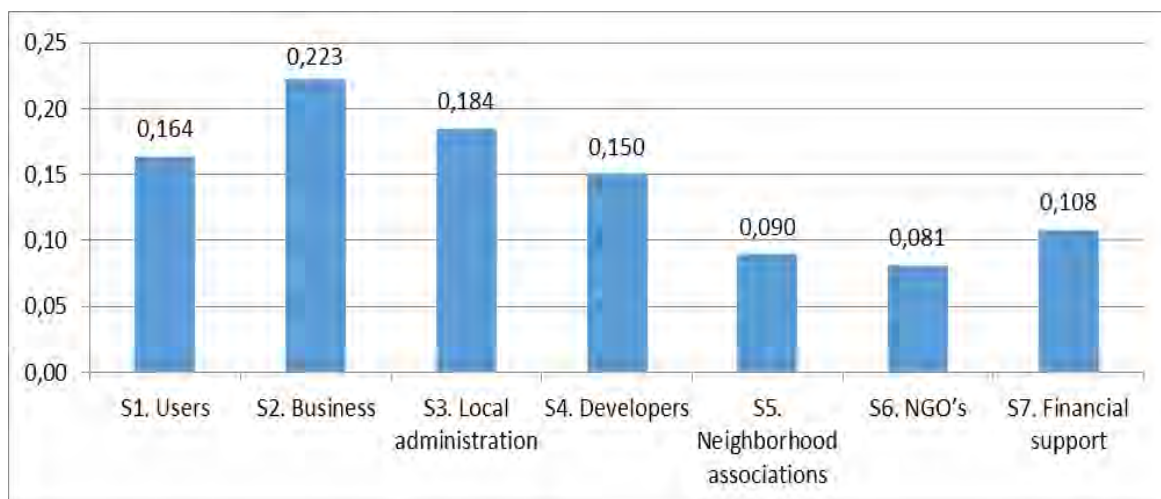
Figure 4: Results for the stakeholders

CONCLUSIONS

In this paper we have provided a novel applications of a MCDA technique to evaluate the stakeholder's influences within a project from the point of view of the responsibility as a key

aspect for its success. The novelty of our model is how the concept of influence is broken down into sixteen criteria, evaluating different aspects that together define an index which measures the influence of stakeholders in term of responsibility in the framework of RRI.

The results of the research herein presented lead to the conclusion that the ANP method, is useful to determine a rank of stakeholder in a research project. Besides, it can be adopted and applied to other types of influence assessment.

Finally, as can be seen in figures 3 and 4, the most influential stakeholder of the project evaluated is "Business", this has sense because this stakeholder generates the leisure or entertainment offer, therefore has the greatest impact.

Acknowledgement

References

[1]   Aragonés-Beltrán, P. et al. 2017. How to assess stakeholders' influence in project management? A proposal based on the Analytic Network Process. *International Journal of Project Management*. (2017). DOI:https://doi.org/10.1016/j.ijproman.2017.01.001.

[2]   Brugha, R. and Varvasovszky, Z. 2000. Stakeholder analysis: a review. *Health Policy and Planning*. 15, 3 (2000), 239–246. DOI:https://doi.org/10.1093/heapol/15.3.239.

[3]   Dahlsrud, A. 2006. How Corporate Social Responsibility is De ned: an Analysis of 37 De nitions. *Corporate Social Responsibility and Environmental Management*. 13, November 2006 (2006), 1–13. DOI:https://doi.org/10.1002/csr.

[4]   European Commission 2011. *DG Research workshop on Responsible Research & Innovation in Europe*.

[5]   Geoghegan-Quinn, M. 2012. *Responsible Research and Innovation. Europe's ability to respond to societal challenges*.

[6]   Mitchell, R.K. et al. 2009. Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really. 22, 4 (2009), 853–886. DOI:https://doi.org/10.5465/AMR.1997.9711022105.

[7]   Prell, C. et al. 2009. Stakeholder Analysis and Social Network Analysis in Natural Resource Management. *Society and Natural Resources*. 22, 6 (2009), 501–518. DOI:https://doi.org/10.1080/08941920802199202.

[8]   Saaty, T.L. 2001. *The Analytic Network Process: Decision Making with Dependence and Feedback*. RWS Publications.

[9]   Von Schomberg, R. 2011. Prospects for Technology Assessment in a framework of responsible research and innovation. *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden*. (2011), 39–61. DOI:https://doi.org/10.1007/978-3-531-93468-6_2.

[10] Stilgoe, J. et al. 2013. Developing a framework for responsible innovation ℧. *Research Policy*. 42, 9 (2013), 1568–1580. DOI:https://doi.org/10.1016/j.respol.2013.05.008.

[11] Strand, R. et al. 2015. *Indicators for promoting and monitoring Responsible Research and Innovation Report from the Expert Group on Policy Indicators*.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 5:
Metaheuristic Optimization

Fast Multi Descent
For Earliness Tardiness Scheduling Problem

Hemmak Allaoua[1],

Department of Computer Science, University of M'sila, 28000, Algeria;

Email: hem_all@yahoo.fr

Abstract

This paper present a new approach to tackle the single machine scheduling of independent jobs where the objective consists to minimize the sum of weighted earliness and tardiness against common due date which proofed as NP-hard problem. Such problems aim to provide a service or a product Just In Time (JIT), earliness and tardiness are penalized and find wide applications in various fields. Indeed, it was as several works subject of combinatorial optimization field in the recent period. The proposed approach called FMD (Fast Multi Descent) consists, firstly, on an experimented new heuristics based on potential problem properties to compute initial solution that is the starting point of the descent iterative process, then, a new variant of descent approach where multiple studied neighbor functions are applied alternatively and randomly in each iteration. The idea is inspired from a climber who has several tools to go down a hill. We show that the provided initial solution quality has a significant impact on the final obtained solution that need just some adjustments to realize significant efficiency. An empirical study is conducted to achieve a good compromise between computing time and quality solutions. Detailed study and comparison with other best results of recently related work justify that our approach give promising results.

Keywords: Combinatorial Optimization; Earliness Tardiness Scheduling; FMD; Hill Climbing Method; Scheduling Problems,

1 Introduction

In recent years, earliness tardiness scheduling problems have been attracting much interest in both academic and industrial communities. That is because Just In Time (JIT) concept used to provide a product or a service to a customer knows more and more applications in many production companies. The topic of this work is tracking single machine early tardy scheduling against common due date by hybridizing a specific heuristic with a variant of hill climbing algorithm which is called FMD (Fast Multi Descent). That consists, firstly, on an experimented new heuristic based on potential problem properties to compute initial solution, that is the starting point of the hill climbing iterative process, then, a new variant of hill climbing approach where multiple studied neighbor functions are applied alternatively and randomly in each iteration. The idea is inspired from a climber who has several tools to go down a hill.

Thus, a new high-level relay hybrid approach is proposed to solve the single machine scheduling problem of independent jobs where the objective is to minimize the sum of earliness and tardiness penalties of jobs having a common due date. The theoretical importance of such problem is due to its complexity as NP-hard problem, [1] [2] [5] [6] and due to its practical importance in just-in-time JIT modern production and service processing. In JIT, policy makers emphasize that a job should be completed as close as possible to its common due date to avoid inventory cost and loss of customer's goodwill. This concept is actually adopted by many economic and industrial companies around the world. The goal is on one hand to realize better results than existing works in the field and on the other hand to give a new idea of methods combination that could be generalized to other situations in the field.

In scheduling on a single machine against a common due date, there are two classes of common due date problems, which have proven to be NP-hard, namely the restrictive and non-restrictive common due date problem. Since, one job at most can be completed exactly at the due date hence; some of the jobs might be completed earlier than the common due-date (d), while other jobs finish late. In this work, the treated problem is such as the restrictive case of the problem where the common due date is less than the sum of the processing times of all the jobs and each job possess different earliness/tardiness penalties. This problem has proven to be the most difficult problem in this area of research.

Reviewing the literature of this kind of problem, it was found that three different properties were embedded into various constructive heuristics to obtain good approximation results, [1, 2, 6,7,21]. However, to capitalize on the potential of such properties, two new additional propensities are proposed with supporting proofs. The set of all mathematical properties are integrated inside a our approach called FMD. The quality of FMD is demonstrated on

---

[1] E-mail: hem_all@yahoo.fr.

a set of 280 benchmark instances available in [6, 7] and compared to the state-of-art algorithms for the problem in [1,6,22,23,24].

In other hand, currently in combinatorial optimization field, hybridizing methods is imposed to fill their gaps and take advantage of their strengths. It is in this context that inserted the present work of combining a heuristic based on the potential problem properties with a variant of the hill climbing method. In terms of hybridization hierarchical clustering, it is a high-level relay hybridization. Three heuristics are proposed and compared for small problem sizes to clear which one achieves the minimum gap to the optimum in order to accelerate our approach and improve its efficiency. In the second part, three neighbor functions are proposed and studied to make faster the hill climbing. The idea is inspired from a climber who has several tools to go down a hill at each crossroad of paths met.

The remaining parts of the paper are organized as follows: Section 2 provides background on the problem including a statement, literature review, and mathematical properties. In Section 3, the implementation details of our proposed FMD approach are presented. Computational results and discussions are provided in Section 4. Finally, we conclude with a summary of findings and perspectives on future work.

2 Problem presentation

The single machine scheduling problem with Early/Tardy jobs around a common due date involves a set of jobs; each job has its own processing time requirements. All jobs must be processed on a single machine and a penalty cost is incurred when a job is completed before (earliness) or after (tardiness) the common due date. The objective is to minimize the summation of earliness and tardiness penalty costs in order to encourage the completion time of each job to be as close as possible to the common due date. The following notations and statements are proposed to understand better the properties of problem. The following notations define the problem statement:

n : integer number of jobs to be scheduled ;
I : set of n jobs: $I = \{1, 2, \ldots, n\}$;
d : common due date of all the n jobs ;
C_i : complete time of job i ;
p_i : processing time of job i ;
$E_i = \text{Max} \{d - C_i, 0\}$ (Earliness of job i);
$T_i = \text{Max} \{C_i - d, 0\}$ (Tardiness of job i);
α_i : penalty per unit time of earliness for job i ;
β_i : penalty per unit time of tardiness for job i ;
h : parameter of common due date tightness used as follows: d = h * T , where: $T = \sum_{k=1}^{n} p_i$; h ∈ { 0.2 , 0.4,
    0.6, 0.8 }.
B : the set of jobs to complete at or before the common due date d ;
A : the set of jobs to complete after d ;

$$|A| + |B| = \begin{cases} n - 1 & \text{if there are a straddled job} \\ n & \text{else} \end{cases} ;$$

A straddled job is a job that starts before d and completes after d. If, in a given sequence, there is a job that completes exactly at d, there will be no straddled job.

- Each job has to be processed on the single machine without interruption;
- Each job is available at time 0;
- Each job must be processed just once;
- For each job i, the processing time p_i, the cost per unit time of earliness α_i the cost per unit time of tardiness β_i are given and assumed integer. The objective function to be minimized can be expressed as the sum of weighted penalties of earliness and tardiness as follows: $\sum_{k=1}^{n} (\alpha_i E_i + \beta_i T_i)$. Any permutation of n jobs will be a feasible solution to the problem and there are an exponential number of such permutations (O(n!)). The optimal sequence is the permutation, which has the minimum objective value among all permutations.

The optimal solution for the single machine problem satisfies the following three optimality properties:

Property 1: It does not contain any idle time between any consecutive jobs.
Property 2: It is V-shaped around the common due date: the jobs completing before or on the common due date are sorted in decreasing order of the ratios p_i/α_i, and the jobs starting on or after the common due date are sorted in increasing order of the ratios p_i/β_i.
Property 3: In an optimal schedule, either the first job starts at time zero or the completion time of one job coincides with the common due date.

The proofs of these properties are established using proof by contradiction in: Kanet [18], Lee and Kim [17], Gordon et al. [13], Feldmann and Biskup [7], Lin et al. [12], Biskup (1999) [19], and Hall and Posner (1991) [18].

The most difficult part of the treated problem usually focuses on how to find the beginning time of optimal sequence, because that will multiply the time complexity of the problem by d-1 (where d is the common due date for all jobs). That means that the beginning time t_0 belongs to [0,d-1].
The three problem properties presented above describe just the optimal sequence without specifying any details about the beginning time of the optimal sequence. That is happening often when h = 0.6 or h=0.8 (in these cases the common due date d is equal to 0.6*sum_pi or 0.8*sum_pi, so, it is quite possible that how scheduling starts t_0 gretaer than 0). Therefore, in this work, two new properties were developed then proofed for determining the beginning time t_0 of the optimal sequence. In one hand, that will improve considerably the solution quality, and, in the other hand, that will reduce clearly the time computing of the algorithm. These properties are called respectively 4 and 5 [26].

Property 4: There is an optimal schedule in which the arithmetic mean $[m = \dfrac{(\sum\limits_{i \in B} \alpha_i C_i + \sum\limits_{i \in A} \beta_i C_i)}{(\sum\limits_{i \in B} \alpha_i + \sum\limits_{i \in A} \beta_i)}]$ is equal to

the common due date d.

Corollary:
a) If d ≤ m the optimal schedule may start at time $t_0 = 0$.
b) If d > m then the optimal schedule may start at time $t_0 = d - m$.

Property 5. If m < d, there is an optimal schedule in which $\sum\limits_{i \in B} \alpha_i = \sum\limits_{i \in A} \beta_i$.

Corollary:

To reduce the size of the space of solutions, it may just seek solutions having $\sum\limits_{i \in B} \alpha_i \approx \sum\limits_{i \in A} \beta_i$.

Example
Size n = 10 jobs ; Instance k = 1 ; (source: Biskup and Feldman benchmarks) [6].

Job i	1	2	3	4	5	6	7	8	9	10
p_i	20	6	13	13	12	12	12	3	12	13
α_i	4	1	5	2	7	9	5	6	6	10
β_i	5	15	13	13	6	8	15	1	8	1

Sum $p_i = T = \sum_{k=1}^{n} p_i = 116$; d = T*h ; h ∈ { 0.2 , 0.4, 0.6, 0.8 }.
Results:

h	d	t_0	Opt.	Optimal sequence
0.2	23	0	1936	4 2 7 3 9 6 5 8 1 10
			Time =	0 d (straddled job 7)
0.4	46	0	1025	4 2 3 7 9 6 5 8 1 10
			Time =	0 d (straddled job 9)
0.6	69	1	841	4 2 3 7 9 6 5 8 1 10
			Time =	1 d (no straddled job)
0.8	92	16	818	4 2 1 3 7 6 9 5 8 10
			Time =	16 d (no straddled job)

Straddled job means a job that start just before d and complete after d.
It is easy to verify that:
1- There are no idle time in optimal schedule. That is because $C_{i+1}=C_i+p_{i+1}$. (Property 1).
2- Jobs in B are sorted by decreasing order of ratio pi/ αi and jobs in set A are sorted by increasing order of ratio pi/ βi. (Property 2) (V- shaped property) . (B set of jobs completing at or before d and A set of jobs completing after d.

3- In optimal job, there is a job that start at 0 or a job completes at d. (Property 3).
4- If $d \leq$ mean then optimal time must start at 0 (as in h = 0.2 and h = 0.4 above). Else optimal schedule may start at $t_0 = d -$ mean. (as in h = 0.6 and h = 0.8 above) Note. It is not always $t_0 = 0$ for h=0.2 , 0.4 and t0 > 0 for h = 0.6 , 0.8 . But it is depending of the instance data. (Property 4).
5- In each optimal sequence above, we have $\sum_{i \in B} \alpha_i \approx \sum_{i \in A} \beta_i$. (Property 5).

3 FMD implementation

FMD approach consists a variant of descent method with several initial solutions and several neighbor functions that is to improve quality solution and the descent faster. Three heuristics for computing initial solution, the starting point of the descent method, then three neighbor functions are used to choice the current solution neighbor.

3.1 Heuristics for initial solution

In this section, the FMD approach will be presented, that is used to solve the single machine early tardy scheduling problem. The five properties described above are integrated as it will be clarified later.
The FMD algorithm process starts by initial solution using heuristics based on potential problem properties. The above properties imply that the treated problem can modeled as partition problem of the set I of jobs on the two sub sets A and B described in the presentation section. That is because property 2 will allow to sort jobs in the appropriate order.
Three heuristics are tried FFH (Fitness Function Heuristic), NNH (Nearest Neigbor Heuristic), GPH (Greedy Partition Heuristic).
FFH is based on defined fitness function *Fitness* : I $\rightarrow$ R that evaluate jobs to belong to sub set B. it is clear that jobs in set B might have small α_i and great β_i inversely for jobs of sub set A (figure1).

Figure 1. Scheduling jobs according greedy algorithms.

So the fitness function for each job i to belong to sub set B is defined as follow:
$$fitness(i) = (mean\beta \times \alpha_i - mean\alpha \times \beta_i)/p_i^{\lambda}.$$
where meanα and meanβ are respectively the mean of costs α_i and β_i, that is for taking in consideration the instance input and λ is defined as follow:
if h < 0.5 then λ=1 else λ=-1, that means :
if h=0.2 or h=0.4 then λ=1 ;
if h=0.6 or h=0.8 then λ=-1 ;
thus : $\lambda = |h-0.5|/(h-0.5)$.

So, the FFH algorithm is described as follow:
Step1. Sort all jobs according increase order of fitness function;
Step2. Apply property 2 to sort jobs of A and B;
Step3. Apply property 4 and 5 to compute the stating time;
Step 3. Compute best solution and its objective.

NNH is inspired from nearest neighbor greedy algorithm for TSP where a distance between two jobs i and j is defined as follow: dist(i,j) = $\alpha_i \times$ max(d-C_i ;0) + $\beta_i \times$ max(C_i-d ;0) ; that is the cost to be added if j is sequenced immediately after i. Since, the nearest jobs will belong to the same sub set. The NNH algorithm is described as follow:
Step1. Find the job with maximum fineness (that is the first job of B);
Step 2. Sort jobs according increase order of distances ;
Step3. Apply property 2 to sort jobs of A and B;
Step4. Apply property 4 and 5 to compute the stating time;
Step 5. Compute best solution and its objective.

GPH is consists on greedy partition algorithm of set I on two sub sets A and B which might satisfy property 5 as close as possible $\sum_{i \in B}' \alpha_i \approx \sum_{i \in A}' \beta_i$. The algorithm is described as follow:

Step1. Let X be a list of all jobs sorted according increase order of α_i ;
Step2. Partition X to two subsets Ax and Bx ;
Step3. Let Y be a list of all jobs sorted according increase order of β_i ;
Step4. Partition Y to two subsets Ay and By;
Step5. Make Ax∩Ay in A and Bx∩By in B goto step 1 for the remaining jobs ;
Step6. Apply property 2 to sort jobs of A and B;
Step7. Apply property 4 and 5 to compute the stating time;
Step8. Compute best solution and its objective.

The set of benchmark which was proposed and designed by Biskup and Feldman [6], [7] and was employed to demonstrate the efficiency and effectiveness of the proposed FMD approach. The benchmark set consists of instances with variable sizes, n=10, 20, 50, 100, 200. 500, 1000 jobs.
Obtained results are compared to the optimum for small instances sizes using deviation ratio R defined as follow:
R= (H-Opt)/Opt where H is the solution obtained by the heuristics above, and Opt is the optimum computed by exact method (as dynamic programming) for instances with small sizes (n=10 and n=20, that are a total of 80 instances).
Tables and curves for comparing results are reported in appendix.

3.2 Descent approach

The descent will use both the heuristics above, then, for a small number of iterations, will apply alternatively and randomly one of three neighbor functions :
SWAP: that consists to swap a job of B with a job of A ;
Right Insertion: that consists to insert a job of A in B and shift all jobs to the right ;
Left Insertion; that consists to insert a job of B in A and shift all jobs to the left.
The FMD process is generally described as follow :
Step1. Compute initial solution with each heristic FFH, NNH , GPH ;
Step2. For each initial solution:
    Repeat
        Choose randomly a neighbor function (SWAP , RI , LI)
        Compute all neighbors of the current solution ;
        Update the current solution ;
        Replace the current solution ;
    Until no improvement ;

4   Computational Results
The produced results are compared to those obtained by Nearchou [1]. It is found that Nearchou's results are the best results in the literature in most cases and it was selected for comparison with the proposed algorithm. For each size of n = 10 to n=1000. Ten different instances with different values of rate h=0.2, 0.4, 0.6, 0.8 to determine the common due date, i.e., there were 80 benchmarks for each size.

Table 1 summarizes the comparative results, starting from the problem's characteristics in column one followed by our CPU time, associated MEGA's average solutions over the ten instances per size, the Nearchou's average solutions, the average of the differences per instances, and to the comments on the number of new Best solutions obtained by the proposed approach (B), the number of Equal solutions (E) and the number of our Worst solutions. The results were obtained on PC machine with Intel I5, 2.4 GHz CPU and 4 GO Ram.

Tables and curves for comparing results are reported in appendix.

5 References

[1] Andreas C. Nearchou, "A differential evolution approach for the common due date early/tardy job scheduling problem", Computers & Operations Research 35 (2008) 1329 – 1343
[2] K. R. Baker, G. D. Scudder, "Scheduling with earliness and tardiness penalties: a review", European Journal of Operational Research, 160 (2005) 190-201.

[3] Ching-Jong Liao, Che-Ching Cheng, "A variable neighborhood search for minimizing single machine weighted earliness and tardiness with common due date", Computers & Industrial Engineering 52 (2007) 404–413.

[4] Chun Nam Cha, Sanggyu Lim and Yong Kyeong Jeong, "Single-Machine Job Scheduling about a Common Due Date with Arbitrary Earliness / Tardiness Penalties Using a Genetic Algorithm", Asia Pacific Management Review (2002) 7(2), 239-254.

[5] Débora P. Ronconi and Márcio S. Kawamura, "The single machine earliness and tardiness scheduling problem: lower bounds and a branch-and-bound algorithm", computational and applied mathematics, Volume 29, N. 2, pp. 107–124, 2010.

[6] M. Feldman, and D. Biskup, "Benchmarks for scheduling on a single machine against restrictive and unrestrictive common due dates", Computer & Industrial Engineering, 28 (2001) 787-801.

[7] M. Feldman, and D. Biskup, "Single-machine scheduling for minimizing earliness and tardiness penalties by meat-heuristic approaches", Computer & Industrial Engineering, 44 (2003) 307-323.

[8] R. Hassin, and M. Shani, "Machine scheduling with earliness and tardiness and non-execution penalties", Computer & Industrial Engineering, 32 (2005) 683-705.

[9] C. M. Hino, D. P. Ronconi, A. B. Mendes, "Minimizing earliness and tardiness penalties in a single machine problem with a common due date", European Journal of Operational Research, 160 (2005) 190-201.

[10] Jorg Lassig et al., "Common Due-Date Problem: "Linear Algorithm for a Given Job Sequence", 2014 IEEE 17th International Conference on Computational Science and Engineering.

[11] S.W. Lin, S. Y. Chou and K. C. Ying, "A sequential exchange approach for minimizing earliness-tardiness penalties of single machine scheduling with a common due date", European Journal of Operational Research, 177 (2007) 1294–1301.

[12] T. Vallée, and M. Yiltizogli, "Présentation des algorithmes génétiques et leurs applications en économie", Mai 2004, V .5.

[13] Valery Gordon, Jean-Marie Proth, Chengbin Chu, "Invited Review A survey of the state-of-the-art of common due date assignment and scheduling research", European Journal of Operational Research 139 (2002) 1–25.

[14] Ramon Alvarez-Valdes, Enric Crespo, Jose Manuel Tamarit, and Fulgencia Villa "Minimizing weighted earliness–tardiness on a single machine with a common due date using quadratic Models". Top (2012) 20:754–767.

[15] Kuo-Ching Ying, "Minimizing earliness–tardiness penalties for common due date single-machine scheduling problems by a recovering beam search algorithm". Computers & Industrial Engineering 55 (2008) 494–502.

[16] D. Biskup, "Single-machine scheduling with learning considerations," European Journal of Operational Research, vol. 115, no. 1, pp. 173–178, 1999.

[17] Ahmad Jafarnejad, Seyed Mehdi Abtahi, Sayyed Mohammad Reza Davoodi, Optimizing the Earliness and Tardiness Penalties in the Single-machine Scheduling Problems with Focus on the Just in Time. International Journal of Academic Research in Business and Social Sciences July 2013, Vol. 3, No. 7.

[18] K. R. Baker, G. D. Scudder, "Minimizing Earliness and Tardiness Costs in Stochastic Scheduling", European Journal of Operational Research, (2013) 236(2):445–452.

[19] Mohammadreza Shahriari, Naghi Shoja, Amir Ebrahimi Zade, Sasan Barak, Mani Sharifi. JIT single machine scheduling problem with periodic preventive maintenance. Journal of Industrial Engineering International. (2016) pp 1-12.

[20] Abhishek Awasthi, Jorg Lassig and Oliver Kramer. Common Due-Date Problem: Exact Polynomial Algorithms for a Given Job Sequence. arXiv (2013).

[21] Allaoua Hemmak, Ibrahim H. Osman, Variable Parameters Lengths Genetic Algorithm for Minimizing Earliness-Tardiness Penalties of Single Machine Scheduling With a Common Due Date, Electronic Notes in Discrete Mathematics. 36 (2010) 471–478.

[22] Allaoua Hemmak, Brahim Bouderah, Hybrid Algorithm for Optimization Problems Applied to Single Machine Scheduling, International Journal of Computer Applications. Volume 66– No.24, March 2013.

[23] Allaoua Hemmak, Brahim Bouderah, A mono crossover genetic algorithm for TSP, Global Journal on Technology, Issue 7 (2015) 109-115.

[24] Allaoua Hemmak, Brahim Bouderah, Sieve Algorithm - A New Method for Optimization Problems, Int. J. Advance. Soft Comput. Appl., Vol. 5, No. 2, July 2013.

[25] Zheng Ning, Chen Tao, Lin Fei. A Hybrid Heuristic Algorithm for the Intelligent Transportation Scheduling Problem of the BRT System, Journal of Intelligent Systems. Volume 24, Issue 4 (Dec 2015), 437–448.

STUDY OF INITIAL GUESS INFLUENCE ON THE QUALITY OF SOLUTIONS ON BINARY GENETIC ALGORITHM IN JOB SHOP SCHEDULING PROBLEM

Valmir Ferreira da Cruz
Universidade Nove de Julho/Industrial Engineering Graduate Program
Francisco Matarazzo, Av. 612, São Paulo, Brazil
E-mail: valmir.vfc@gmail.com

Fabio Henrique Pereira
Universidade Nove de Julho/Informatics and knowledge management Graduate Program
Francisco Matarazzo, Av. 612, São Paulo, Brazil
E-mail: fabiohp@uni9.pro.br

Abstract: The purpose of this study is to evaluate the influence of the initial guess to generate the Genetic Algorithm population of solutions of scheduling problems in relation to the quality and feasibility of the solutions. The scheduling problem is defined as to find the sequence of operations on the machines that optimize some performance measure as, for example, the use of resources and the total processing time (makespan). It is common to treat such problems with the use of metaheuristics as genetic algorithm mainly due to its computational complexity. This work carried out experiments with a set of literature instances, varying the sequencing rule used in the generation of initial solutions. Usually, rules from the literature have been tested and identified a hybrid rule that generates a smaller number of non-feasible solutions and the number of instances that have reached the optimal makespan.

Keywords: Scheduling, Job Shop, Genetic Algorithm, Initial Guess

1 INTRODUCTION

Among the main production engineering problems, the production scheduling problem has been studied, nowadays, by a large number of researchers (KURDI, 2015; ASADZADEH, 2015; AMIRGHASEMI & ZAMANI, 2015; KUNNATHUR *et al.*, 2004; HEINONEN; PETTERSSON, 2007). This problem consists of finding the best operations scheduling to be done in the production line, seeking to optimize the time on each machine, in such a way to reduce to the most the idle time as well as the best position of the various machines in the production, for example. According to BO PENG *et. al* (2014), the job-shop scheduling problem (JSSP) is, apart from being a notable and hard to tackle problem, one of the scheduling problem most important, which come up in situations, where the set of activities following irregular flow patterns, should be carried out by a set of scarce resources.

Production scheduling when has a large number of machines and/or Jobs, demands much processing time to evaluate all of the possible solutions, ending up where there are cases that make impossible to draw close to an optimal solution in acceptable computational time. Such a problem is classified as NP-Hard, which is not resolved in acceptable polynomial time.

Problems with such a degree of complexity allow for a larger number of possible combinations when the crossing between jobs and machines. According to Lukaszewicz (2005), the solutions space for a JSSP is made up of $(n!)^m$ possible sequencing, where n, is the number of Jobs and, m, the number of machines. Therefore, in a scenario with 10 jobs and 15 machines, there is $2.49 * 10^{98}$, possible combinations. However, many of these combinations are taken as infeasible, that is, not possible of being run in practice.

To evaluate the viability and the makespan of each one of these combinations looking for the best solution may require an impractical computational time, even not allowing for the conclusion of the task due to computational scarce resources. Therefore, overall, techniques

set forth for the evaluation of a limit number of such possible combinations are used purposing to select the best solutions among the evaluated only. These techniques, which propose to find the sub-optimal solution among an acceptable computational time, are called metaheuristics.

Studies, with metaheuristics application in scheduling problems, are discussed next subsection. Further, in the article, it is presented a production environment classification highlighting the job-shop environment, genetic algorithm general aspects, material and methods, results and conclusions.

1.1 Metaheuristics application on scheduling problems

Many researches, which address JSSP with metaheuristics utilization, have been published all over the years. Gao et al. (2015), used the Ants' Colony metaheuristics to handle the engineering re-manufacturing problem; Saidi-Mehrabad et al. (2015), used the same mataheuristics, to deal with the transport time problem, among the machines. As recent works using the genetic algorithm metaheuristics on the handling of JSSP, we can name: Asadzadeh (2015); Amirghasemi & Zamani (2015); Qing-dao-er-ji et al. (2013) and Kurdi (2015). These Works illustrate the scientific community interest on the metaheuristics use. Within this context, several approaches have been developed varying diverse parameters and GA operators, among them the solution representation (ABDELMAGUID, 2010; GRASSI et al., 2016).

Abdelmaguid (2010), assessed the different representations based on real numbers and on a list of integer numbers. The author points out that the integer numbers lists, which represents in a direct way a solution for JSSP, are the ones more employed, but may present a high number of infeasible solutions. On the other hand, there are representations which yield feasible solutions solely; though they hold deficiencies with regards to generation of new solutions different from the preceding ones (MODOLO et al., 2015).

The fact is that the different representations yield different results in GA scheduling problems, interfering with feasible solutions output. In this sense, Grassi et al. (2016) proposed a Genetic Algorithm with binary representation using a dynamic seed concept (DSGA) proposing a reduction of the infeasible solutions numbers. On the work the author shows that the proposed representation cuts down the infeasible solutions proportion in comparison with the traditional GA elitism, since the initial seed, represents a feasible solution. In order to guarantee the solutions feasibility, the author generates the initial seed out of FIFO (First In First Out) rule, but does not investigate what are the effects of generating that initial guess based on other rules, especially those which, contrary to FIFO, consider the scheduling job processing time. Thus, the present article purposes to study the GA initial guess effect in the binary representation regarding the scheduling solutions quality relating to the infeasible solutions proportions, processing time, and the obtained makespan value. The hybrid scheduling jobs, on top of the input order of such Jobs, are tested: FIFO + SPT and FIFO + LPT (FIFO + Shortest Processing Time e FIFO + Longest Processing Time, respectively).

2 PRODUCTION SCENARIO AND SCHEDULING PROBLEM

Lukaszewicz (2005) defined the production scheduling as an attribution process of one or more resources for the execution of certain activities which, in their execution, will require a given amount of time. In industrial environment, the machines represent the resources and, the Jobs, are the activities that are processed in each machine. So, a job is a set of one or more tasks. Regarding the way, in which the Jobs are distributed for the different machines, and, whether these Jobs are different between themselves, the production scenarios are given different classifications (ALLAHVERDI et al., 2008).

In a job shop environments each order is unique with pre-determined routes, which are unlike one from another, being the object of this study. Job-shop scheduling tasks have the following characteristics:

- Unique order with pre-determined route;
- Each order is processed, at least, once on each machine.

The problems within this environment are known as Job-shop scheduling problems (JSSP) and due to its computational complexity, metaheuristics techniques are used such as, for instance, the genetic algorithm, presented in the next section.

3. GENETIC ALGORITHM

The Genetic Algorithm (GA), introduced by Holland (1975), is a technique based on the species evolution, proposed by Charles Darwin, and has been broadly applied to solve NP-Hard class problems. Especially, the utilization of this algorithm as an optimization technique, for JSSP scheduling environment, is largely known (ASADZADEH, 2015; AMIRGHASEMI & ZAMANI, 2015; QING-DAO-ER-JI et al., 2013; KURDI, 2015).

GA diverges from other heuristics methods, for its distinctive characteristics: 1) It deals and operates with a set of known points, called population, rather than isolated points; and 2) It operates in a codified solutions space, rather than straight into the search space with the needed data, using the value of an objective function, called aptitude; It uses probabilistic transition rules instead of deterministic (GOLDBERG, 1989).

The technique, applied by a GA, consists in right from an initial population, to calculate the fitness value of each point, called individual or chromosome. From the crossover and mutation genetic operators, GA creates new generations inserting the chromosomes into the current population, thereby promoting random changes to the end of accessing a smaller search space and, at the same time, preventing from being restricted to minimum or maximum sites. This passage, determines the mechanism, which will cover two or more existing chromosomes, to create two or more offspring.

4. MATERIAL AND METHODS

This study is based on the Dynamic Seed Genetic Algorithm (DSGA) presented by (GRASSI et al., 2016). DSGA uses the GA classical, at an internal level, where the prospective solutions are generated by permuting a feasible initial solution, called seed. In the original work, the authors used as a way to guarantee the initial solution viability, a seed created based upon the dispatch rule FIFO (first in, first out), and permutation relying on the binary chromosome. The DSGA approach is still made up of an external level, in which the best solution found at the internal level, after a given number of GA offspring, is used to update the seed used for the iterations (GRASSI et al., 2016).

The experiments carried out in this study, purposes to evaluate the effects caused on the results of: Makespan, Gap between the yielded makespan and the best one already achieved in literature, Processing Time, Not feasible solutions ratio.

From a set of JSSP samples, known as LA (Lawrence, 1984), initial seeds were generated for the LA01 to LA10 problems, with the following characteristics:

Non feasible seed, FIFO (First-In, First-Out), FIFO + SPT (Shortest Processing Time), FIFO + LPT (Largest Processing Time). The implementation was accomplished using the GA library GAlib (1996).

The parameters used by the GA execute the experiments are shown in Table 1. The objective function runs the evaluation of the yielded solutions using a modified Dijkstra algorithm to calculate the critical path (SHANKAR e SIREESHA, 2010).

Table 1: GA parameters used on the experiments

Parameter	Value
Representation of solution	Binary 2D (DSGA)
Selection	Rolete
Replacement of solutions	*Steady State*
Replacement rate	90%
Population size	10
Crossover	*OnePointCrossover (OP)*
Crossover rate	90%
Mutation	Inversion of bit
Mutation rate	1%
Number of Generations	25 x 200
Stop criterion	Number of generations
Fitness function	Long path based on a modified Dijkstra algorithm (SHANKAR and SIREESHA, 2010).

Twenty-four experiments were made of each one of the LA01 to LA10 samples, where each sample was run individually and, at the end, the averages of each factor studied were evaluated as well as the overall means between all instances.

5. RESULTS

The obtained results were compared between themselves with the purpose to identify the initial guess effects on the solution quality. Table 2 results present for each tested problem and each initial seed (*Seed*) generating approach, the *gap* (difference) between the best solution found by the algorithm and the best result found in literature (best). It was observed that in all experiments started with infeasible initial seed, the algorithm did not manage to converge to a feasible solution, which is signaled with * symbol.

In the LA01 sample is possible to notice that, from a not feasible seed, the algorithm did not converge to the global optimal and 100% of its solutions were not feasible as shown in Table 3. This behavior is repeated in all of the other problems.

Table 2: Numerical results for tested problems.

Problem	Infeasible	FIFO	FIFO+SPT	FIFO+LPT
LA01	*	2	2	4
LA02	*	0	0	0
LA03	*	14	14	14
LA04	*	17	11	10
LA05	*	0	0	0
LA06	*	0	0	0
LA07	*	0	0	0
LA08	*	0	0	0
LA09	*	0	0	0
LA10	*	0	0	0

On the other hand, when the process starts with a feasible solution in LA01, only 9% of the solutions were not feasible and did not yield *gap* in relation to the best solution found in literature. The method also reached the optimal (*gap* = 0) for the LA02 e LA05-LA10 problems. It's worth noticing, however, that the optimal value is not achieved in all of the executions. In the LA02 sample, for instance, the initial solution with dispatch rule FIFO+LPT, outdid the others, yielding a larger number of feasible solutions, reaching the optimal solution as twice as the times with the same processing time.

In the LA03 and LA04 samples, the method did not find the optimal solution in all of the cases. For the LA03 the experiments which departed from a feasible solution, obtained results very similar, regardless of the dispatch rule. In the case of LA04, however, there was a difference to be pointed out, which is the smallest *gap* in the utilization of the FIFO+LPT dispatch rule. Thus, it is possible to assert that the utilization of a feasible initial seed based upon the FIFO+LPT hybrid rule, despite not finding the optimal solution, found the best suboptimal solution in an acceptable computational time. Finally, for the LA07 instance, despite the small difference in the optimal solutions quantity, that is, with *makespan* equal, found in literature, one can say it came to a technical tight, mainly if the mean *makespan* is taken into account, which showed irrelevant variation.

It is also important to point out that, starting the process with the utilization of a solution which follows the FIFO dispatch rule, the algorithm fetched a higher number of times the optimal solution with regards to the FIFO+SPT and FIFO+LPT rules with a processing time, if not equal, very close to. A summary of all the results is presented in Table 3.

Table 3: Mean of results and improvement in relation to the FIFO approach

Parameter	Not feasible	FIFO	FIFO + SPT	FIFO + LPT	Improvement
% not feasible	100%	7,2%	7,2%	7,1%	1,3%
Mean gap	*	3,1	2,5	2,4	22,6%
Mean mkp	*	779,6	779,4	779,0	0,1%
Mean of feasible solutions	*	27480,4	27497,2	27488,6	0,1%
Mean of not feasible	29025,7	1881,2	1878,4	1862,5	1,0%
Min mkp	*	772	771,4	771,3	0,1%
Mean time	05:41	10:00	06:40	06:37	33,7%

6. CONCLUSION

By analyzing the experiments' results, it is possible to verify that the initial guess suitable choice has positive influence on the quality of the obtained solutions. Besides the algorithm finding the optimal solutions in the majority of the experiments, in those where the initial seed was not feasible, the genetic algorithm did not manage to converge to a feasible solution. In this scenario, the problem of the non-convergence was duplicated in all the test samples yielding 100% of not feasible solutions. The experiments made with the feasible initial seed usage presented, within the genetic algorithm output solutions universe, a high ratio of feasible solutions, ranging from 91 to 96%.

The usage of a feasible initial solution reliant on the FIFO+LPT hybrid rule stood out of the others, showing performance gains in all of the analyzed parameters, when compared to the results of the smaller performance approach, with the greatest highlight for the spent mean time, on the experiments execution, and, for the overall *gap* decreasing at 33.7% and 22.6%, respectively. This, indicates that an initial guess, which uses all the problem information (job order arrival and processing time), may yield better results in comparison with the use of the FIFO rule.

References

[1] Abdelmaguid, T. F. 2010. Representations in genetic algorithm for the job shop scheduling problem: a computational study. Journal of Software Engineering and Applications, 3: 1155-1162.

[2] Allahverdi A.; Ng C. T.; Cheng T. C. E.; Kovalyov M. Y. 2008. A survey of scheduling problems with setup times or costs. European Journal of Operational Research, 187: 985-1032.

[3] Beasley J. Operations research library, 2005. http://people.brunel.ac.uk/~mastjjb/jeb/orlib/files/jobshop1.txt> [Accessed 17/12/2015].

[4] Peng, B. Lü, Z., Cheng, T.C.E. 2014. A tabu search/path relinking algorithm to solve the job shop scheduling problem. Computers & Operations Research, 53(2015): 154–164.

[5] GAlib: A C++ library of genetic algorithm components [online]. Available from: http://lancet.mit.edu/ga/dist/ [Accessed 27/10/2015].

[6] Goldberg, D. E. 1989. Genetic algorithms in search, optimization and machine learning. New York: Addison-Wesley.

[7] Graham, R. L.; Lawler, E. L.; Lenstra, J. K.; Kan, A. H. G. R. 1979. Optimization and approximation in deterministic sequencing and scheduling: a survey. Annals of Discrete Mathematics. 5: 287-326.

[8] Grassi, F. Schimit, P. H. T. and Pereira, F. H. 2016. Dynamic seed genetic algorithm to solve job shop scheduling problems. In IFIP International Conference on Advances in Production Management Systems, 170-177.

[9] Heinonen, J., Pettersson, F. 2007. Hybrid ant colony optimization and visibility studies applied to a job-shop scheduling problem. Applied Mathematics and Computation, 187(2): 989-998.

[10] Holland, J. H. 1992. Adaptation in Natural and Artificial Systems. 2. Ed. The MIT Press.

[11] Jain, A. S., Meeran, S. 1999. Deterministic job-shop scheduling: past, present and future. European Journal of Operational Research, 113(2): 390-434.

[12] Kunnathur, A. S., Sundararaghavan, P. S., e Sampath, S. 2004. Dynamic rescheduling using a simulation-based expert system. Journal of Manufacturing Technology Management, 15 (2): 199-212.

[13] Lawrence, S. 1984. Resource constrained project scheduling: an experimental investigation of heuristic scheduling techniques (Supplement). PhD diss., Carnegie-Mellon University.

[14] Asadzadeh, L. 2015. A local search genetic algorithm for the job shop scheduling problem with intelligent agents. Computers & Industrial Engineering, 85: 376–383.

[15] Lukaszewicz, P. P. 2005. Metaheuristics for job shop scheduling problem, comparison of effective methods. PhD diss., Aarhus School of Business.

[16] Amirghasemi, M. Zamani, R. 2015. An effective asexual genetic algorithm for solving the job shop scheduling problem. Computers & Industrial Engineering, 83: 123–138.

[17] Kurdi, M. 2016. An effective new island model genetic algorithm for job shop scheduling problem. Computers & Operations Research, 67: 132-142.

[18] Pinedo, M. L., 2008. Scheduling: Theory, algorithms, and systems. New York: Springer.

[19] Qing-dao-er-ji, R. Wang, Y. Wang, X. 2013. Inventory based two-objective job shop scheduling model and its hybrid genetic algorithm. Applied Soft Computing, 13: 1400–1406.

[20] Modolo, V.; Menezes, F. M.; Grassi, F.; Pereira, F. H. 2015. The Influence of the Crossover Operator on Genetic Algorithms Applied to the Job Shop Scheduling Problems. In: XXI International Conference on Industrial Engineering and Operations Management, Aveiro. Proceedings of ICIEOM. Rio de Janeiro: ABEPRO, 1:1-8.

[21] Wang S. F.; Zou Y. R. 2003. Techniques for the job shop scheduling problem: a survey. Systems Engineering - Theory & Practice, 23: 49-55.

AN ANN AND GA APPROACH FOR DEMAND FORECASTING AND ROUTING FOR CASH MANAGEMENT

Alev Taskin Gumus
Yildiz Technical University, Department of Industrial Engineering
34349 Besiktas, İstanbul, Turkey
E-mail: ataskin@yildiz.edu.tr

Erkan Celik
Munzur University, Department of Industrial Engineering
62000 Tunceli, Turkey
E-mail: erkancelik@munzur.edu.tr

Furkan Ömerustaoğlu
Yildiz Technical University, Department of Industrial Engineering
34349 Besiktas, İstanbul, Turkey
E-mail: furkanomerustaoglu@gmail.com

Abstract: Forecasting of cash demand and cash transportation are critical processes for all bank asynchronous transfer modes (ATMs). In this paper, an application is presented to accurate forecasting of cash demand and minimizing the transportation cost. In the first step of the proposed application, artificial neural network is implemented for forecasting the amount of withdrawn and deposited money from an ATM in one day. Hence, a more accurate way is provided to determine the amount of money that will be loaded into the ATM, and opportunity costs are minimized. Then, we applied genetic algorithm for minimizing the transportation cost of cash deposit to the ATMs.

Keywords: artificial neural networks, genetic algorithms, forecasting of cash, cash transportation and vehicle routing

1 INTRODUCTION

Although the number of asynchronous transfer modes (ATMs) offer easy access for customers; forecasting of demand for cash and cash transportation are critical processes for all bank ATMs. Cash demand in ATMs needs to be forecasted exactly like other products in vending machines [1,2]. If the forecasts are wrong, they induce costs. If the forecast is too high, unused cash is stored in the ATM incurring costs to the bank. If the forecasts are correct and the ATM is available, the customers can get service just in time and they don't need to go to other competitors. The rate of being advised of the bank increases; also the bank's brand awareness and value increases. For these reasons, forecasting is the first and maybe most important part of this process. Artificial neural networks (ANNs) are used for forecasting operations in this paper because the estimation error is lower than other statistical forecast methods.

Cash transportation vehicle routing and scheduling are serious activities in security carrier operations. Not only does it affect the effectiveness of vehicle usage and safety during cash conveyance, but also the carrier's operating costs [3]. In addition, the cash transportation vehicle routing and scheduling problems involve complicated analyzes among many time-windows and space constraints which are highly correlated to each other, coupled with certain side constraints. It is difficult to apply the traditional integer programming techniques (e.g., traditional vehicle routing and scheduling models) to construct and to efficiently solve these types of problems. The second part of the paper is cash transportation. In this paper, genetic algorithm (GA) is used for transportation operation. GAs are search algorithms designed to mimic the principles of biological evolution in natural genetic system [4]. The rest of the paper is organized as follows: Section 2 presents the fundamentals of ANNs and GAs. The application is detailed in Section 4. Finally, conclusion is given in Section 4.

2 METHODOLOGY

In this section, the fundamentals of ANNs and GAs are briefly introduced.

2.1 Artificial Neural Networks

ANNs are the distributed processing systems which have been inspired by the biological nerve system [5]; that have been widely implemented as pattern recognition, function approximation optimization, simulation, prediction, among many other application areas. It is also preferred for the forecasting model because benefit from the advantage of no requirement for any assumptions, extrapolating from historic data to generate forecasts and solving the complex nonlinear problems successively. ANNs composed of an input layer, some hidden layers and an output layer. Each layer has a certain number of neurons which are the basic processing elements of ANN [5]. There are numerous algorithms available for training neural network models; the most popular of them is the back-propagation algorithm, which has different variants. Standard back propagation is the gradient descent algorithm. An ANN with a back-propagation algorithm learns by changing the connection weights, and these changes are stored as knowledge [5]. Backpropagation algorithm is essentially a network of simple processing nodes arranged into different layers as input, hidden and the output. The input layer propagates components of a particular input vector after weighting these with synaptic weights to each node in the hidden layer. At each node, these weighted input vector components are added. Each hidden layer computes output corresponding to these weighted sums through a non-linear/linear function (e.g. Logsig, Tansig and Purelin). These functions are known as transfer functions. The transfer functions are defined by expressions given below:

$$\text{Logsig transfer function: } f(n) = \frac{1}{1+\exp(-n)} \tag{1}$$

$$\text{Tansig transfer function: } f(n) = \frac{2}{1+\exp(-2n)} - 1 \tag{2}$$

$$\text{Purelin transfer function: } f(n) = n \tag{3}$$

where $f(n)$ = output of the transfer function, and n = weighted sum of inputs [14].

Thus, each of the hidden layer nodes compute output values, which become inputs to the nodes of the output layer. At nodes of output layer, also a weighted sum of outputs of previous layer (hidden layer) is obtained and processed through a transfer function. Thus, the output layer nodes compute the network output for the particular input vector. In this paper, output nodes use linear transfer function. In the training algorithm, the main objectives are: a) to determine the output of the neural network for a given input data, b) to find the difference of the obtained output from the desired output, and then, c) to adjust the weights of synapses to minimize the difference. The weights are adjusted to minimize the error by propagating the output error backward through the network. The errors for the hidden layer nodes can be computed by assigning a portion of the error at each output layer node to the hidden layer node, which fed that output node. The amount of error due to each hidden layer node depends on the size of the weight assigned to the connection between the two nodes [6].

2.2 Genetic Algorithms

GAs are search algorithms which are designed to mimic the principles of biological evolution in natural genetic system. GAs are also known as stochastic sampling methods, and they can be used to solve difficult problems in terms of objective functions that possess 'bad' properties, such as multi-modal, discontinuous, non-differentiable, etc. These algorithms maintain and

manipulate a population of solutions and implement their search for better solutions based on 'survival of the fittest' strategy [4]. Biologically inspired operators like crossover and mutation are applied on these strings to yield a new generation of strings. The process of selection, crossover and mutation continues for a fixed number of generations or till a termination condition is satisfied [7].

GAs have many advantages over the traditional optimization methods. In particular, GAs do not require function derivatives and work on function evaluations alone; they have a better possibility of locating the global optimum because they search a population of points rather than a single point and they allow for consideration of design spaces consisting of a mix of continuous and discrete variables. In addition, GAs provide the decision-makers a set of acceptable optimal solutions (rather than a single solution) from which they can select the most appropriate one. [8]. One of the main disadvantages of GA techniques is that, although as global optimization techniques, they have good initial convergence characteristics, but they may slow down considerably once the region of optimal solutions has been identified [8]. Elitism is commonly used to improve the convergence of GA. It has been well established that elitist evolutionary algorithms (EAs) have better convergence characteristics than non-elitist EAs [8]. Generally, elitism preserves the current best solution(s) and transfers them to subsequent generations. In this work, elitism is implemented by simply carrying the best solution of the population to the next generation. In addition, to further improve the convergence of the GA, a neural network, trained with the population of individuals and their fitness values, is used to find even a better solution than the current best solution and transfer it to the next generation as well.

3 APPLICATION

The aim of the application is to establish a distribution network combining all these processes. An imaginary bank is established only serving in Istanbul and the total number of ATMs is determined which are necessary for the bank. These ATMs are determined for deposited (D) and withdrawn (W) money for 365 days. The forecasted data is obtained using ANNs for 366th day and distribution network is obtained using a GA by these forecast results.

3.1 Daily Demand Forecast with ANNs

Firstly, the number of ATMs is determined using the information in the Table 1 [9]. The number of ATMs is set as 1 ATM per 50,000 people and the total number of ATMs is found as 305. Hour coefficients (HC) are determined as shown in the Table 2.

Table 1: District information table

	District	District Name	Area	Population	ATM	LQC	PDC
	0	Arnavutköy	450.35	215.531	5	3	1
	1	Avcılar	42.01	407.240	9	5	5
	2	Bağcılar	22.36	752.250	16	3	9
	3	Bahcelievler	16.62	602.931	13	5	9
Europe	…		…	…	…	…	…
	21	Silivri	869.52	155.923	4	5	1
	22	Sultangazi	36.3	505.190	11	1	5
	23	Sisli	10.71	274.420	6	9	7
	24	Zeytinburnu	11.59	292.313	6	5	7
	25	Adalar	11.05	16.166	1	7	3
	26	Atasehir	25.2	405.974	9	7	7
	27	Beykoz	310.36	248.056	5	5	1
Asia	…	…	…	…	…	…	…
	36	Tuzla	123.63	208.807	5	5	3
	37	Ümraniye	45.31	660.125	14	5	5
	38	Üsküdar	35.33	534.636	11	7	7

Hour	HC	Hour	HC
1:00	7	13:00	9
2:00	5	14:00	9
3:00	4	15:00	8
4:00	2	16:00	8
5:00	1	17:00	7
6:00	1	18:00	7
7:00	2	19:00	7
8:00	3	20:00	7
9:00	3	21:00	8
10:00	3	22:00	9
11:00	4	23:00	9
12:00	9	0:00	9

Deposited and withdrawn money is calculated taking into consideration the coefficients in the Table 1 and Table 2 for each ATM. These calculation methods are:

For the amount withdrawn:

$$\sum_0^{23}((HC + R_{max}(30 \times PDC)) \times (20 + R_{max}(5 \times LQC))) \qquad (4)$$

For the amount deposited:

$$\sum_0^{23}(((\tfrac{HC}{4}) + Rmax(7.5 \times PDC)) \times (40 + Rmax(8 \times LQC))) \qquad (5)$$

Where $R_{max}(x)$ is generate a random number between 0 and x, HC is hour coefficient, PCD is population density coefficient, LQC is life quality coefficient. Total deposited and withdrawn amounts of money are generated using these equations (4, 5) for each ATM on a daily basis. The first part of the equation is the number of people, and the second part is the amount of money. A portion of the result is shown in Table 3. The results are divided into a large number, such as 100,000 to be used for ANN.

Table 3: A portion of the generated data

Arnavutköy					
1. ATM		**2.ATM**		**3.ATM**	
W	**D**	**W**	**D**	**W**	**D**
0.015418	0.005101	0.014335	0.004557	0.012722	0.004765
0.012621	0.005073	0.015266	0.005428	0.011706	0.004908
0.013728	0.005051	0.012899	0.004821	0.015254	0.004815
0.012244	0.004564	0.011268	0.004621	0.011682	0.005339
0.012471	0.00551	0.013325	0.004908	0.013194	0.004939
…	…	…	…	…	…
0.014948	0.005624	0.012762	0.004609	0.015588	0.005291
0.013266	0.005091	0.011146	0.004661	0.012508	0.005139
0.012708	0.005961	0.013488	0.005169	0.017095	0.004753
0.013211	0.005262	0.00989	0.004603	0.015356	0.005322
0.014025	0.005938	0.012624	0.004696	0.009427	0.004224

After the data-generating process, ANN is implemented for forecasting of cash management. Firstly, the network is designed for the application of ANN. Designed network structure is as: 1 input layer, 1 input neuron; 1 hidden layer, 10 hidden neurons; 1 output layer, 1 output neuron. In this network, learning rate is 0.4, momentum coefficient is 0.01, iteration number is 5000. Sigmoid function and purelin function are used out of hidden layer and output layer, respectively. MSE values for training, testing and forecasted data are presented in Table 4. The MSE value for training is found between 10^{-10} and 10^{-7}. MSE for testing is found between 10^{-7} and 10^{-6}. After determining the forecast values, the vehicle routing process is started.

Table 4: The mean squared error for train, test and forecasted data

| | Arnavutköy | | | | | |
| | 1.ATM | | 2.ATM | | 3.ATM | |
	W	D	W	D	W	D
Training	3.17E-08	4.20E-10	5.09E-08	2.38E-09	3.09E-07	7.39E-10
Testing	3.89E-06	2.20E-07	9.57E-06	3.09E-07	6.44E-06	3.11E-07
Forecasted	0.012028	0.00516	0.01511	0.00554	0.014264	0.004714

3.2 Vehicle Routing Process with GA

Firstly, we determined distances between the district and Maslak (39) headquarters. A portion of these distances are shown in Table 5.

Table 5: The distances for Maslak

	0	1	2	3	4	5	6	7
0	0	42524	25167	37244	38089	29509	26578	42707
1		0	26529	17729	12422	28388	24703	33431
2			0	7218	14495	9124	13731	24755
3				0	5598	16213	9447	20500
4					0	19084	18213	20876
5						0	19704	30728
6							0	16866
7								0

Secondly, we determined average speed values for vehicle routing operations in Istanbul. These values are shown in Table 6.

Table 6: Average speeds

Hour	Speed (km/h)	Hour	Speed (km/h)
1:00	50	13:00	20
2:00	60	14:00	30
3:00	70	15:00	50
4:00	80	16:00	40
5:00	80	17:00	20
6:00	60	18:00	30
7:00	30	19:00	30
8:00	20	20:00	30
9:00	30	21:00	50
10:00	40	22:00	60
11:00	50	23:00	40
12:00	30	0:00	40

After these operations, GAs are designed and population is generated according to this design. Districts are identified as the genes of individuals in the population. An individual is similar to the following: [39-0-1-2-3-4-5-6-39]. Each district is represented by a number code for that particular district. The population is composed of 50 individuals. The mutation rate and elite individual number is determined as 5%. The generation number is determined as 40. The fitness function is set for selection of the individual with high demand and low total time. When loading money into an ATM is set to 10 minutes. To calculate the total time:

$$D_{ij}/V_h + (NA_i \times 10) \tag{4}$$

Where D_{ij} is distance from i to j and NA_i is the number of ATMs in the district i.
To calculate the total number of vehicles needed:

$$39/(AND - 1) \tag{5}$$

Where *AND* is the average number of district.

Finally, we initialize the average number of district and start time using the application interface. This interface is shown in Figure 1a. If the average number of district is set to 5, the required route and the number of vehicles is 10 and the start time is set to 02:00; the results are as shown in Figure 1b and Table 7. Total processing time is 5059 minutes.
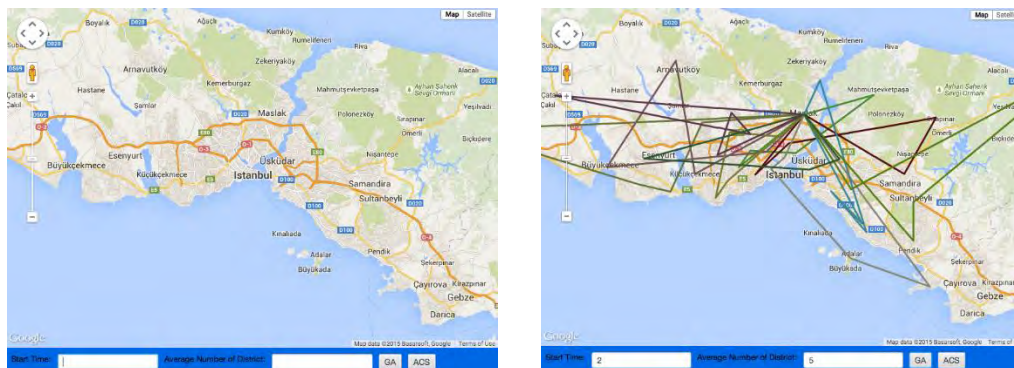
Figure 1: a) Application interface   *b)* Result of the GA on the map

Table 7: Result of the GA

	Distance (m)	Time (min)	Route
Vehicle 1	122000	334	[39-27-7-6-17-39]
Vehicle 2	166000	400	[39-24-23-28-33-39]
Vehicle 3	178000	492	[39-14-10-0-19-39]
Vehicle 4	120000	527	[39-18-12-4-3-39]
Vehicle 5	204000	396	[39-9-15-25-36-39]
Vehicle 6	268000	554	[39-5-1-8-21-39]
Vehicle 7	145000	559	[39-20-30-29-31-39]
Vehicle 8	245000	644	[39-26-35-34-32-39]
Vehicle 9	166000	597	[39-11-16-22-2-39]
Vehicle 10	139000	556	[39-13-38-37-39]
Total	1753000	5059	

4 CONCLUSION

In this paper, we evaluate the results of the application that understanding the major costs of the banking sector, which the opportunity and transportation costs can be minimized. In the first step of the proposed approach, ANN is applied to forecast the amount of withdrawal and deposited for ATM. We used back propagation ANN algorithm and the designed ANN is constructed based back propagation algorithm. ATM demand forecasts for next year's first day is found using annual demand data in ANN method. Thus, the gain is provided by the opportunity cost minimization can be used in other fields. In the first step of the proposed approach, GA has been decided to apply for solving the problem. The shortest routes to reach the bank ATMs are determined by the GA. Thus, the transport and distribution costs are minimized and the amount of time and labor is made available in other fields. For future work, by changing GA parameters could be achieved better results. Similarly, the experiences of the project can be easily applied on similar routing and optimization problems.

References

[1] Venkatesh, K., Ravi, V., Prinzie, A., Van den Poel, D., 2014. Cash demand forecasting in ATMs by clustering and neural networks. European Journal of Operational Research, 232:383–392.

[2] Castro, J., 2009. A stochastic programming approach to cash management in banking. European Journal of Operational Research, 192(3): 963–974.

[3] Yan, S., Wang, S., Wu, M., Roth, J.P., 2012. A model with a solution algorithm for the cash transportation vehicle routing and scheduling problem. Computers & Industrial Engineering, 63: 464–473.

[4] Changyu, S., Lixia, W., Qian, Li., 2007. Optimization of injection molding process parameters using combination of artificial neural network and genetic algorithm method:, Journal of Materials Processing Technology, 183: 412–418.

[5] Ata, R., 2015. Artificial neural networks applications in wind energy systems: a review. Renewable and Sustainable Energy Reviews, 49: 534–562.

[6] Prasad, R., Pandey, A., Singh, K.P., Singh, V.P., Mishra, R.K., Singh, D., 2012. Retrieval of spinach crop parameters by microwave remote sensing with back propagation artificial neural networks: A comparison of different transfer functions. Advances in Space Research, 50: 363–370.

[7] Maulik, U., Bandyopadhyay, S., 2000. Genetic algorithm-based clustering technique. Pattern Recognition, 33: 1455-1465.

[8] Javadi, A.A., Farmani, R., Tan, T.P., 2005. A hybrid intelligent genetic algorithm. Advanced Engineering Informatics, 19: 255–262.

[9] T.C. Istanbul Metropolitan Municipality, Istanbul districts area and population data. http://www.ibb.gov.tr/tr-TR/kurumsal/Pages/IlceveIlkKademe.aspx, Accessed: 10 March 2017.

A VARIABLE NEIGHBOURHOOD SEARCH BASED HEURISTIC FOR BUFFER ALLOCATION PROBLEM IN PRODUCTION LINES

Mehmet Ulaş Koyuncuoğlu
Pamukkale University, Department of Information Processing Center
Pamukkale University, Kinikli Campus, Denizli, Turkey
E-mail: ulas@pau.edu.tr

Leyla Demir
Pamukkale University, Department of Industrial Engineering
Pamukkale University, Kinikli Campus, Denizli, Turkey
E-mail: ldemir@pau.edu.tr

Abstract: The buffer allocation problem, i.e. how much buffer storage to allow and where to place the buffer, is an important research issue in designing production lines. In this study, we present a new solution approach based on variable neighbourhood search to determine the optimal buffer allocations for a serial production line with unreliable machines. The objective is to maximize the throughput of the line under the constant total buffer size constraint. To evaluate the throughput of the line simulation is used. The performance of the proposed heuristic approach is demonstrated by a numerical example.

Keywords: Buffer allocation problem, Production lines, Variable neighborhood search.

1 INTRODUCTION

A production line is composed of machines in series and the buffer areas between these machines as seen in Figure 1. The items pass through all the machines in the same sequence. The performance of production lines is affected by either variable processing times or machine failures. The effects of these variations can be reduced by using storage buffers between the machines.
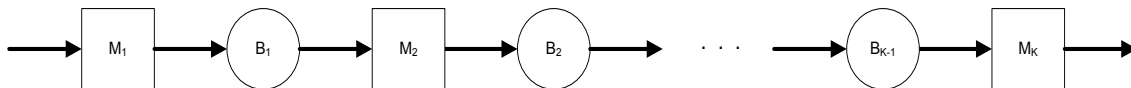
Figure 1: A production line consisting of K machines and K-1 buffers

The main reason for having storage buffers is to allow sequential machines to operate nearly independently of each other. If allocating buffers between the machines are allowed the idle time due to starving and blocking is reduced. In this way, the production rate of the line is increased. On the other hand, buffering requires additional capital investment and floor space and it may be expensive. Including buffers in a production line also increases in-process inventory. Because of this trade-off finding optimal buffer configurations is an important optimization problem in designing production lines.

In this study, we propose a new solution approach based on variable neighborhood search for solving buffer allocation problem in unreliable production lines. Variable neighborhood search (VNS) is a meta-heuristic for solving combinatorial optimization problems. To the best of our knowledge the VNS has not been employed before for solving BAP. Since there are successful applications of this method on combinatorial optimization problems such as vehicle routing ([1], [11], [18], [27]) and knapsack problems [28] it is thought that it

produces good results for also buffer allocation problem which can be formulated as a knapsack problem.

The rest of this study is organized as follows. The next section explains the buffer allocation problem. The details of the proposed solution approach are given in section 3. Section 4 presents the performance of the proposed VNS-based heuristic algorithm by a numerical example. Finally, the concluding remarks and some future research directions are given in section 5.

2 THE BUFFER ALLOCATION PROBLEM

The buffer allocation problem (BAP) deals with finding optimum buffer configurations, i.e. the optimum size and the location of buffers, to achieve a specified objective under certain constraints. The buffer allocation problem is mainly formulated as three types of mathematical structures. The first type focuses on the throughput maximization, the second type employs the minimization of the total buffer size as an objective function, and the final formulation aims to minimize the amount of WIP (Work-in-Process) in the production line. The reader can refer to Demir et al. [10] for the details of BAP formulations.

The first study in research of buffer allocation problem is presented by Koenigsberg [14] and since then many studies have been published in this area. A comprehensive analysis of mathematical models describing the effect of the buffers in production lines, the reader can refer to the studies of Buzacott and Shanthikumar [3] and Papadopoulos et al. ([20], [22]). In addition to these studies, Dallery and Gershwin [7] and Papodopoulos and Heavey [21] provide detailed information on behaviour of production lines with buffers.

Demir et al. [10] present a detailed analysis of the studies on BAP published up to 2013. In their study, BAP is classified according to main four categories, i.e. the topology of the line studied, the reliability of the machines in the line, the objective function considered in the problem and the methods used to solve the problem. Later, another survey study on BAP is presented by Weiss et al. [28] who classified BAP according to the decision problems, solution approaches, and test instances. They have constructed a scheme according to the objective function, constraints, test sizes, exact solution algorithms, and integrated solution approaches. The reader can refer to these two studies for more information on BAP.

Chow [5] stated that the buffer allocation problem is difficult for two reasons: (1) there is no algebraic relation between the throughput of the line and buffer sizes; and (2) the combinatorial nature of the problem. In general, to solve the buffer allocation problem a generative method and an evaluative method are employed in an iterative manner as depicted in Figure 2. In this solution process, an evaluative method is used to obtain the performance values of the line, e.g., throughput, average WIP etc. The generative method is then communicated to the evaluative method to obtain optimum/near-optimum buffer configurations. In the following subsections, these methods are briefly explained.
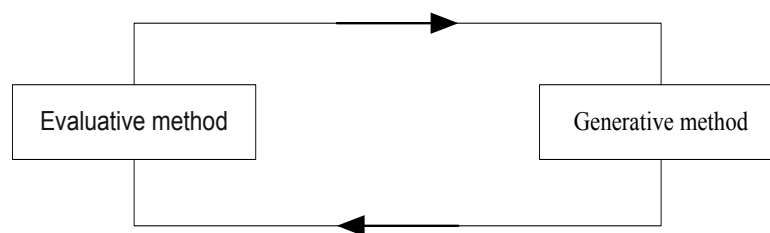
Figure 2: General solution process of BAPs

2.1 Evaluative methods

Mainly two methods are used as an evaluative method for solving BAPs: analytical methods and simulation. Exact analytical results can be obtained only for short production lines and they are usually based on the queuing models. When the long production lines are considered, generally approximate evaluative methods such as decomposition method, the aggregation method, and the generalized expansion method are used. Among these methods, the decomposition method is the most widely used evaluation method as it reaches the solution quickly and its accuracy is very high ([8], [8], [12], [23], [24], [25]). However, it can be applicable only under the certain assumptions such as geometric or exponentially distributed failure and repair rates.

Because it needs some restrictive assumptions analytical methods are not computationally efficient in dealing with real world BAPs. If the objective is to model a large and complex real system, simulation provides many advantages in comparison to analytical methods. As long as the components of the system and their logical relations are well understood, simulation modeling is very flexible in terms of model development. Successful examples of real world buffer allocation problem can be found in [4], [15], and [16].

2.2 Generative methods

Generative methods focus on finding optimal buffer configurations to achieve the specific objective. Various techniques are used as generative method in BAP. The method called "complete enumeration" in the literature where all possible solutions are evaluated is the simplest optimization method. However, this method can be applied to only the small sized problems. As the total amount of buffer and the number of machines in the line increases, the solution space also exponentially increases and it is impossible to obtain solution in a reasonable time by complete enumeration. Hence, different search methods or meta-heuristic method is widely employed for solving BAP.

Some research employ problem specific heuristics while the others used well-known meta-heuristic search methods such as genetic algorithms [15], tabu search ([8], [9], [25]), and simulated annealing ([16], [19]).

In this study a recent meta-heuristic method known as variable neighbourhood search is employed for solving BAP in unreliable production lines. To the best of our knowledge the VNS has not been used before for solving BAP. Hence, in this study, the performance of the VNS is tested for BAP whether it produces good results as it is expected. The proposed solution method based on VNS are presented in the next section.

3 PROPOSED SOLUTION APPROACH

VNS is a simple and effective meta-heuristic search method that aims to solve combinatorial optimization problems and it is proposed by Mladenovic [17] and Hansen and Mladenovic [18]. VNS is based on a single solution and it uses static objective function and various neighbourhood structures. The algorithm systematically changes the neighbourhood structures while searching the solution space. VNS is employed different combinatorial problems such as p-median problems ([2], [18][18]), vehicle routing problems ([1], [11], [26]), and knapsack problems [27].

While designing VNS algorithm, first of all, the neighbourhood structures to be used is determined. Moreover, the type of local search which is employed while evaluating neighbourhood solutions is also be determined. The initial solution is generated randomly, or

it can be generated according to a specific rule, which improves the performance of the algorithm. Finally, as in all meta-heuristics, the stopping criterion should be defined.

We adopt the VNS method for solving BAP as follows. The initial solution is generated randomly. The neighbourhoods are generated by using increment-decrement strategy which increases the size of buffer one unit in a randomly selected buffer location i, while decreasing one unit in another location j so as to keep the total buffer size constant. All possible neighbourhoods are generated employing this strategy. After that the throughput values of each generated configurations, i.e. neighbourhood solutions, are obtained by simulation. As a local search Lin-Kernighan heuristic is used to improve these solutions. The algorithm is terminated when the maximum allowable iteration number is reached. The flowchart of the proposed VNS-based heuristic algorithm is given in Figure 3.
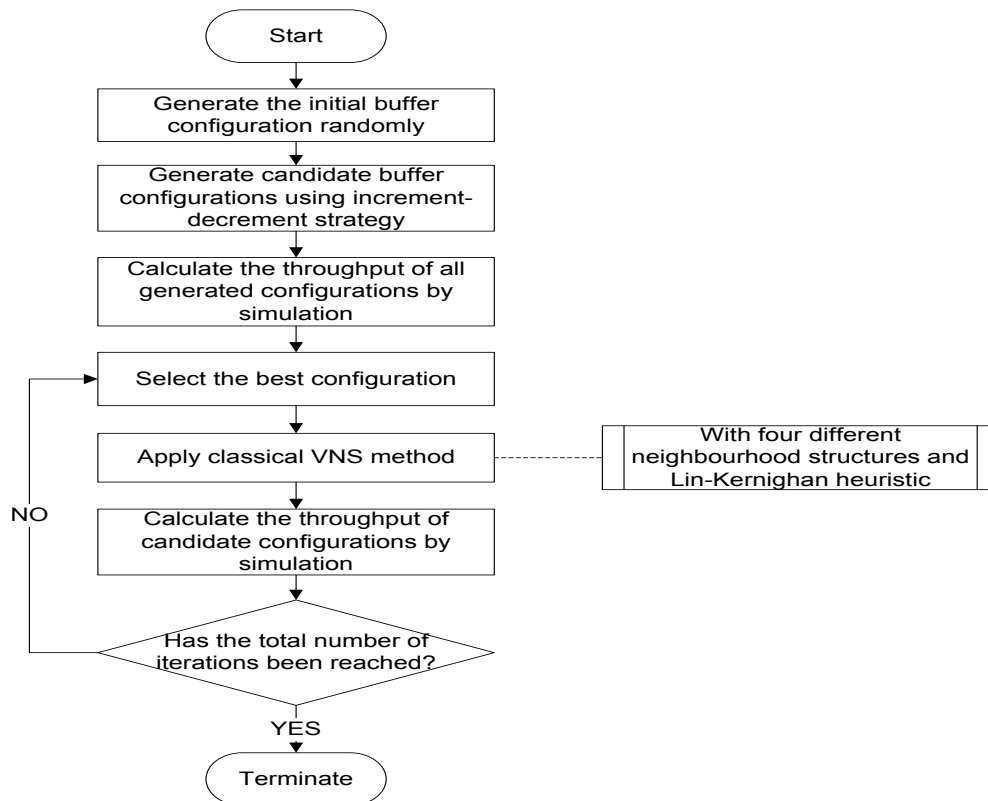
Figure 3: Proposed VNS-based heuristic algorithm

4 NUMERICAL EXAMPLE

The performance of the proposed solution approach is tested on a five-machine production line initially proposed by Ho et al. [13]. Same example is later used by Gershwin and Schor [12] and Demir et al. [8]. It is assumed that all machines have same deterministic processing times which are one time unit. The reliability parameters of the machines are given in Table 1.

Table 1: Reliability parameters of five-machine production line

Machine	1	2	3	4	5
MTTR=$1/r_i$	11	19	12	7	7
MTBF=$1/p_i$	20	167	22	22	26

The total buffer size is set to 31 and the lower bound on each buffer location is set to 4 as in the referred publications. The simulation is run for 100 000 parts and 50 replications. The experiments are carried out on a computer having 2.40 GHz Pentium (R) 4 CPU processor and 4 GB of RAM. The comparative results are shown in Table 2.

Table 2: Results of comparative experimental study.

Case	Buffer i				N	Throughput
	1	2	3	4		
Ho et al (1979).	5	11	8	7	31	0.4931
Gershwin & Schor (2000)	7	10	10	4	31	0.4962
Proposed VNS-based Algorithm	7	10	10	4	31	0.4962

As it is seen in Table 2, the proposed VNS-based heuristic algorithm produces similar results as in the study of Gershwin and Schor [12]. This result encourages us that the proposed VNS-based heuristic algorithm is applicable for large-sized problems especially dealing with real-world cases. So it can be concluded that the proposed VNS-based heuristic algorithm produces very promising results.

5 CONCLUSION

The buffer allocation problem in unreliable production lines is studied in this paper. The objective is to maximize the throughput of the line under total buffer size constraint. To achieve this objective simulation is employed as an evaluative method and a VNS-based heuristic algorithm is proposed as a generative method. The performance of the proposed algorithm is tested on a benchmark problem taken from the literature. The results show that the proposed VNS-based heuristic produces very promising results.

In this study, the performance of the proposed VNS-based heuristic is tested only on a five machine production line. It is planned that this study is extended to longer lines such as 20 and 40 machine lines to prove the efficiency of the proposed approach. Moreover, as a future study, it is aimed that the proposed approach is applied to real-world problems.

References
[1] Braysy, O. 2003. A reactive variable neighbourhood search for the vehicle-routing problem with time windows. INFORMS Journal on Computing, 15: 347-368.
[2] Brimberg, J., Mladenovic, N. 1996. A variable neighborhood algorithm for solving the continuous location allocation problem. Studies in Locational Analysis, 10: 1–12.
[3] Buzacott, J. A., Shanthikumar, J. G. 1993. Stochastic models of manufacturing systems, New Jersey: Prentice-Hall.
[4] Can, B., Heavey, C. 2012. A comparison of genetic programming and artificial neural networks in metamodeling of discrete-event simulation models. Computers & Operations Research, 39(2): 424–436.
[5] Chow, W. M. 1987. Buffer capacity analysis for sequential production lines with variable process times. International Journal of Production Research, 25(8): 1183–1196.
[6] Costa, A., Alfieri, A., Matta, A., Fichera, S. 2015. A parallel tabu search for solving the primal buffer allocation problem in serial production systems. Computers & Operations Research, 64: 97–112.

[7] Dallery, Y., Gershwin, S. B. 1992. Manufacturing flow line systems: A review of models and analytical results. Queuing Systems, 12 (1-2): 3–94.

[8] Demir, L., Tunalı, S., Løkketangen A. 2011. A tabu search approach for buffer allocation in production lines with unreliable machines. Engineering Optimization, 43 (2): 213-231.

[9] Demir, L., Tunalı, S., Eliiyi, D. T. 2012. An adaptive tabu search approach for buffer allocation in unreliable non-homogeneous production lines. Computers & Operations Research, 39: 1477-1486.

[10] Demir, L., Tunalı, S., Eliiyi, D. T. 2014. The state of the art on buffer allocation problem: A comprehensive survey. Journal of Intelligent Manufacturing, 25 (3): 371-392.

[11] Fleszar, K., Osman, İ. H., Hindi, K. S. 2009. A variable neighbourhood search algorithm for the open vehicle routing problem. European Journal of Operational Research, 195(3): 803-809.

[12] Gershwin, S. B., Schor, J. E. 2000. Efficient algorithms for buffer space allocation. Annals of Operations Research, 93(1): 117–144.

[13] Ho, Y.C., Eyler, M.A., Chien, T.T. 1979. A gradient technique for general buffer storage design in a production line, International Journal of Production Research 17(6): 557–580.

[14] Koenigsberg, E. 1959. Production lines and internal storage - A review. Management Science, 5: 410–433.

[15] Köse, S. Y., Kılınçcı, O. 2015. Hybrid approach for buffer allocation in open serial production lines. Computers & Operations Research, 60: 67–78.

[16] Köse, S. Y., Demir, L, Tunalı, S., Eliiyi, D. T. 2015. Capacity improvement using simulation optimization approaches: A case study in the thermo-technology industry, Engineering Optimization, 47 (2): 149-164.

[17] Mladenovic, N. 1995. A variable neighbourhood algorithm - a new metaheuristic for combinatorial optimization. Presented at Optimization Days, Montreal.

[18] Mladenovic, N., and Hansen, P. 1997. Variable neighbourhood search. Computers & Operations Research, 24: 1097-1100.

[19] Nahas, N., Ait-Kadi, D., Nourelfath, M., 2006. A new approach for buffer allocation in unreliable production lines. International Journal of Production Economics, 103: 873-881.

[20] Papadopoulos, C. T., O'Kelly, M. E. J., Vidalis, M. J., Spinellis, D. 2009. Analysis and design of discrete part production lines. New York: Springer Science+Business Media.

[21] Papadopoulos, H. T., Heavey, C., 1996. Queuing theory in manufacturing systems analysis and design: A classification of models for production and transfer lines. European Journal of Operational Research, 92: 1-27.

[22] Papadopoulos, H. T., Heavey, C., Browne, J. 1993. Queuing theory in manufacturing systems analysis and design, London: Chapman and Hall.

[23] Shi, C., Gershwin, S. B. 2009. An efficient buffer design algorithm for production line profit maximization. International Journal of Production Economics. 122: 725-740.

[24] Shi, C., Gershwin, S. B. 2016. A segmentation approach for solving buffer allocation problems in large production systems. International Journal of Production Research, 54(20): 6121–6141.

[25] Shi, L., Men, S. 2003. Optimal buffer allocation in production lines. IIE Transactions, 35: 1-10.

[26] Şevkli, A. Z., Güler, B. 2017. A multi-phase oscilated variable neighbourhood search algorithm for a real-world open vehicle routing problem. Applied Soft Computing, 58: 128-144.

[27] Taşgetiren, M. F., Pan, Q. K., Kızılay, D., Suer, G. 2015. A differential evolution algorithm with variable neighborhood search for multidimensional knapsack problem. IEEE Congress on Evolutionary Computation (CEC), 2797-2804.

[28] Weiss, S., Schwarz, J. A., Stolletz, R. 2016. The Buffer Allocation Problem: Formulations, Solution Approaches, and Test Instances. Econometrics: Econometric & Statistical Methods - Special Topics eJournal, Working papers.

A ROBUST OPTIMIZATION APPROACH FOR BETTER PLANNING OF A RETAIL SUPPLY CHAIN PRODUCT FLOW
(EXTENDED ABSTRACT)

Tea Vizinger
Faculty of logistics, University of Maribor, tea.vizinger@um.si
Tomaž Kokolj
tomaz.kokolj@gmail.com
Janez Žerovnik
Faculty of Mechanical Engineering, University of Ljubljana, janez.zerovnik@fs.uni-lj.si

Abstract: Paper discusses a robust optimization approach that addresses the inventory allocation problem from the tactical level of planning of a retail supply chain product flow. Previous research work shows that several metaheuristics can be used to provide solutions of the multi-objective optimization problem including a stochastic linear program. The local search based algorithms were applied to minimize the sum of costs of three criteria. The solutions were of very good quality, but the time needed was prohibitively large. In this work we introduce new heuristics that dramatically improves the time complexity of the algorithms and also allows more flexibility in choosing the final solution that need not always be minimizing the sum of costs.

Keywords: Robust optimization, Stochastic demand, Tabu search, Distribution flow, Flow coordination

1   INTRODUCTION

A robust stochastic model for tactical planing of a retail supply chain product flow was first introduced in [6]. The idea was to align the decisions reflecting the ordering policies that are in the retail companies usually taken independently by several decision makers. Warehouse managers' orders are naturally based on price and availability of a product, while store managers' orders are usually based on the actual requirements of the merchant. Needless to say that ordered quantities from the external suppliers that depend on the stock market prices and not on the actual requirements may result in higher or lower stock levels, causing overstocking effects or lost sales (state of stock-out with possible lower sale realization). To integrate the tactical and the operational decisions, an idea of an on-going optimization was introduced in [6]. In the approach proposed, tactical planning for a chosen period (month, season, etc.) defines the appropriate inventory levels at warehouses and stores, and consequently the allocation of resources between retailers' facilities. The replenished quantities are on the operational level defined with the difference between an actual inventory and the pre-defined maximum level of a certain store inventory. Note that the tactical plan already determines the necessary stock levels, the warehouses demand becomes almost deterministic. With precise tactical planning, retailers may be able to contract constant supply quantities, which may result in a lower unit price and related higher profit. At the stores' side of the supply chain, precisely pre-defined inventory levels prevents stock accumulation, which may result in increased products' quality and related customer service level.

The model for a retail supply chain product flow coordination, which was introduced in [6] corresponds to the optimization of three conflicting criteria (distribution costs, overstocking effects and lost sales). Direct distribution costs are going to be realized later, but not in the exact calculated amount, hence the rough estimates for transport and holding costs are used. On the other hand, we are not certain that the chosen plan to sell the distributed quantities

will be exactly realized. If not, we will face the second type of "cost", i.e. indirect supply risk "cost". Because both type of supply risk (overstocking effects and lost sales) can not appear at the same time (they are mutually exclusive) it is reasonable to treat all three costs separately.

The optimization problem that arises from the model is a computationally hard problem. Several versions of the local search heuristics were adopted in [6], including iterative improvement (a basic form of local search), tabu search and threshold accepting. The best performance was shown by the tabu search heuristic that proved to provide very good solutions on the instances tested. However, the computational time for a single product instances of moderate size take a lot of time. Regardless of the fact that this calculations are to be performed only occasionally, it is important to have a faster method if possible. Furthermore, we will try to improve the objectives' diversity using some other (cost) metric. Goal function defined as a simple or a weighted sum of the criteria has already been shown to be an inappropriate choice.

The paper is further organized as follows. An insight into related work is provided in Section 2. The problem formulation and the new heuristic solution procedure is provided in Section 3. Computational experiment is described in Section 4. Conclusions and interesting research avenues for future research work are given in Section 5.

2 RELATED WORK

The problem we deal with in this paper interrelates two usually separately treated problems; a single-item lot sizing problem and a resource allocation problem. To the best of our knowledge, the introduced stochastic model in [6] is the first that tackles the coordination problem at the tactical level of planning. Treated inventory allocation problem may be seen as a generalized material flow problem, where the goal is to minimize the distribution costs of goods delivered from some supply points to a number of destination points [1]. Nevertheless that routing and further operational decisions are excluded from our core tactical planning, we will further briefly refer to the combined inventory routing problem (as there exist similarities to the treated inventory allocation problem).

To tackle the inventory routing problem, the literature mostly introduces the use of mixed integer optimization, multi objective optimization and stochastic programming (e.g. [3, 4]). The idea is to find the appropriate policy with minimal costs of distribution, minimal overstocking effects and maximal customer service level. Liu and Papageorgiou [4] interpret customer service level with the percentage of a customer demand satisfied on time. Therefore, lower customer service level causes lost sales or lost customers, resulting in profit loss of the supply chain.

Most of the applied stochastic models are two stage programs, and are used to deal with the demand uncertainties when assigning probability distributions [2]. A stochastic transportation problem may be transformed into a deterministic by removing the demand constraints, which are used to introduce new cost function related to the expected extra cost (resulting difference between the delivered amount and the actual demand). Nevertheless that risk is in our work measured with the cost metric, the direct and indirect costs should not be summed up, because they rise from totally different origin. In contrast to the well known Newsvendor approach for coordination of the supply chain flow we are not limited to use only a simple or a weighted sum of the considered criteria.

For the time prohibitive stochastic programs, the use of heuristic approaches (that provides good feasible solutions) have become very popular. Beside stochastic programming, Grossman and Guillén-Gosálbez [3] provide an insight into the robust optimization and probabilistic programming. In many practical situations one may not be able to identify the underlying probability distributions or such a stochastic description may simply not exist [5]. In such a situation it is reasonable to fit a suitable probability distribution for each parameter based on an experts subjective knowledge derived from past experiences and feelings. Uncertainties are

therefore unknown-but-bounded quantities, while constraints are satisfied for all realizations of the uncertain parameters. Therefore in robust programming, not every scenario represents a feasible solution. When uncertainty is once realized, the solution obtained from the robust optimization ensures that a set of constraints are satisfied with a certain probability. Introducing a probabilistic level of constraint satisfaction corresponds to the probabilistic programming also known as chance-programming.

3   ROBUST OPTIMIZATION APPROACH

In the approach outlined below, a retail distribution flow is represented as a multi objective discrete optimization problem. A typical retail supply chain consists of one or several warehouses $i \in (1, ..., I)$ and a set of stores $j \in (1, ..., J)$, where dc_{ij} represents distribution cost, and x_{ij} is the quantity of distributed product. There are two types of vertices. The vertices of the first type have positive weights a_i that correspond to a fixed supply available at each origin or warehouse. The second type of vertices have positive weights C_j that model fixed inventory holding capacities of stores. In addition, we are given stores' demand as random variables b_j.

Remark. Customers' shopping habits are modelled by random variables b_j with given probability distributions. Here, we mainly consider discrete distributions and assume that we are given hypothetical distributions based either on knowledge of stores' managers or on information from the system on recently sold quantities.

Previously used heuristic solution procedures applied to the stochastic model for better planning of a retail supply chain product are time prohibitive, therefore we further apply a new heuristic solution procedure. A robust optimization approach is performed in three consecutive steps, where we first generate an initial solution (having at most some percent of lost sales or overstocking effects realization). Among the limited number of testing scenarios we iteratively seek for solutions with minimal supply risks. For the set of best solutions (with minimal supply risk) we evaluate the distribution costs. This approach is closely related to the robust programming, at least what concerns the generation of scenarios. Moreover, the objective function to be minimized is no longer represented as a sum of all three costs, but as a sum of supply risks. This allows us to exclude the time prohibitive linear programming from the iterative improvement, which enormously sped up the used heuristic solution procedure.

3.1   Generating an initial solution

As mentioned above, we first randomly generate an initial solution (or, it may be known from the past experience). Let us assume for a moment that we are familiar with the quantities that are to be sold at store j (in reality, actual demand is only known a posteriori), given in a vector: $k = (k(1), k(2), ..., k(j), ..., k(J))$, where $k(j)$ gives the quantity of the product to be sold at store j. Each $k(j)$ is assumed to take a limited number of values, which are obtained from possible sales realizations for each store j (i.e. from past experience). When generating an initial solution we therefore first generate a scenario k, and use $X(k)$ as the initial solution, where $X(k)$ is the optimal solution of the linear problem solving the deterministic transportation problem corresponding to the scenario k. A solution X is a distribution plan given by a matrix: $X = [x_{ij}]_{i \in I, j \in J}$.

A set of feasible solutions is denoted by χ. A solution X is feasible, if it satisfies the inventory holding capacity of stores C_j, while complies with the supply available at each origin or warehouse a_i. More formally, the quantities of the product that should be available at stores, i.e. x_{0j} are represented with the sum of x_{ij} over all warehouses i. Consequently, the amount of product that will be transported from warehouse i, x_{i0} is the sum of x_{ij} over all stores j.

The generated initial solution X has to be further checked whether it coincides with the specified robust criteria (maximum allowable supply risks). The cost of overstocking effects

OS is calculated as:

$$OS\left(X,k\right) = \sum_j \left(x_{0j} - k\left(j\right)\right) \cdot c_{OS}\left(j\right), \qquad (1)$$

and the cost of lost sales LS is defined as:

$$LS\left(X,k\right) = \sum_j \left(k\left(j\right) - x_{0j}\right) \cdot c_{LS}\left(j\right). \qquad (2)$$

In (3) and (4), $c_{OS}(j)$ is the cost of overstocking effects for a unit of product at store j, and $c_{LS}(j)$ is the cost of lost sales for a unit of product at store j.

As the optimization criteria depend on the a priori unknown values of actual stores' demand we can only optimize the expected values. The expected cost of overstocking effects is simply:

$$Exp\left(OS\left(X\right)\right) = \sum_k p\left(k\right) \cdot OS\left(X,k\right) \qquad (3)$$

and the expected cost of lost sales is:

$$Exp\left(LS\left(X\right)\right) = \sum_k p\left(k\right) \cdot LS\left(X,k\right) \qquad (4)$$

i.e. as the weighted sum of the costs over all scenarios. Assuming that the probability distributions of k are given by (independent) probability distributions at each store, for the random variables $k(j)$, the probability of a scenario k is clearly:

$$p\left(k\right) = \prod_j p\left(j, k\left(j\right)\right). \qquad (5)$$

Expected overstocking effects are calculated for the initially generated solution as well as for the last possible solution (maximal distributed quantities). If the ratio between $Exp(OS(X))$ for the tested solution and $Exp(OS(X))$ for the last possible solution is less than some predefined percentage (for example 30 percent), we may with reasonable confidence accept the initially generated solution. Furthermore, we may seek for feasible solutions with at most e.g. 40 percent of lost sales, where we consider the ratio between the $Exp(LS(X))$ for the tested solution and the $Exp(LS(X))$ for the first possible solution (having minimal distributed quantities).

3.2 Tabu search

For the initial solution with at most e.g. 30 percent of overstocking effects and/or with at most e.g. 40 percent of lost sales realization we apply the tabu search heuristic solution procedure. Starting from current (initial) feasible solution, tabu search generates a random neighbor, and moves to the new solution based on the difference in the goal function. A local change is defined as follows. Given the current solution (the distribution plan X), we randomly select a store and randomly change the amount to be delivered (for ± 1 unit). The procedure moves to a new solution if the value of the goal function is improved, i.e. in our case minimized compared to the objective function value of the previous solution. The objective function to be minimized is here represented as a sum of supply risks ($Exp(OS(X)) + Exp(LS(X))$). Needless to say, the best known solutions are kept in memory and reported as the final results.

3.3 Solution evaluation

For the feasible solutions with minimal supply risks we further solve a linear program and evaluate the distribution costs $DC(X)$. Depending on the company's policy, the preferred solution may for example be the one that appears most often in the set of solutions (with minimal supply risks):

$$\min_{X \in \chi} \left(Exp\left(OS\left(X\right)\right) + Exp\left(LS\left(X\right)\right)\right), \qquad (6)$$

or the one with minimal total costs (trade-off between the direct and indirect costs):

$$\min_{X \in \chi} \left(Exp\left(DC\left(X\right)\right), \left(Exp\left(OS\left(X\right)\right) + Exp\left(LS\left(X\right)\right)\right)\right). \qquad (7)$$

4   NUMERICAL EXAMPLE

Let us consider a retail supply chain with 2 warehouses and 100 stores. Chosen product (bananas) is usually packed into basic units (packages), each weighting $18kg$. It is assumed that transportation between warehouses and stores is provided once per day, and that the stores may order only a whole number of packages, as the package represent a basic transportation unit. For a unit of product (package) we use cost estimates (in Euro) for daily distribution (transport, warehouse) and supply risks (overstocking effects, lost sales). Realistic figures are chosen according to interviews with practitioner and are estimated to: $c = 0.02$ Euro per package km, $h_w = 0.25$ Euro per package per day, $h_s = 0.5$ Euro per package per day, $c_{OS} = 5$ Euro per package per day, and $c_{LS} = 6$ Euro per package per day. Beside that, the entries d_{ij} of the distance matrix (in km) are generated randomly from the interval [10,150] that serves as entries when calculating the transportation costs ($c_{ij} = c \cdot d_{ij}$). Furthermore, the distribution cost (holding and transportation) from warehouse i to store j per unit is computed as $dc_{ij} = c_{ij} + h_w + h_s$.

The problem is generated on the assumption that the demand distributions may take four possible values (ranges) from the discrete interval [1,20], where each value from this interval has the same chance of being selected. As each of b_j, $j = 1, 2, .., 100$ has exactly four possible values, there are 4^{100} possible scenarios (possible sales realizations). The probability estimates (of possible sales realizations) to each of the demand classes are generated from the interval [0,1] so that the sum of probabilities for each store equals to 100 percent. The first value (p_1) was generated from the interval [0,1], the second value (p_2) from the interval $[0, 1 - p_1]$, the third value (p_3) from the interval $[0, 1 - p_1 - p_2]$, and $p_4 = 1 - p_1 - p_2 - p_3$.

In the implementation used here, we first generate a scenario k, and use $X(k)$ as the initial solution, where $X(k)$ is the optimal solution of the linear program solving the deterministic transportation problem corresponding to the scenario k. The solution is obtained by (independent) random realizations of the random variables b_j. Solutions with higher probability for sales realization are thus more likely to be generated.

Each heuristic was run 10 times for 3000 iterations. For all of the 10 randomly generated initial solutions we have allowed at most 60 percent of lost sales and overstocking effects realizations. The tabu search heuristic from 6 of the 10 initial solutions returns the result, corresponding to the scenario with total costs of 2736 Euro, therefore it seems likely that this solution is the best solution of this instance. Minimized are as well as expected overstocking effects as expected lost sales, consequently minimized are also the costs of distribution.

Tabu search algorithm has been programmed in the Python environment (the code was not optimized). Search runs were executed on an Intel Xeon E3-1230 v3 (8MCache, 3.3GHz). For the experimental case of 100 stores, 10 initially selected solutions and for each solution performed 3000 tabu search iterations, computation lasted roughly 10 seconds. Computation

speed was increased significantly compared to the previously used solution procedures (40 minutes). Thus the implementation that separately improved the distribution cost and the cost of risks dramatically decreases computational time. In addition, note that the solutions obtained are for comparable quality and it is even easier to adapt the method to optimize cost functions that are not a (simple or weighted) sum of the criteria.

5 CONCLUSION

In the paper, we have introduced and tested a robust optimization approach for tactical planning of a retail supply chain product flow. We have considered a product, whose sales figures are independent from other products' sales figures. At first, we have randomly generated initial solutions (representing distribution plans with defined inventory levels and allocation of resources), having at most some predefined percent of supply risks realizations (robust conditions). Then we used a tabu search algorithm and seek for solutions with minimal supply risks (overstocking effects and lost sales). The solutions were further evaluated due to the distribution costs assessments. It was shown that excluded time prohibitive linear programming from the iterative improvement solution procedure enormously speeds up the computations. Beside improvements regarding the computational speed and a better definition of an appropriate cost function, there are also many other interesting research avenues for future research work. First of all, we shall try to upgrade the model so it will be capable to deal with the products distribution plans, whose sale figures are dependent. Naturally, the optimization problem will be harder, therefore we might also consider possible improvements of the heuristic solution procedure. For the future we also left out the natural extensions that would include a dynamical self-adapting mechanism. Here the comparative model for the operative planning is one of the interesting future research avenues, where the distribution quantities are going to be defined with a difference between an actual inventory and the pre-defined maximum level of a certain inventory (result from the tactical planning).

References

[1] Anholcer, M. (2016). Bi-criteria Stochastic Generalized Transportation Problem: Expected Cost and Risk Minimization. Multiple Criteria Decision Making online publication, 1 November 2016, doi:10.22367/mcdm.2016.11.01.

[2] Franca, R.B., Jones E.C., Richards, C.N. and Carlson, J.P. (2010). Multi-objective stochastic supply chain modelling to evaluate tradeoffs between profit and quality. International Journal of Production Economics, 127, 292–299.

[3] Grossman, I.E. and Guillén-Gosálbez, G. (2010). Scope for the application of mathematical programming techniques in the synthesis and planning of a sustainable processes. Computers and Chemical Engineering 34(9), 1365–1376.

[4] Liu, S. and Papageorgiou, L.G. (2013). Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. Omega 41(2), 369–382.

[5] Sarimveis, H., Patrinos, P., Tarantilis, C.D. and Kiranoudis, C.T. (2008). Dynamic modelling and control of supply chain systems: A review. Computers and Operations Research, 35, 3530–3561.

[6] Vizinger, T. and Žerovnik, J. (2017). A stochastic model for better planning of a product flow in retail supply chains. Submitted to the Journal of the Operational Research Society in June 2017 (in review process).

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 6:
MRP and Related Systems Approach to Industrial Engineering and Services

MRP THEORY SUPPORTING TRADE-OFF BETWEEN INVESTMENTS IN COLLABORATIVE ROBOTS AND PRODUCTION IN FOREIGN COUNTRIES FOR A WATER PUMPS SUPPLY CHAINS

Daria Battini, Martina Calzavara, Fabio Sgarbossa and Alessandro Persona
University of Padova, Department of Management and Engineering, Vicenza, Italy
E-mail: daria.battini@unipd.it

Abstract: In Italy, the number of people aged 65 and older is growing. To keep a sustainable pension system, the retirement age of industrial workers is rising. While their functional capacities are decreasing, better ergonomic solutions are required, preserving their productivity level. Collaborative robots and smart workstations could be a solution. The second solution is to move a part of production to foreign countries, e.g. Hungary and Romania. In the article, we are investigating the water pumps supply chain and the trade-off between two solutions: to invest in the collaborative robots or to move the production plants to Hungary or Romania, where the human resources are available. The paper presents the study of the impact of these two solutions on the Net Present Value (NPV) on the base of the extended MRP (EMRP) model. Investments in smart workstations seem the better solution.

1 INTRODUCTION

Working populations in European Member States are ageing and cities are shrinking [14, 7]. In short, if most people continue to retire at around 60 years of age, the European labour force will shrink by around three million per year over the period 2020 to 2035 as reported by EU-OSHA [13]. As written in [13], the EU had set itself strategy objectives to increase labour market participation of older workers. However, practical limits arise, e.g. health problems and acquired financial security (also dependent on available reverse mortgages [4, 5, 6] regarding different wishes for time management). There is a need for a new and more comprehensive policy design to counter the shortage of workers in the future, particularly to keep workers in employment for longer. This cannot be done by cutting pension entitlements to force people to work for longer, but must be thought of as a systemic change, *"thinking about work from a life-course perspective"* [13] the improvements of poor workplace ergonomics which influence productivity is the best approach to solve these problems [1, 21]. According to one of the four principles of Industry 4.0 there is a need to support humans by conducting a range of tasks that are unpleasant, too exhausting, or unsafe thanks to new ergonomics-oriented equipments and collaborative robots. Consequently, currently there is a growing demand of applications with arm-based robots [12]. Robots will become a safe help to reduce fatigue and cognitive stress during the human work, without substituting expert workers. To invest in such systems or to de-localize the production system in East of EU is the question to which we shall answer based on extended MRP model.

2 THE BASIC MODEL

We shall consider an option that investments in collaborative robots increase the labour cost from $c_{L,i}(1+\alpha_{R,i})$ to $c_{L,i}(1+\alpha_{E,i})$, where $\alpha_{E,i} \cdot c_{L,i}$ are factored into the program of investments into ergonomics and $\alpha_{R,i} \cdot c_{L,i}$ are factored into the program of earlier retirement. There are two options of decisions: (a) the ergonomics could be improved enabling also seniors to work longer or (b) early retirement schemes are introduced [4, 5] and some activities are allocated in foreign country, where human resources are less expensive but perturbations on the roads

and border crossings could add additional uncertainties. Any of these decisions could influence the quality and quantity of production in the workplace and also influence the lead time in a supply chain [3]. The benefit to the total supply chain could be evaluated through using the Net Present Value (NPV) approach only, because investments and activities have different dynamics on the time horizon of the production and distribution in total supply chain. Our analysis is based on *MRP theory,* as found in the papers of Grubbström [15,17], first summarised at Storlien in 1997 [16], later being extended to the global supply chain [3], [10], [11] in which the location is also considered [9], [18], [19] [20], [22], including regional characteristics, such as the cost of labour and ageing of European population. According to the basic MRP theory [15], [16], [17], shortly described below, the *j*-th process is run on activity level (in node *j)* having intensity P_j, the volume of required inputs of item *i* is $h_{ij}P_j$ per time unit. The total of all inputs may then be collected into the column vector **HP**. The net production is determined as $(\mathbf{I} - \mathbf{H})\mathbf{P}$. In general, **P** is a time-varying vector-valued function. In MRP systems, *lead times* could be easily studied simultaneously in a total supply chain if using the Laplace Transforms methodology. If $P_j(t)$ is the rate of items *j* planned to be completed at time *t*, then the quantity $h_{ij}P_j(t)$ of items *i* need to be available for production or reloading in a time unit which is lead time τ_j in advance of time *t*, i.e. at time $(t - \tau_j)$, sent from the previous activity cell at time $(t - \tau_{ij} - \tau_i)$. While item *i* is assumed to be located previously at location *i* it will be available for activity *j* at location *j* before activity $P_j(t)$ starts, and it will need a certain time τ_{ij} to arrive there. For our needs, we consider an *assembly* system of water pumps. This process can be perturbed in nodes by reducing the functional capacities of aging workers or on roads because of custom duties on borders [3], [2], road works, accidents and other reasons, which could influence cascaded risk [8] The input requirements are given as transforms in the perturbed *generalised transportation-production-input matrix*, which is denoted $\breve{\mathbf{H}}'(s)$. Thus, the requirements for the production plan $\tilde{\mathbf{P}}(s)$ written as $\breve{\mathbf{H}}'(s)\tilde{\mathbf{P}}(s)$ are specified in the frequency domain where the net production $\tilde{\mathbf{x}}(s)$ will conveniently be written as follows:

$$\breve{\mathbf{H}}'(s) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ h_{21}e^{s\tau_{21}(1+\delta_{21})} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1}e^{s\tau_{n1}(1+\delta_{n1})} & h_{n2}e^{s\tau_{n2}(1+\delta_{n2})} & \cdots & 0 \end{bmatrix} \begin{bmatrix} e^{s\tau_1(1+\delta_{\tau_1})} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_n(1+\delta_{\tau_n})} \end{bmatrix}$$

$$\tilde{\mathbf{x}}(s) = (\mathbf{I} - \breve{\mathbf{H}}'(s))\tilde{\mathbf{P}}(s) \tag{1}$$

For *cyclical processes*, which repeat themselves in constant time intervals Γ_j, $j = 1, 2, \ldots, n$, the plan $\tilde{\mathbf{P}}(s)$ is written [10],[11], [17] and here extended because of ageing to the perturbed

$$\tilde{\mathbf{P}}(s) = \tilde{\mathbf{t}}'(s)\tilde{\Gamma}'(s)\hat{\mathbf{P}} = diag(e^{\frac{-st_1(1+\sigma_1)}{1-e^{-s(\Gamma_1+\Delta\Gamma_1)}}}, \ldots, e^{\frac{-st_n(1+\sigma_n)}{1-e^{-s(\Gamma n+\Delta\Gamma_n)}}}) \cdot \hat{\mathbf{P}} \tag{2}$$

Here $\tilde{\mathbf{t}}'(s)\tilde{\Gamma}'(s)$ is the product of perturbed matrices of starting moment of activities $\tilde{\mathbf{t}}(s)$ and total cycle $\tilde{\Gamma}'(s)$, $\hat{\mathbf{P}}$ is a vector of constants: for instance, batch sizes to be produced in each process during one of the periods $\Gamma'_j = \Gamma_j + \Delta\Gamma_j$, $j = 1, 2, \ldots, n$. Furthermore, in the above

equation (2), $t_j(1+\sigma_j)$, $j = 1, 2, \ldots, n$, are the points in time when the first of each respective cycle starts being perturbed for $t_j\sigma_j$ because of declining functional capacities of workers at j-th nodes. We shall use the approximation for perturbed expression describing $\tilde{\mathbf{P}}(s) = \tilde{\mathbf{t}}(s)'\tilde{\mathbf{T}}(s)'\hat{\mathbf{P}}$:

$$\tilde{\mathbf{P}}(s) = \left[\frac{e^{-st_1(1+\sigma_1)}\hat{P}_1}{1-e^{-s(\Gamma_1+\Delta\Gamma_1)}} \quad \cdots \quad \frac{e^{-st_n(1+\sigma_n)}\hat{P}_n}{1-e^{-s(\Gamma_n+\Delta\Gamma_n)}} \right]^T \square \frac{1}{s}\left[\frac{e^{-st_1(1+\sigma_1)}\hat{P}_1}{\Gamma_1+\Delta\Gamma_1} \quad \cdots \quad \frac{e^{-st_n(1+\sigma_n)}\hat{P}_n}{\Gamma_n+\Delta\Gamma_n} \right]^T \qquad (3)$$

Let us collect the economic values of items into a perturbed price vector **p**, which is a row vector, as follows:

$$\mathbf{p} = \mathbf{p}^1 + \Delta\mathbf{p} = \left[p_1^1(1+\delta_1), p_2^1(1+\delta_2), \ldots, p_n^1(1+\delta_n) \right] \qquad (4)$$

where δ_i is the relative reduction of prices of item i $(\delta_i \leq 0)$, because the worker in the production unit i is no longer able to assure the best quality of products, due to his advanced age or because we have replaced this worker with a new one without needed specific skills on distant location. By investments in better ergonomic conditions in workplaces, the reduction of negative δ_i could be achieved. The choice between better products with higher price and lower transportation costs and reduced lead-times on one hand and not to invest in improvements of ergonomic conditions, on the other hand and retired earlier, can be achieved based on the NPV evaluation. According to the *Net Present Value Theorem*, the NPV of the cash flow is obtained by replacing the complex frequency s with the continuous interest rate ρ, for example for NPV of production NPV_{prod}, ordering and fix costs per cycle NPV_{ord}, and transportation NPV_{tr}, the total NPV: NPV_{tot} for perturbed system is:

$$\text{NPV}_{tot} = \text{NPV}_{prod} + NPV_{ord} + NPV_{tr} = \sum_{i=1}^{n}\left(p_i(1+\delta_i) \right)\tilde{x}_i(\rho) -$$

$$-\left\{ \frac{\mathbf{K}+\mathbf{E}^T\mathbf{\Pi}}{\rho} \right\}\left[\frac{e^{-(\rho)t_1(1+\sigma_1)}\hat{P}_1}{\Gamma_1+\Delta\Gamma_1} \quad \cdots \quad \frac{e^{-(\rho)t_n(1+\sigma_n)}\hat{P}_n}{\Gamma_n+\Delta\Gamma_n} \right]^T \qquad (5)$$

In (5) $\mathbf{K}$ is a row vector of the setup costs and other fixed costs of cycle also costs of annuities for other investments except investments in the collaborative robots. In (5) $\mathbf{\Pi}$ is transportation matrix:

$$\mathbf{\Pi} = \begin{bmatrix} 0 & 0 & \ldots & \ldots & 0 \\ h_{2,1}b_{2,1}\tau_{2,1} & 0 & \ldots & \ldots & 0 \\ \vdots & h_{3,2}b_{3,2}\tau_{3,2} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ h_{n,1}b_{n,1}\tau_{n,1} & \ldots & \ldots & h_{n,n-1}b_{n,n-1}\tau_{n,n-1} & 0 \end{bmatrix} \qquad (6)$$

Where b_{ij} is cost of transportation of one unit of item i one hour on the rout from i to j so that the cost of transportation one item all along this rout is $b_{ij}\tau_{ij}$. Considering a long-term profit of a supply chain and its NPV we also need to include the direct costs of labour including early retirement premiums and investments in robots into account. The total NPV (NPV_{tot}) is reduced for the payments to the labour in individual places of activities with production or distribution intensity P_i, including the part of earnings which goes to the occupational pension funds, and the amount which goes to the annuity stream of investment

in ergonomics to support workers, like investments in robots and other new equipment which improve ergonomics of workers. Here the investments to the robots will be considered as a part of earnings (tied to work) and not to fix costs.

3 THE COSTS OF HUMAN RESOURCES AND COLLABORATIVE ROBOTS, INFLUENCING NPV

The early retirement age, as determined in an occupational pension scheme and trade-off between additional pension and investments in ergonomics, could be achieved by increasing the contributions from gross earnings $c_{L,i}$ to the extra occupational pension schemes $\alpha_{R,i} c_{L,i}$ for workers at the activity cell i and/or to the part of income which goes to the annuities for robots and other investments in ergonomics $\alpha_{E,i} c_{L,i}$, calculated relatively to the costs of labour. If early retirement age would be decision, a part of production should go in foreign country, which means additional costs of transportation and exposure to higher risks. Thus, if the labour cost would increase from $c_{L,i}(1+\alpha_{R,i})$ to $c_{L,i}(1+\alpha_{E,i})$, where $\alpha_{E,i} c_{L,i}$ is factored into the program of collaborative robots and other investments into ergonomics, the ergonomics could be improved enabling seniors to work longer, but $\alpha_{R,i} c_{L,i}$ may lower the retirement age of a worker at the workplace i. The costs of work which include collaborative robots in the total NPV could be written as $c_{L,i}(1+\alpha_{E,i})L_i P_i$. Here L_i is the number of employed at i. Also $c_{L,i}\alpha_{E,i}$ includes maintenances and depreciation costs per cycle, where $c_{L,i}(1+\alpha_{R,i})$ is the cost of one unit of work, which also includes the part of gross earnings $\alpha_{R,i}$ that is sent to the occupational pension fund, but this costs are connected with costs of labour in distant country and transportation costs to that locations, to be able to retire earlier. The amount of annuity for investments into the ergonomic improvements, including investments in robots. We need to write the NPV of the cost of labour and annuities for improvement of ergonomic environment, keeping production at home with negligible transportation costs (a), and early retirement and export the production in foreign country with substantial additional transportation costs (b).

Therefore, the NPV of the profit of total supply chain in case of local production in Italy (a) and in case of extension of the supply chain to foreign country (b) is the following:

$$\text{NPV}_{profit}(a) = \text{NPV}_{prod} - NPV(E) - NPV_{ord} = \sum_{i=1}^{n}\left(p_i(1+\delta_i)\right)\tilde{x}_i(\rho) - \frac{1}{\rho}\sum_{i=1}^{n} c_{L,i}(1+\alpha_{E,i})L_i \frac{e^{-(\rho)t_i(1+\sigma_i)}\hat{P}_i}{\Gamma_i + \Delta\Gamma_i}$$

$$-\left\{\frac{\mathbf{K}}{\rho}\right\}\left[\frac{e^{-(\rho)t_1(1+\sigma_1)}\hat{P}_1}{\Gamma_1 + \Delta\Gamma_1} \quad \cdots \quad \frac{e^{-(\rho)t_n(1+\sigma_n)}\hat{P}_n}{\Gamma_n + \Delta\Gamma_n}\right]^T$$

$$\text{NPV}_{profit}(b) = \text{NPV}_{prod} - NPV(R) - NPV_{ord} - NPV_{tr} = \sum_{i=1}^{n}\left(p_i(1+\delta_i)\right)\tilde{x}_i(\rho) -$$

$$-\frac{1}{\rho}\sum_{i=1}^{n}\left\{\left[c_{L,i}(1+\alpha_{R,i})L_i + c'_{L,i}L'_i\right]\frac{e^{-(\rho)t_i(1+\sigma_i)}\hat{P}_i}{\Gamma_i + \Delta\Gamma_i}\right\} - \left\{\frac{\mathbf{K}+\mathbf{E}^T\mathbf{\Pi}}{\rho}\right\}\left[\frac{e^{-(\rho)t_1(1+\sigma_1)}\hat{P}_1}{\Gamma_1 + \Delta\Gamma_1} \quad \cdots \quad \frac{e^{-(\rho)t_n(1+\sigma_n)}\hat{P}_n}{\Gamma_n + \Delta\Gamma_n}\right]^T$$

$$\tag{7}$$

The owners of a supply chain will choose the chain for which: $\max\left\{\text{NPV}_{profit}(a), \text{NPV}_{profit}(b)\right\}$ is achieved.

4 THE NUMERICAL EXAMPLE

The numerical example is getting inspiration from an Italian manufacturer of water pumps. The product design and production is all made in Italy, but because of ageing of human resources and shortage of new workers they have two options:
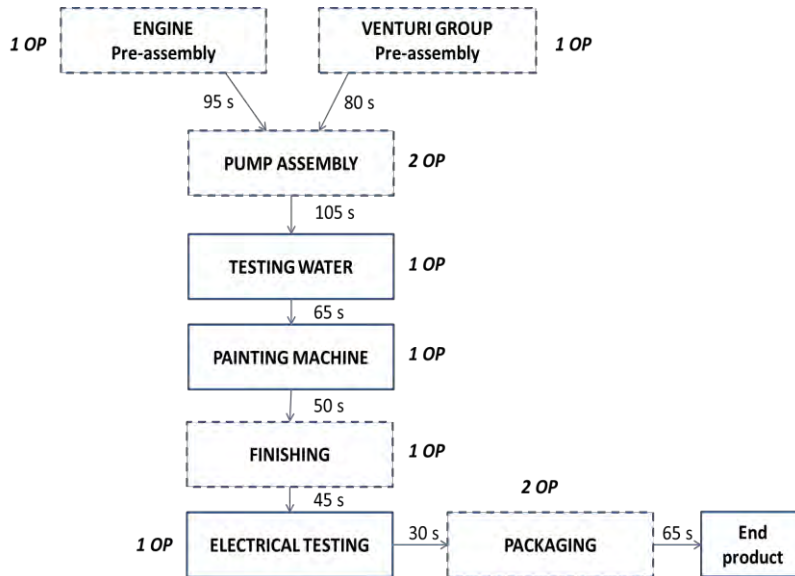
(a) to buy the collaborative robots and extend their retirement age or

(b) to retire them earlier, to pay in early retirement schemes and to open new factory or in Hungary (close to Budapest: 11 hours distance, 0.16 Euros/item) or in Bulgaria (close to Sofia: 17 hours, 0.22 Euros/item).

Figure 1: The production graph with process times and workers number

The norm-production τ_i per item is given in fig. 1. The raw material is coming from Italy and in the case of establishing production units in Hungary or Bulgaria, two semi-products would be sent there and final product would be transported back. Environmental restrictions and pollution fees are not considered in our calculations (7). The manpower cost in the assembly line in Italy are: 26,5 €/h, additional to pension fund for earlier retirement 20% and the investments in 1 collaborative robot could range from 20,000 to 60,000 €/robot which requires 20% yearly depreciation and 10% of maintenance costs which reduce the production time for 20% or more. The manpower costs are 3,1 €/h in Bulgaria and 7 €/h in Hungary. The product price to the final costumer is assumed equal to 250 euros per pump. Following the equation (7) we can see that in case of 4% interest rate or lower it is better to keep the production in Italy and to buy the collaborative robots for each human operator involved in each station highlighted by dashed line squares in Fig. 1. Even in the case that investments in production systems in foreign country would be covered by European structural funds. The new robot will provide the needed flexibility of the interfaces between parts and workstations.

5 CONCLUSION

As far as the authors know this is the first attempt that Extended MRP model is applied to demonstrate the trade-off between investments in ergonomics and collaborative tools or production in foreign country, where the costs of human resources are lower and structural funds are available for invest in production in less developed countries (where cost of human resources is not so high and resources are still available). To solve the problems of declining functional capacities of ageing population, European companies have option to invest in collaborative robots, to provide a better working environment, or to export the part of

production on East. In this specific case study, the best solution is to buy them. Since environmental taxes and fees are not included in calculation this will be the subject of our further research.

References

[1] Battini D., Faccio M., Persona A., Sgarbossa F. 2011, New methodological framework to improve productivity and ergonomics in assembly system design, Int. J. of Industrial Ergonomics, 41, 30–42.

[2] Bogataj, D., Bogataj, M., 2007. Measuring the supply chain risk and vulnerability in frequency space. Int. J .Prod. Econ., 108(1-2): 291-301

[3] Bogataj, D., Bogataj, M. , 2011. The role of free economic zones in global supply chains - a case of reverse logistics. Int. J .Prod. Econ., 131(1): 365-371

[4] Bogataj, D., 2013. Pensions and home ownership in the welfare mix for older persons. In: ZADNIK STIRN, L. (ed.), et al. SOR '13 proceedings, Ljubljana: Slovenian Society Informatika, SOR, 2013, pp. 281-286

[5] Bogataj, D., Vodopivec, R., Bogataj, M., 2013. The extended MRP model for the evaluation and financing of superannuation schemes in a supply chain. Technological and economic development of economy, 19(S1): S119-S133

[6] Bogataj, D., Ros-McDonnell, D., Bogataj, M. 2015. Reverse mortgage schemes financing urban dynamics using the multiple decrement approach. Springer Proceedings in Mathematics & Statistics, vol. 135. pp. 27-47

[7] Bogataj, D., Ros-McDonnell, D., Bogataj, 2016a, Management, financing and taxation of housing stock in the shrinking cities of aging societies. Int..J. Prod. Econ., 181(A):. 2-13

[8] Bogataj, D., Aver, B., Bogataj, M., 2016b. Supply chain risk at simultaneous robust perturbations. Int..J. Prod. Econ, 181(A): 68-78

[9] Bogataj, M., Grubbström, R.W., Bogataj, L. 2011. Efficient location of industrial activity cells in a global supply chain, Int..J. Prod. Econ., 133(1): 243-250.

[10] Bogataj, M., Grubbström, R.W. 2012. On the representation of timing for different structures within MRP theory. Int..J. Prod. Econ.140(2): 749-755

[11] Bogataj, M., Grubbström, R.W. 2013. Transportation delays in reverse logistics, Int..J.Prod.Econ., 143(2) : 395-402

[12] Estevez, E., Garcia, AS Garcia, J.G., Ortega, J.G. 2017. An UML based approach for designing and coding automatically robotic arm platforms. Revista Iberoamericana de automatica e informatica industrial 14(1): 82-93

[13] EU-OSHA, Cedefop, Eurofound and EIGE, 2017. Joint report on Towards age-friendly work in Europe. Publications Office of the European Union, Luxembourg.

[14] European Commission, 2015. The 2015 Ageing Report. Brussels: DGECFIN

[15] Grubbström, R.W., 1967. On the application of the Laplace transform to certain economic problems, Manag. Sci., 13: 558-567

[16] Grubbström, R.W., Bogataj, L., 1998. (Eds.), Input–Output Analysis and Laplace Transforms in Material Requirements Planning, Storlien, 1997. FPP, Portorož, 1998.

[17] Grubbström, R. W., 2007. Transform methodology applied to some inventory problems. Z. Betriebswirtsch, 77 (3): 297-324

[18] Kovačić, D., Bogataj, M. 2013. Reverse logistics facility location using cyclical model of extended MRP theory. Central European Journal of Operations Research, 21(1): 41-57

[19] Kovačić, D., Usenik, J., Bogataj, M. 2017. Optimal decisions on investments in urban energy cogeneration plants. Int.j.prod.econ., 183: 583-595

[20] Kovačić, D., Hontoria, E., Ros McDonnell, L., Bogataj, M. 2015, Location and lead-time perturbations in multi-level assembly systems of perishable goods in Spanish baby food logistics. Central European Journal of Operations Research, 23(3): 607-623.

[21] Otto A., Scholl A. 2011. Incorporating ergonomic risks into assembly line balancing, European Journal of Operational Research, 212, 277–286

[22] Usenik, J. Bogataj, M. 2005. A fuzzy set approach for a location-inventory model. Transportation planning and technology, 28(6): 447-464

AGE MANAGEMENT OF HUMAN RESOURCES

David Bogataj
Department of Management and Engineering, University of Padua,
Stradella S. Nicola, 3, 36100 Vicenza VI, Italy
david.bogataj@unipd.it

Marija Bogataj
CERRISK - Zavod INRISK, Vrtača 9, 1000 Ljubljana, Slovenia
marija.bogataj@guest.arnes.si

Abstract: The focus of this paper is the age management of total supply chains through identifying how functional decline of ageing workforce influences time delays that could appear simultaneously. We highlight how introduction of a proper age management of human resources could support mitigating of a supply chain risk using the Net Present Value (NPV) approach in Extended Material Requirements Planning, where the simultaneous perturbations of timing influence the NPV. It is explained how to define a better structure of costs for human resources based on the known trajectory of functional decline of workers. The method for evaluation of impact of the structure of contributions to national and occupational pension schemes, health insurance schemes, long-term care insurance schemes and investments in workplace ergonomics is given. Established actuarial principles are used to link future age-related liabilities with current payroll contributions. We show how the introduction of an active age management in the decision support models of supply chains is essential to achieve sustainability and competitiveness.

Keywords: supply chain, human resources, ageing, insurance, multiple decrement model, cobots, age management

1 INTRODUCTION

The ageing has implications for employment, working conditions, living standards and welfare of industrial workers. Therefore, age management is becoming key area of activities by which human resources are managed within organisations with an explicit focus on ageing via public policy or collective bargaining (Walker, 2005) [10]. Europe 2020 strategy states as employment target: "75% of people aged 20–64 to be in work". Challenges are highlighted in recently published EUROFUND study *Towards age-friendly work in Europe* [5]. The retirement age in public pension schemes will reach 70 years by 2060 in many EU member states. Many industrial and logistics workers will not be able to work to the increased retirement age. They require better workplace ergonomics or possibility of early retirement. The expected healthy life years at birth is 17 years shorter than the overall life expectancy for men and 22 years for women. In this time window workers receive pensions and many persons are dependent on the help of others [7].

The challenge of an ageing workforce brings various aspects that need to be addressed to achieve sustainability and competitiveness of supply chains in an environment of aging and shrinking workforce. Introducing smart process technologies such as smart manufacturing cells, supporting robots (cobots) and other improvements in worker's ergonomics can help workers to work longer, while contributions to pension, health, and long term care (LTC) funds improve the life of workers after retirement. To achieve the sustainability and competitiveness of a supply chain in aging societies, the managers of a chain must ensure the availability of adequate funds for age management in combination with higher safety stock.

In the presented methodology, we show how to measure the influence of aging on the perturbation of NPV in production and logistics, as results of the simultaneous perturbations of timing of financial flows, information flows, flows of items and market perturbations. These perturbations can be better evaluated simultaneously through the Input/Output analysis,

Laplace transforms and the NPV expression which enables us to forecast and control the physical and financial flows simultaneously. In each activity cell of a supply chain, the functional decline of workers is different. For each work place exist the optimal retirement age. Retirement age at workplace can increase in the case of investment in workplace ergonomics. The trajectory of functional decline of workers and investments in the workplace ergonomics should be carefully studied. Functional decline of workers influence parameters of pension schemes, health and long term care insurance schemes (see Fig. 1).
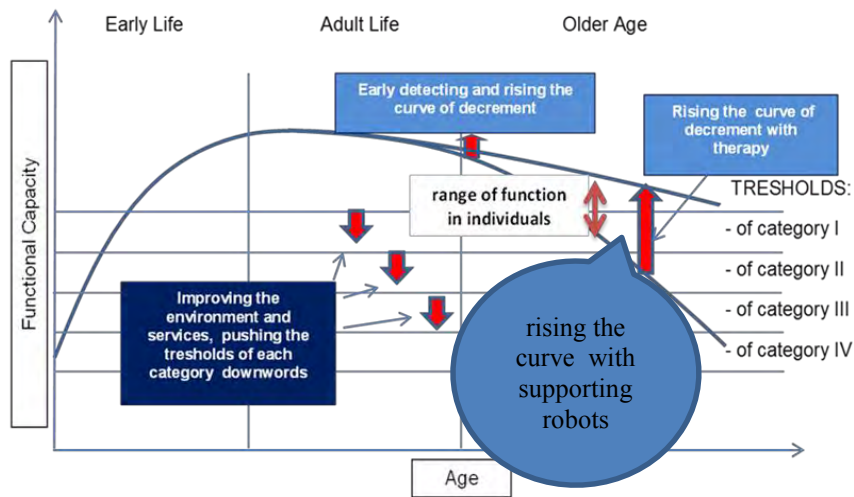
Figure 1: The dynamics of functional capacities

Physically demanding work influences functional decline of workers [5,7], therefore contributions for occupational pension and LTC are dependent on the type of workplace [7]. Due to functional decline, many workers will not be able to work till increased retirement age. Organisations can provide early retirement occupational pension from the moment when worker cannot perform his work and achieve required productivity to the moment when workers is entitled to public pension. To be able to work longer, ageing workforce requires support at their workplace by the development of smart supporting environment, like smart production cells, in some cases also by investing in collaborative robots (cobots), like explained in Battini et al. [1, 2, 3] or Sgarbossa et al. [9]. The optimal level of social contributions and investments in ergonomics could be determined by decision support models based on MRP theory and actuarial mathematics. National an occupational contribution rates are input to the structure of the labour costs in the extended MRP model, where one of the choices in this structure is also an investment in ergonomics.

2 THE IMPACT OF RETIREMENT AND LTC POLICIES ON NPV IN SUPPLY CHAINS

2.1 Lead time perturbations of the Net Present Value

These policies will affect quality and prices **p** of items produced by supply chain**,** as well as the earlier retirement of workers. Regarding fig.1 the matrices $\widetilde{\boldsymbol{\pi}}, \widetilde{\boldsymbol{\tau}}$, influence the added value all according to the equation of the NPV of profit in a supply chain, as developed by Bogataj et all.[4] and further developed by Battini et all. [1]. The profit depends on time delays, therefore on retirement plan, which determine yearly contribution to retirement fund for pension**,** $c_L. \alpha_R$, health care**,** $c_L. \alpha_H$, and LTC, $c_L. \alpha_L$ as well as on contributions for investments in ergonomics tied to the costs of labour $c_L. \alpha_E$: $\widetilde{\boldsymbol{\tau}} = \widetilde{\boldsymbol{\tau}}(c_L, \alpha_R, , \alpha_H, \alpha_L, \alpha_E)$, through

the impact on: (a) the NPV of selling the products, $NPV_{prod} = \text{NPV}(\tilde{\tau}(c_L, \alpha_R,, \alpha_H, \alpha_L, \alpha_E))$; (b) the NPV associated with the length of the cycle $NPV_{ord} = NPV_{ord}(\sigma_t, \Delta\Gamma)$ affected by the functional capacities and ergonomics of workers, therefore on the retirement, health care, LTC schemes and investments in ergonomics, (c) the NPV of transportation between activity cells NPV_{tr} again depending on length of the cycle, but especially on the perturbations in the transportation matrix: $NPV_{tr} = NPV_{tr}(\sigma_t, \Delta\Gamma, \tilde{\Pi})$. We can remark that the functional capacities of workers, labour costs and investments in ergonomics through contributions to retirement schemes, health and LTC insurance as well as investments in ergonomics influence NPV of the profit in a supply chain as written in Battini et al.[1, equation (5)]:

$$\text{NPV}_{tot} = \text{NPV}_{prod} + NPV_{ord} + NPV_{tr} = \sum_{i=1}^{n}\Big(p_i\,(1+\delta_i)\Big)\tilde{x}_i(\rho) -$$

$$-\left\{\frac{\mathbf{K} + \mathbf{E}^T\tilde{\Pi}}{\rho}\right\}\left[\frac{e^{-(\rho)t_1(1+\sigma_1)}\hat{P}_1}{\Gamma_1 + \Delta\Gamma_1} \quad \cdots \quad \frac{e^{-(\rho)t_n(1+\sigma_n)}\hat{P}_n}{\Gamma_n + \Delta\Gamma_n}\right]^T \qquad , \qquad (1.1)$$

Here the discount factor ρ consist of the continuous interest rate ρ_o and the growth of the flow in the system ω : $\rho = \rho_o - \omega$. Earlier retirement and applying ergonomics to the workplace by increasing investments in cobots, smart production cells and other environment influences (a) reduction of the potential for accidents which improve $\Delta\Gamma_i, \tau_i, \sigma_i$; (b) reduce the potential for injury and ill health, also improving $\Delta\Gamma_i, \tau_i, \sigma_i$ and (c) improving performance and productivity (increase the level of $\mathbf{p}$ and ω in (1.1)). The supply chain could be long-term sustainable if the following condition is achieved:

$$\sum_{i=1}^{n} c_{L,i}\left(1 + \alpha_{Ri} + \alpha_{Hi} + \alpha_{Li} + \alpha_{Ei}\right) L_i \tilde{x}_i(\rho) \leq NPV_{tot} \qquad (1.2)$$

Where L_i is the number of workers per time unit per item in the activity cell i. Which means that structure of $\alpha_{Ri} + \alpha_{Hi} + \alpha_{Li} + \alpha_{Ei}$, consisting of contribution rates regarding net salary, like contributions to pension (α_{Ri}), health (α_{Hi}), LTC(α_{Li}) funds and also investments in ergonomics (α_{Ei}) in total chains, when the net salary is $c_{L,i}$ for worker at *i-th* activity cell. This influences also on the level of losses which are results of lower quality of production $\sum_{i=1}^{n} p_i^1 \delta_i \tilde{x}_i(\rho)$ and losses as the results of perturbations in timing because of declining functional capacities of workers $\mathbf{K}\tilde{t}(\rho)\{\mathbf{I} - \tilde{r}(\rho)\}\tilde{\Gamma}(\rho)$, as follows from the paper of Bogataj et al [4]. Here $\tilde{r}(\rho)$ is the matrix of lead time perturbations.

2.2 Optimum level of annuities required for LTC – the case for Slovenian cohorts

In the papers of Kavšek and Bogataj [7] and Rogelj and Kavšek [8] the authors have developed a model of the actuarial NPV of the expenditures for LTC and examined trajectories of decrease of functional capacities (fig.1) for older people who have been working on physically demanding jobs and those who did not. They have developed a model for estimating the NPV of lifetime expenditure for LTC for those two groups of older workers. For the calculation of expenses, they used standardised prices for home care, assisted living and institutional care determined by the responsible ministry in 2016. They found that the NPV of the expected

expenditures for long-term care for worked in workplaces where hard labour is required, is equal to 181.555,36 € and for normal work is equal to 127.147,83 €.

3 TRANSITIONS MATRIX DUE TO DECLINING FUNCTIONAL CAPACITY

3.1 Multiple decrements model

We can differentiate national and occupational age-related expenditures determined in national pension schemes, national health insurance schemes and national LTC insurance schemes, and occupational age related expenditures which determine occupational pension schemes, occupational health insurance schemes and occupational long term care insurance schemes and investments in ergonomics. Probability of transition $q_x^{(i,j)}$ is probability that worker will move from state of functional capacity i to state of functional capacity j in the period between age x and age x+1 while $p_x^{(i,i)}$ is probability that worker will stay in the same state of functional capacity. Therefore the following notation will be used: i determines state of functional capacity of worker and marks required category of ergonomic support or care that a worker needs ($i = 0$ to 6); j determines state of functional capacity and marks category of ergonomic support or care to which worker moves due to decreased functional capacity (move from state of functional capacity i to state of functional capacity j or death ($j = 1$ to 7); (i, j) refers to transitions from state i to the state j. Here $i = 0$ means worker is fully functional and does no need support of ergonomic environment, $i = 1$ denotes state of functional capacity where worker needs special improvements of ergonomic environment and $i = 2$ is state where functional capacity declines so that the worker is not able to perform his work and needs to enter early retirement. Further (i=3,4,5,6 are states of functional capacities in which worker is dependent on help of others and denotes intensity of care in categories I to IV and the final state (j=7) is death. $S_x^{*(i)}$ is measured number of workers of age x in the state i. We can calculate this probability of transitions by observing $M_x^{(i,j)}$ - the number of workers x years old who have moved from state i to state j in the examined period. $q_x^{(i,j)}$ is probability of transition for worker of age x from category of care i to category of care j.: $q_x^{(i,j)} = M_x^{(i,j)} / S_x^{*(i)}$. $S_x^{(i)}$ is the number of workers in the state i that are x years old at the beginning of the examined period.

Distribution of x years old workers regarding the type of working conditions and care can be described with the following vector S_x:

$$S_x = \begin{bmatrix} S_x^{(0)} & S_x^{(1)} & S_x^{(2)} & S_x^{(3)} & S_x^{(4)} & S_x^{(5)} & S_x^{(6)} & S_x^{(7)} \end{bmatrix} \tag{3.1}$$

Probabilities of the transition are given by matrix $\mathbf{Q_x}$:

$$\mathbf{Q_x} = \begin{bmatrix} p_x^{(0,0)} & q_x^{(0,1)} & q_x^{(0,2)} & \dots & \dots & q_x^{(0,7)} \\ 0 & p_x^{(1,1)} & q_x^{(1,2)} & \dots & \dots & q_x^{(1,7)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & p_x^{(5,5)} & q_x^{(5,6)} & q_x^{(5,7)} \\ 0 & 0 & 0 & 0 & p_x^{(6,6)} & q_x^{(6,7)} \\ 0 & 0 & 0 & 0 & 0 & p_x^{(7,7)} \end{bmatrix} \tag{3.2}$$

It shows the distribution of the cohort of workers who are x years old and some are dependent on the help of other, regarding the required intensity of support or care defined by states of functional capacities above. Distribution of workers x+1 years old by states of functional capacities regarding different categories of required support and care c_j due to functional decline can be calculated as:

$$S_{x+1} = S_x \mathbf{Q_x} = \begin{bmatrix} S_x^{(0)} & S_x^{(1)} & S_x^{(2)} & ... & ... & S_x^{(7)} \end{bmatrix} \cdot \mathbf{Q_x} \tag{3.3}$$

Yearly expenditures per capita for workers in different states of functional capacities are expressed by vector:

$$\mathbf{C} = \begin{bmatrix} 0 & c_E & c_0 & c_1 & c_2 & c_3 & c_4 & 0 \end{bmatrix}^T \tag{3.4}$$

The contribution rates for health, *LTC* and pensions in national *PAYG* systems are dependent on the demographic structure of the population. Where notation for yearly expenditures in the national health *NYHE*, long-term care *NYLE* and pension *NYRE* :

$$\alpha_{Hi}^{na}(t) = \frac{NYHE(t)}{m\overline{c}}, \quad \alpha_{Li}^{na}(t) = \frac{NYLE(t)}{m\overline{c}}, \quad \alpha_{Ri}^{na}(t) = \frac{NYRE(t)}{m \cdot \overline{c}}.$$ Here m is the number of insured

who are contributing in the national insurance schemes and $\overline{c}$ is the average wage. The total value of α_{Ki} is the sum of contribution rates to national and occupational schemes:
$\alpha_{Ki} = \alpha_{Ki}^{na} + \alpha_{Ki}^{o}; K \in \{H, L, R\}$.

a_x^{L+R+E} denotes actuarial present value of annuities for long term care, pensions and investments in ergonomics,

$$a_x^{L+R+E} = \sum_{i=0}^{l} (\mathbf{S}_{x+i-1} \cdot \mathbf{Q}_{x+i} \cdot \mathbf{C}^T) \cdot e^{-\rho i}; \quad l = 100 - x \Leftarrow H, L, R; \quad l = x + 5 \Leftarrow E \tag{3.5}$$

Here $\mathbf{C}$ is dependent on functional capacity of workforce. This functional capacity is dependent on occupational early retirement age, workplace ergonomics, and therefore, act as feedback on quality and consequently on vector $\mathbf{p}$ and timing, therefore on $\tilde{\tau} = \tilde{\tau}(c_L, \alpha_R, , \alpha_H, \alpha_L, \alpha_E)$, $\boldsymbol{\sigma}_t, \Delta\boldsymbol{\Gamma}$ and perturbations of $\tilde{\boldsymbol{\Pi}}$. Each supply chain has to consider the relationship between these parameters in each individual activity cell and on the branches of the graph of supply chain. The value of annuity stream at *i-th* activity cell for repayment of investments in ergonomics *E* if the repayment period is n years, is equal to:

$$\alpha_{E,i} = \frac{\dfrac{E_i}{\ddot{a}_{\overline{n}|}} = \dfrac{E_i}{\dfrac{v^n - 1}{v - 1}} = \dfrac{E_i}{\dfrac{e^{-\rho_o n} - 1}{e^{-\rho_o} - 1}}}{c_{L,i}} \tag{3.6}$$

Let us denote: rates of the administrative costs: γ_1 - in the period of payment of premiums and γ_2 - in the period of payment of early pensions as noted in actuarial mathematics, by: $_np_x$ - probability that person x years old will survive n years; $\ddot{a}_{x:\overline{m}|}$ - actuarial present value of n years term annuity for person x years old and $p_x^{ltcI}, p_x^{ltcII}, p_x^{ltcIII}, p_x^{ltcVI}$ -probabilities that worker, *x* years old, is in category of care I,II,III,IV. Therefore, for early retirement we write the contribution rate for early retirement occupational pension as:

$$\alpha_{Ro} = \alpha_R - \alpha_R^{na} = \frac{P_{x:\overline{n|}}}{c_{Li}} = \frac{\dfrac{c_{Li}\cdot(1+\gamma_2)\cdot {}_np_x\cdot \ddot{a}_{\overline{x:m|}}}{(1-\gamma_1)\cdot \ddot{a}_{\overline{x:n|}}}}{c_{Li}} = \frac{(1+\gamma_2)\cdot {}_np_x\cdot \ddot{a}_{\overline{x:m|}}}{(1-\gamma_1)\cdot \ddot{a}_{\overline{x:n|}}} = \frac{(1+\gamma_2)\cdot {}_np_x\cdot \sum_{i=0}^{m-1}{}_ip_x\cdot e^{-\rho i}}{(1-\gamma_1)\cdot \sum_{i=0}^{n-1}{}_ip_x\cdot e^{-\rho i}}, \qquad (3.7)$$

Contribution rate α_{Lo} for LTC annuity with the net present value $LTCa_{65}$ in occupational scheme where contributions are paid for 40 years and benefits are paid from age 65 is:

$$\alpha_{Lo} = \alpha_L - \alpha_L^{na} = \frac{{}_{40}p_{25}\cdot e^{-\rho 40}\cdot(1+\gamma_2)\cdot LTCa_{65}}{c_{Li}\cdot(1-\gamma_1)\cdot \sum_{i=0}^{n-1}{}_ip_x\cdot e^{-\rho i}} =$$

$$= \frac{{}_{40}p_{25}\cdot e^{-\rho 40}\cdot(1+\gamma_2)\cdot \sum_{j=0}^{100-65} {}_jp_{65}\cdot e^{-\rho_o i}\cdot\left(p_{65+j}^{(ltcI)}\cdot c_3 + p_{65+j}^{(ltcII)}\cdot c_4 + p_{65+j}^{(ltcIII)}\cdot c_5 + p_{65+j}^{(ltcIV)}\cdot c_6\right)}{c_{Li}\cdot(1-\gamma_1)\cdot \sum_{j=0}^{40-1}{}_jp_{25}\cdot e^{-\rho i}} \qquad (3.8)$$

The total contribution rate $\alpha_i = \alpha_{Ri} + \alpha_{Hi} + \alpha_{Li} + \alpha_{Ei}$ written in inequality *(1.2)* at continuous interest rate ρ_o is equal to:

$$\alpha_i = \frac{NYHE(t)}{m\overline{c}} + c_{L,i}\cdot E_i(e^{-\rho_o}-1)/(e^{-\rho_o n}-1) + \frac{(1+\gamma_2)\cdot {}_np_x\cdot \sum_{i=0}^{m-1}{}_ip_x\cdot e^{-\rho_o i}}{(1-\gamma_1)\cdot \sum_{i=0}^{n-1}{}_ip_x\cdot e^{-\rho_o i}} + \frac{NYRE(t)}{m\cdot\overline{c}} +$$

$$+ \frac{\sum_{j=0}^{100-65}{}_jp_{65}e^{-\rho_o j}\cdot\left(p_{65+j}^{(ltcI)}\cdot c_3 + p_{65+j}^{(ltcII)}\cdot c_4 + p_{x+j}^{(ltcIII)}\cdot c_5 + p_{x+j}^{(ltcIV)}\cdot c_6\right)}{c_{Li}\cdot(1-\gamma_1)\cdot \sum_{j=0}^{40-1}{}_jp_{25}\cdot e^{-\rho j}} + \frac{NYLE(t)}{m\cdot\overline{c}} \qquad (3.9)$$

4 A NUMERICAL EXAMPLE

Let us consider a worker whose net monthly salary is 1000 €. To keep his productivity on workplace he needs support of cobot at the age 60 and will retire at 65 according to the earlier retirement occupational pension scheme and later at 70 he will retire under the national PAYG system. From the cost of care, calculated in Kavšek and Bogataj [7], the yearly expenditures for care is given in vector $\mathbf{C} = \begin{bmatrix} 0 & E & |\mathbf{C}^*| \end{bmatrix}$, where sub-vector $\mathbf{C}^*$ can be written as: $\mathbf{C}^* = \begin{bmatrix} 0 & 8701.20 & 11685.60 & 15166.80 & 16732.80 & 0 \end{bmatrix}$. We shall later add investments in ergonomics and contributions to pension fund, and contribution for health insurance.

$$\mathbf{Q}_{65} = \begin{bmatrix} 0.93664 & 0.02662 & 0.01167 & 0.01172 & 0.00176 & 0.01159 \\ 0 & 0.83415 & 0.04867 & 0.05111 & 0.05355 & 0.01252 \\ 0 & 0 & 0.9035 & 0.03993 & 0.04405 & 0.01252 \\ 0 & 0 & 0 & 0.76037 & 0.22711 & 0.01252 \\ 0 & 0 & 0 & 0 & 0.8344 & 0.1656 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{Q}_{95} = \begin{bmatrix} 0.80664 & 0.15662 & 0.01167 & 0.01172 & 0.00176 & 0.01159 \\ 0 & 0.83415 & 0.04867 & 0.05111 & 0.05355 & 0.01252 \\ 0 & 0 & 0.9035 & 0.03993 & 0.04405 & 0.01252 \\ 0 & 0 & 0 & 0.76037 & 0.22711 & 0.01252 \\ 0 & 0 & 0 & 0 & 0.8344 & 0.1656 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

We have created 7 transition matrices for the ages 65 to 95, taking 5-years cohorts and got the values $\mathbf{Q}_x$. Following the average functional decline of the human resources and the impact of the physically demanding work on the functional decline, we calculated the required participation to the LTC fund. At the continuous interest rate $\rho_o = 0.006$ the actuarial net present value $LTCa_{65}$ for minimum requirements at the adequate long term care according to the Slovenian care standards and cost structure is 121348 € for the man of 65 years old when he retires and receives early retirement occupational pension. To these costs we add annuities for investments in ergonomics, and also yearly health and pension costs. In case of 25,000 € investment in a cobot for worker in his 60s, working additionally 5 years, calculating also 5

years repayment period for cobot which is the same as depreciation period the annuity is 5060 € which is 24% of the average monthly salary. Therefore the net present value of investments in robots and participation to LTC is together: NPV_a=146348 €, without contributions to health and pensions. These two additional contributions can be added easily as actuarial present value calculated in early retirement scheme and additionally for PAYG pension and health insurance schemes, calculated for PAYG systems as an average on the national level (3.7).

$$\alpha_{Ri}^{na}(t) = \frac{NYRE(t)}{m \cdot \bar{c}} = 36.37\%; \quad \alpha_{Ri}^{o} = 10.96\% ; \quad \alpha_{Lo} = 2,84\% ; \quad \alpha_{Hi}^{na}(t) = \frac{NYHE(t)}{m\bar{c}} = 16.5\%$$

$\alpha_{Hi}^{na} = 42,17\%$. It means that total age related expenses for person who has monthly 1000 € of net pension is $\alpha_i = 108.84\%$ of net salary, it means 1088.40€.

5 CONCLUSIONS AND PLAN FOR FURTHER RESEARCH

We have presented how the introduction of a proper age management of human resources could reduce time delays and therefore support mitigating of a supply chain risk using the Net Present Value approach in Extended MRP model. Therefore the contributions to ergonomics, occupational pension schemes for earlier retirements and LTC schemes should be developed. The numerical example for Slovenian worker with an average net salary shows that the contributions to all these funds require more than 100% of net salary to achieve sustainable supply chains and an adequate ergonomics, pensions, LTC and health insurance. In the next 40 years, this percentage will rise with the further aging of the population. Therefore this value is so high that the further research needs to include these contributions to the evaluation of costs of human resources including investments in smart production cells and cobots.

References
[1] Battini, D., Glock, C.H., Persona, A., Sgarbossa, F. 2015. Ergo-Lot-Sizing: Considering Ergonomics in Lot-Sizing Decisions. IFAC-PapersOnLine, 48(3): 326-331
[2] Battini, D., Calzavara, M., Persona, A., Sgarbossa, F. 2015. Linking human availability and ergonomics parameters in order-picking systems IFAC-PapersOnLine, 48(3): 345-350
[3] Battini D., Calzavara, M., Sgarbossa F., Persona A. 2017, MRP theory supporting trade-off between investments in collaborative robots and production in foreign countries for a water pumps supply chains, In: Zadnik Stirn, L. (ed.), et al. SOR '17 proceedings, Ljubljana: SDI-SOR, In Press.
[4] Bogataj, D., Battini D., Calzavara, M., Persona A. 2017. Investments in workplace ergonomics from the supply chain approach. Proceedings 24th International Conference on Production Research, DEStech Publications, In Press.
[5] Dubois, HJ, Jean-Marie W., Mathijn Vermeylen, G. 2017. Towards age-friendly work in Europe: a life-course perspective on work and ageing from EU Agencies, EU-OSHA Cedefop Eurofound EIGE.
[6] Cutler, J. 2014. Managing the Impact of Long-Term Care Needs and Expense on Retirement Security Monograph, Illinois: Society of Actuaries.
[7] Kavšek, M. Bogataj, D. 2017. Towards new quality standards of long-term care in Slovenia. Revija za univerzalno odličnost, 6(1): 11-24.
[8] Rogelj, V., Kavšek, M. 2017. Contributions to The Long-Term Care insurance fund for workers who hold physically demanding and labour-intensive jobs in supply chains. SOR'17 Proceedings, In press.
[9] Sgarbossa, F. Battini, D., Persona, A., Vizentin, V. 2016. Including Ergonomics Aspects into Mixed-Model Assembly Line Balancing Problem. In Goonetilleke, R., Karwowski, W. (Eds.), Advances in Physical Ergonomics and Human Factors, Proceedings of the AHFE 2016, Florida. USA: Springer Books.
[10] Walker, A. 2005. 'The emergence of age management in Europe'. International Journal of Organisational Behaviour, Volume 10(1), 685-697.

MEASUREMENT OF TEMPERATURE AND HUMIDITY FOR FURTHER DEVELOPMENT OF SMART COLD SUPPLY CHAINS

Domen Hudoklin
University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia
E-mail: domen.hudoklin@fe.uni-lj.si

Abstract: For better cold supply chain management advances in technology are needed for temperature and humidity control. At the sensory level, temperature control is studied to be complemented by humidity measurements. This paper presents a tool for measuring the impact of precise humidity reports on the management of cold supply chains. The evaluation of these new technologies is proposed through a Net Present Value (NPV) approach of extended material requirements planning models. The paper discusses on how these innovative technologies could contribute to an increase in the NPV of activities in the supply chains of perishable goods in general and how to measure this impact.

Keywords: input-output analysis, supply chain control; wireless sensors; accurate environment monitoring; humidity monitoring, time delay; perishability; extended MRP model

1 INTRODUCTION

While the globalization has made the relative distances between world regions much smaller, the freight can more likely be damaged in one of the complex transport operations involved in otherwise greater physical separations. For a range of goods labelled as perishables, particularly drug and food, their quality degrades with time due to unavoidable chemical reactions. The rate of these reactions can be mostly mitigated with lower temperatures and proper humidity. To ensure that cargo does not become damaged or compromised throughout this process, businesses in the pharmaceutical, medical and food industries are increasingly relying on the smart cold supply chain (SCSC), which involves thermal and refrigerated packaging methods and the logistical planning to protect the integrity of these shipments from field to forks. Analyzing services to the ageing European population, also from the supply chain perspective [7, 9] it is expected that development of smart homes will increase and the developed domotics will enable control of such cargo also by final buyer. There are several means in which cold chain products can be transported, recently already in smart containers, and as a novelty, delivered in humidity and temperature controlled parcel delivery boxes in smart homes [12], especial convenient for older inhabitants. Many countries have established food safety regulations, such as product temperature regulation along the supply chain; obligatory recording of air and product temperature in refrigerated vehicles, in production cells and loading-reloading places; and standardized equipment traceable to international measurement standards. As the regulations differ from country to a country, European regulations and directives have been harmonized for prescribed temperature of frozen products at most steps of the production and distribution chain. The European Community directive EN 92/1 CEE has prescribed air temperature recording in transportation, warehousing and retail display. There is also an ''Agreement on the international carriage of perishable foodstuffs and special equipment to be used for such carriage'' – ATP [26] established in 1970 and last amended in 2014, by which contracting parties have to take the measures necessary to ensure that equipment, used for carriage of perishable goods meets the technical standards described in the agreement. The technological potential in the management of cold supply chains (CSC) is studied by introducing a smart measurement device, which would extend the existing temperature measurements with humidity measurements with additional focus on the accuracy of measurement results. As mentioned in [3,4,24,25] instrument's capacity shall also be taken into account and also optimized calibration and reports on uncertainty is important as described

in [1,2]. By adding data-logging and access to data through cloud computing, it is expected to have a significant impact on a CSC management and the Net Present Value (NPV) in the supply chains, as explained in [8], and further developed in [11]. The papers are based on the previous results of risks evaluation [6,7,9,10] and they show how strong influence on NPV have delays and accuracy in detection of changes in perishability dynamics and proper forecasting of remaining shelf life when the product is bought by final user. This raises the incentive for evaluation of investment in such technology, so the paper investigates how these innovative technologies could contribute to an increase in the NPV of activities in CSCs.

2 MEASUREMENT DEVICES IN COLD SUPPLY CHAINS

The ATP gives details on temperature measurements that needs to be carried out, including the traceability of the measurements and also procedures for thermal evaluation with respect to prescribed insulation. On the other hand, perishables are not influenced only by temperature but also other very important environmental parameter – humidity, and consequently, moisture. Table 1 shows an optimal storage humidity for certain fruit and vegetables as presented in [23].

Table 1: An optimal storage humidity for certain fruit and vegetables

Commodity	*Shelf life / days*	*Relative humidity / %*
Asparagus	14-28	90-100
Carrots	28-180	90-100
Cucumbers	10-14	95
Banana	7-28	85-95
Strawberries	2-7	90-95

Humidity can especially become an issue, when vapor condenses during temperature shocks or during more latent non-even temperature/humidity distribution. Namely, the change in temperature for 1 °C can cause a relative humidity change as high as 7 %, which could result in unwanted condensation for longer periods of time [20]. Therefore, if the humidity levels are close to condensing limits, it is important to note that temperature distribution inside the container alone can cause the condensation without being detected by the controlling thermometer. For this reason, proper accurate humidity measurements are important to implement. Taken into account the sharply defined natural threshold in humidity – the condensation – and the sensitivity, mentioned above (up to 7%/°C), it is important to define the target uncertainty of the measurements in order to be able to advice the optimal system with respect to its cost and broad use. The characterization needs to be performed by a precision state-of-the-art reference instruments, which are traceable to humidity measurement (physical) standards. The end of such metrological pyramid is represented by a primary measurement standard (temperature, humidity and other quantities), which are compared at the international level and coordinated by the International Bureau of Standards (BIPM). One such example at the national level is a primary dew-point generator that is maintained at the University of Ljubljana, Faculty of Electrical Engineering, Laboratory of Metrology and Quality (LMK). It has been presented in [19] and evaluated as in an Euramet comparison [18] with other national measurement institutes worldwide. Figure 1 shows the LMK primary dew-point generator. Until the consumption, the distribution chain of food products is made up of many different stages. The environmental conditions vary in each step, especially during the loading and unloading outside air-conditioned warehouses, when can exceed the upper or lower limits.

This brings a need for an implementation of a smart device that would measure and trace different environmental parameters. Besides measuring, the device would also integrate geolocation with time-stamping and communication with the network cloud, as a part of an internet of things (IoT). This would provide the base for real-time decision making and also it would provide a ground for recently introduced concept of ''cold traceability''. This concept helps tracing different groups of perishable goods like poultry and other meat, fish, fruits and vegetable, confectionery, ice cream and other dairy products, which are transported under different cooling requirements. The question arises, how the cargo handling can be improved in order to meet the final requirements?

Figure 1: The LMK primary dew-point generator, which represents a national reference for humidity measurements

Also, what should be the conditions and restrictions of a dynamic system management so that the final consumers will buy the food with remaining shelf life (RSL) long enough? Are the static limitations of the perturbations at each moment as good as dynamic restrictions? To monitor the chain flows and to achieve the optimal control in the procedures of cooling and improvements, the skeleton of the model on the basis of input–output analysis in frequency domain, where timing is easier manipulated, is suggested. This approach was first presented in [15], 50 years ago, in 1967. The first overview through the MRP Theory was given 20 years ago, in [16] and later developed further in [17]. In parallel, games theory has been introduced and risk approach [5], fuzzy modelling [22], and also transportation matrices. The latter have enabled extension of the model with modelling of the impact of activity cells location in global supply chains on NPV of total networks [6]. Also, the evaluation of human resources and ergonomics in the chain have been enabled by introduction of production functions [3] and close loop of a supply chain has been evaluated based on MRP Theory regarding economic-environmental equilibrium [21,22]. The results of cold chain management (CCM) are finally expressed and valuated in the Net Present Value (NPV) of the final delivery, reduced by the costs of production, distribution, additional cooling and deterioration of goods in the criterion function. The risk of faster perishability and varying time delays is conveniently analyzed using NPV as criterion function, derived from annuity stream in frequency domain. With the development of the IoT, the idea of moving activity cells with sensors have been presented in [11]. In this paper the perishability dynamics was assumed based on variation of transportation lead time and other time delays at given changes of temperature, but humidity was not considered in that paper. Therefore, fine tuning of the data is added in this paper, considering variation of humidity at different temperatures with its impact on the NPV.

Currently, DHL is developing a technology for the next generation of cold chain visibility, forecasting near real-time visibility; full condition sensing, which monitors temperature and humidity, indicates shock and light; geo positioning; proof of delivery; real-time alert notifications via SMS/e-mail [7,8]. These solutions would benefit from combining them with a decision support system, which would contribute to increase NPV of activities in a supply chain. Besides, the measurements don't seem to consider all aspects of accurate humidity measurements. As mentioned in [11,13] through the use of smart devices possessing multiple micro-sensors deployed in large numbers over wide areas covered by supply networks, an unprecedented capability exists for monitoring, tracking, and controlling fruit and vegetable or

other food packages, pallets, parcel box or individual items as cargo in transportation or warehousing. In controlling environment conditions, wireless sensors could measure physic-mechanical properties. Although sensors have a long history, the realm of wireless sensors is relatively new. Regarding parcel boxes in Germany, where delivery points also at home (domotics) continue to multiply [12], logistics providers face new challenges. They need to find creative cost-effective solutions that provide value for the end customer and operational efficiency for the logistics provider. IoT in the last mile can connect the logistics provider with the end recipient in exciting ways as it drives dynamic new business models. IoT infrastructure for Cyber-Physical systems as presented in [11] could create the "last mile" optimized collection from the mail boxes. Sensors placed inside the box detect whether it is empty and, if so, transmits a signal that is then processed in real time.

3   EXTENDED MRP MODEL OF COLD SUPPLY CHAIN CONTROLLED ALSO BY HUMIDITY SENSORS

3.1 The basic theory

We have to know the correlation between temperature dynamics, humidity and perishability $\boldsymbol{\beta} = \left[\beta_{i,j}(\chi, \vartheta) \right]$ to protect the integrity of shipments and to improve NPV of all activities involved in CSC. To achieve the optimal control in the procedures of cooling the cargo, Input–Output approach first studied by Grubbström [15, 16, 17] and further developed to the CSCM model [8, 11] should give the basic approach to precise feedback control and managerial decision making.

According to the basic MRP theory the j-th process is run on activity level P_j, the volume of required inputs of item i is $h_{ij}P_j$ and the volume of produced (transformed, or uploaded, unloaded, conserved, detected as changed in quality) outputs of item k is $g_{kj}P_j$. The total of all net output may then be collected into the column vector $(\mathbf{G} - \mathbf{H})\mathbf{P}$. We assume $\mathbf{P}$ (and thereby net production) will be a time-varying vector-valued function. In CSCM systems, *lead times* between nodes τ_{ij} and inside nodes τ_i and related perishability dynamics which determine the remaining shelf life RSL at the destination. The remaining shelf life (RSL) is the time from a certain moment of measurement until the time when the product is still acceptable for consumption by a customer. Depending on supply-chain ambient conditions (temperature: ϑ and humidity: χ, quality will decrease over time at varying rates, described by dynamics $\alpha(\chi, \vartheta)$, where in time τ *RSL* changes by $\alpha(\chi, \vartheta).\tau$. If the decay acceleration rate changes at the moments $t + \tau_i$, we can write:

$$RSL(t + \tau) = RSL(t) - \sum_i \alpha(\chi, \vartheta)_i \tau_{i,j} , \qquad (1)$$

where $\sum_i \tau_{i,j} = \tau$ and j is the index of a child node of the i-th chain node on the route.

In [11], for the case of simplicity, we have considered a *linear logistic* system, for which the components of activity j need to be in place τ_j time units before completion (packaging, sorting…) and sent from parent node i to j having floating point of sensor in j' for additional time delay τ_{ij} in advance. The matrix of SCS graph, denoted by $\tilde{\mathbf{H}}(s, \boldsymbol{\beta}(\chi, \vartheta), \boldsymbol{\delta})$ include floating points between each parent and child node of activity cells where changes of perishability are detected in time distance $\delta_{i,j}$, here on the bases of more precise humidity detection.

$$\tilde{\mathbf{H}}(s,\boldsymbol{\beta}(\chi,\vartheta),\boldsymbol{\delta}) = \begin{bmatrix} 0 & 0 & \ldots & \ldots & \ldots. & & 0 \\ h_{21}e^{s\beta_{21}^*\delta_{21}} & 0 & 0 & 0 & \ldots\ldots & & 0 \\ h_{31}e^{s\beta_{31}\tau_{31}} & h_{32}e^{s\beta_{32}(\tau_{31}-\delta_{21})} & 0 & 0 & \ldots\ldots & & 0 \\ & & h_{43}e^{s\beta_{43}^*\delta_{43}} & 0 & \ldots. & & 0 \\ & & h_{53}e^{s\beta_{53}\delta_{53}} & h_{54}e^{\beta_{54}s(\tau_{53}-\delta_{43})} & : & \cdots\cdots\cdots & \cdots\cdots\cdots \\ \cdots & \cdots & \cdots & \cdots\cdots & \cdots: & : & : \\ 0 & 0 & 0 & 0 & \ldots & h_{n-1,n-2}e^{s\beta_{n-1,n-2}^*\delta_{n-1,n-2}} & 0 \\ 0 & 0 & 0 & 0 & \ldots & h_{n,n-2}e^{s\beta_{n,n-2}\delta_{n,n-2}} & h_{n,n-1}e^{s\beta_{n,n-1}(\tau_{n,n-2}-\delta_{n-1,n-2})} \end{bmatrix}, \quad (2)$$

and lead times in activity cells are described by matrix $\quad \boldsymbol{\tau} = diag\left[e^{s\tau_1},...,e^{s\tau_n}\right] \qquad (3)$

In [11] we have assumed that items have unit economic values, which could be different in different nodes in the graph of a supply chain, but also depend on customer remaining shelf life (CRSL=RSL at final node) while delivered to a customer in domotic systems. We also assumed that if suddenly cargo becomes highly exposed to risk of decay and smart devices recognize that, the system could report to near city, which host child node, and smart city can organize the transactions for such cargo locally at lower but acceptable prices, or in the worst case the city can organize a disposal of rotten goods. Therefore, we shall write the price vector in dependence of the CRSL, being a row vector, dependent on temperature ϑ and humidity χ trajectories from origin of the cargo to the destination:

$$\mathbf{p}(CRSL(\chi,\vartheta)) = \left[p_1, p_2,..., p_n\right]_{(\chi,\vartheta)} \qquad (4)$$

Introducing the cyclical behaviour of logistic activities, where $\mathbf{P}$ is vector of constant batch sizes, describing the total amount to be delivered during the period T while activity j start at time t_j after start of cycle, the production vector in frequency domain if frequency is replaced by the continuous interest rate, is written:

$$\tilde{\mathbf{P}}(\rho) = diag\left[e^{-\rho t_1} \quad e^{-\rho t_2} \quad ... \quad e^{-\rho t_n}\right]\cdot(1-e^{-\rho T})^{-1}\mathbf{P} \qquad (5)$$

If we write $\mathbf{K}$ as sum of setup, ordering and fix costs per cycle appearing at each node, also in floating points (could be 0), collected into the row vector: $\mathbf{K} = \left[K_1, K_2,..., K_n\right]$ and if vector of timing is $\tilde{\mathbf{v}}(s) = \left[\tilde{v}_1(s),...., \tilde{v}_n(s)\right]^T$, following the procedure of Bogataj et al. (2017) we can write

$$\text{NPV}(CRSL(\chi,\vartheta)) = \mathbf{p}(CRSL(\chi,\vartheta))\left(\mathbf{I}-\breve{\mathbf{H}}(\rho,\chi,\vartheta,\boldsymbol{\delta})\right)\tilde{\mathbf{P}}\left(\rho\right)-\mathbf{K}\tilde{\mathbf{v}}(\rho)-NPV_{tr} \quad . \qquad (6)$$

In (6) the NPV of transportation costs, if b_{ij} is cost of transportation of one item from i to j is

$$NPV_{tr} = \mathbf{E}^T \cdot \begin{bmatrix} 0 & 0 & \ldots & \ldots & 0 \\ h_{2,1}b_{2,1}\tau_{2,1} & 0 & \ldots & \ldots & 0 \\ : & h_{3,2}b_{3,2}\tau_{3,2} & 0 & \ldots & 0 \\ : & : & : & \ldots & : \\ h_{n,1}b_{n,1}\tau_{n,1} & \ldots & \ldots & h_{n,n-1}b_{n,n-1}\tau_{n,n-1} & 0 \end{bmatrix}\cdot\tilde{\mathbf{P}}(\rho) \qquad (7)$$

At a given CSC depends on time delays in detection of deterioration and on proper forecasting of the development of deterioration, which all depend on quality of the smart devices, also on precise measurement of humidity and proper and quick adaptation of temperature and humidity. Therefore, supply chains and cities need smart devices to control, optimally decide and to communicate with agglomerations in the surroundings of CSC to sell quickly the products which are still of acceptable quality, to mitigate the consequences of the risk realization.

3.2 What the numerical example of Bogataj et al. (2017) tell us

In [11] we have presented a case study where we supply Berlin with lettuce from Murcia, Spain. Each day the truck with a smart container of 500 boxes, each having 20 kg of lettuce, which is harvested in Murcia region, is packaged on the fields nearby and immediately sent from the field to the warehouse located 35 hours far from these fields and 2 hours from Berlin. The temperature in the container was assumed to be close to 4^0C. In such a case the shelf life of this kind of lettuce is five days at 90% of humidity. Here we assumed that the decay acceleration factor α at 4^0C is 1. Decay acceleration factor α means the decrease in shelf life in one day when cargo is exposed to a given ambient temperature. It means that each day the RSL falls for one day. If the decay dynamics would be without perturbations there was assumed NPV equal to 2,135,423 *Euro*, but if the stipulated quality would fall under threshold, the costs of landfill and all logistics would be 1,708,338 Eur. We did not consider that also humidity can be lower. If the *CRSL*, stipulated in the contract, is fallen under threshold with probability $\pi = 0.03$, the expected net present value of such a supply chain is $(1-\pi) \cdot 2{,}135{,}423 - \pi \cdot 1{,}708{,}338 Eur = 2{,}020{,}110 Eur$. But if we suppose that also humidity could fall from 90% to 70% so that the threshold is reached and CRSL fall under threshold with probability $\pi = 0.05$, in this case the expected *NPV* is 1,943,235 *Eur*. The difference in these two numbers help us to decide how much we are willing to invest in better humidity measurement and control and organisation a market for lower quality food along the road.

4 CONCLUSION

A demonstration is given on how a postharvest supply chain management, introducing sensitive sensors of humidity and temperature, can contribute to mitigation of risks in CSC when exploring new technologies available for better CSCM and better communication with final buyers. It can contribute to more agile actions towards postharvest loss prevention.

In CSC lead time and other time delays influence NPV, therefore, we have to see how new investments in better control increase NPV and if this difference is higher than investments in smart control, communication and optimization devices than we have to consider to invest and to develop a new generation support using precise humidity measurement devices as well to a market along the main road for cargo of reduced quality before it goes to landfill.

References

[1] Beges G, Drnovsek J, Pusnik I, Bojkovski J, 2002. Measurement uncertainty, traceability and evaluation of test results in testing laboratories. Measurement science & technology, 13:565-572

[2] Beges G, Drnovsek J, Pendrill L R, 2010. Optimising calibration and measurement capabilities in terms of economics in conformity assessment. Accreditation and Quality Assurance, 15 (3): 147-154

[3] Beges G, Drnovsek J, Ogorevc J, Bojkovski J, 2015. Influence of different temperature sensors on measuring energy efficiency and heating-up time of hobs. Int. j. of Thermophysics, 36 (2/3)

[4] Beges G, Rudman M, Drnovsek J, 2011. Evaluation of flat surface temperature probes. Int. j. of Thermophysics, 32 (1/2), 397-406

[5] Bogataj, D., Bogataj, M., 2007. Measuring the supply chain risk and vulnerability in frequency space. Int. J .Prod. Econ., 108(1-2): 291-301

[6] Bogataj, D., Bogataj, M., 2011. The role of free economic zones in global supply chains - a case of reverse logistics. Int. J .Prod. Econ., 131(1): 365-371

[7] Bogataj, D., Vodopivec, R., Bogataj, M., 2013. The extended MRP model for the evaluation and financing of superannuation schemes in a supply chain. Technological and economic development of economy, 19(S1): S119-S133

[8] Bogataj, D., Bogataj, M., 2015. Floating points of a cold supply chain in an environment of the changing economic growth. In: ZADNIK STIRN, L. (Ed.), et al. SOR '15 proceedings, Ljubljana: Slovenian Society Informatika, SOR: pp. 47-52.

[9] Bogataj, D., Ros-McDonnell, D., Bogataj, M., 2016, Management, financing and taxation of housing stock in the shrinking cities of aging societies. International journal of production economics, 181: 2-13

[10] Bogataj, D., Aver, B., Bogataj, M., 2016. Supply chain risk at simultaneous robust perturbations. International journal of production economics, 181(A): 68-78

[11] Bogataj, D., Bogataj, M., Hudoklin, D., 2017. Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model. Int. j. prod. econ, doi.org/10, 1016/j.ijpe.2017.06.028, available online June 2017.

[12] DHL parcel delivery boxes now in apartment buildings. 2015. DHL Global, http://www.dhl.com/en/press/releases/releases_2015/group/dhl_parcel_delivery_boxes_now_in_apartment_buildings.html (Accessed 29/6/17)

[13] DHL. 2017. Internet of Things in Logistics. DHL Trend Research. (Accessed 29/6/2017) http://www.dhl.com/en/about_us/logistics_insights/dhl_trend_research.html

[14] Grgić G, Pušnik I, 2011, Analysis of thermal imagers. International journal of thermophysics, vol. 32 (1/2): 237-247

[15] Grubbström, R.W., 1967. On the application of the Laplace transform to certain economic problems, Manag. Sci., 13: 558-567

[16] Grubbström, R.W., Bogataj, L., 1998. (Eds.), Input–Output Analysis and Laplace Transforms in Material Requirements Planning, Storlien, 1997. FPP, Portorož, 1998.

[17] Grubbström, R. W., 2007. Transform methodology applied to some inventory problems. Z. Betriebswirtsch, 77 (3): 297-324.

[18] Heinonen M, et al. 2012. Investigation of the equivalence of national dew-point temperature realizations in the -50 °C to +20 °C range. International journal of thermophysics, 33 (8/9)

[19] Hudoklin D, Bojkovski J, Nielsen J, Drnovšek J, 2008. Design and validation of a new primary standard humidity sensors. Measurement: 41 (9), 950-959

[20] Kentved A.B., Heinonen M., Hudoklin D., 2012, Practical study of psychrometer calibrations, Int. J. of Thermophysics, 33(8–9): 1408–1421

[21] Kovačić, D., Bogataj, M., 2017. Net present value evaluation of energy production and consumption in repeated reverse logistics. TEDE 23(6): 877-894

[22] Kovačić, D., Usenik, J., Bogataj, M., 2017. Optimal decisions on investments in urban energy cogeneration plants - extended MRP and fuzzy approach to the stochastic systems. Int. J .Prod. Econ., 183(B): 583-595

[23] Paull R.E., 1999, Effect of temperature and relative humidity on fresh commodity quality, Postharvest Biology and Technology 15: 263–277

[24] Pušnik I, Drnovšek J., 2005, Infrared ear thermometers - parameters influencing their reading and accuracy. Physiological measurement, vol. 26, 1075-1084.

[25] Pušnik I, Miklavec A, 2009, Dilemmas in measurement of human body temperature. Instrumentation science & technology, 37(5): 516-530

[26] UNECE, 2014. Agreement on the international carriage of perishable foodstuffs and special equipment to be used for such carriage, ATP (ECE/TRANS/232/Corr.1).

50 YEARS OF THE MRP THEORY

Danijel Kovačić
MEDIFAS, Mednarodni prehod 6, SI-5290 Šempeter pri Gorici, Slovenia
kovacic.danijel@gmail.com

Abstract: The idea of MRP theory started 50 years ago, with Grubbström's article published in Management Science under the title *"On the Application of the Laplace Transform to Certain Economic Problems",* where he firstly presented that time delays can be successfully introduced into modelling of systems of production economics, using Laplace transforms. In the sixties, the practical use of Laplace transform and annuity streams calculations were better known in electrical engineering and actuarial science. Therefore, this was a pioneering work in the field of material requirements planning (MRP), which provided the basis for calculating the net present value (NPV) of a system in which the continuous interest rate appears as the discount rate for cash flows. NPV approach is particularly convenient when delays and risks should be evaluated in the production systems. The paper aims to present some main achievements in the development of MRP theory in occasion of the anniversary of 50 years after the issue of this first article, and 20 years after the first conference on MRP theory in *Storlien.*

Keywords: MRP Theory, input-output analysis, Net Present Value (NPV), Laplace transformations.

1 INTRODUCTION

The concept of production becomes particularly relevant with the emergence of the industrial revolution. Since then, various attempts have been made to scientific research and set up theoretical models, including the treatment of inventories [24]. The concept of Material Requirements Planning (MRP), which had already been well-established in practical use, was first described in details by Orlicky [47]. His professional work refuted the previous belief about the impossibility of scientific treatment of such systems, and raised motivation for further scientific research, such as MRP theory.
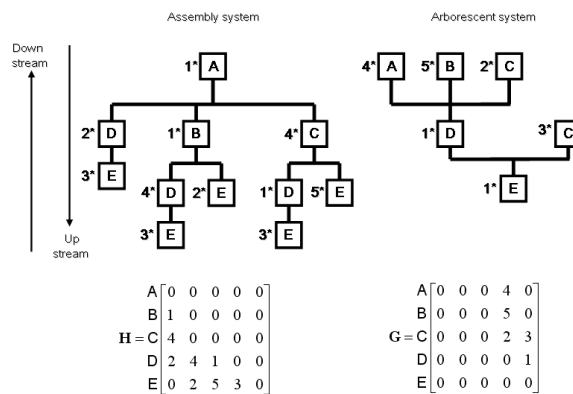
Figure 1: Examples of assembly and arborescent systems, written in a mathematical form of input-output matrices **H** and **G**, respectively [29].

MRP theory systems are captured within the Bill of Materials (BOM), where elements on upper level are assembled using one or multiple elements from preceding level (Fig 1). Tree structure of the product composition can also be written using the gozinto graphs, which were first introduced into the production models by Vazsonyi [54], [55]. Further, the input-output approach [46] based on the structure of the products written in the matrix form is much more suitable for mathematical modeling of production problems (Fig 1), which gives one of the pillars for MRP theory [28].

The idea of MRP theory started 50 years ago, with Grubbström's [20] work where he firstly presented that time delays can be successfully introduced into production economics using Laplace transforms. It is a pioneering work in the field of MRP, which provides the basis for calculating of the net present value (NPV) of a system in which the continuous interest rate is used as discount rate of cash flows. The value of an element from any level of the BOM is drawn up by adding the partial added value of elements from al preceding levels, which can be studied using the Leontief inverse. Approach with a NPV is more convenient than an average cost approach when dealing with production and storage problems. It allows easier evaluation of costs of time delays and their perturbations, and also evaluation of exposure to risks in the system. Among the earlier developments of the MRP theory are papers of Grubbström and his doctoral students at Institute of Production Economics at Linköping Institute of Technology [21], [23], [31], [32], [38]. The MRP theory at the early beginning was treated as a theoretical concept only, but was later found suitable to study the problems of the real world too; i.e. Grubbström applied it in the case of paper production [22]. Grubbström and Ovrin [34] are the first to use z-transformation in the MRP model, which is a discrete equivalent of Laplace transformation, to describe inventories and timings of production activities arising from lead times. They presented a generalized input matrix $\tilde{\mathbf{H}}(s)$ that, in addition to the input requirements, includes time delays of production at each activity cell: $\tilde{\mathbf{H}}(s) = \mathbf{H}\tilde{\boldsymbol{\tau}}(s)$ where $\tilde{\boldsymbol{\tau}}(s)$ is the diagonal matrix of delays inside each activity cell. Grubbström and Molinder [33] moved from discrete approach to a continuous time, using Laplace transforms. In their works we can find the fundamental equations of the MRP theory, which are the basis for all further research of system (Fig. 2).
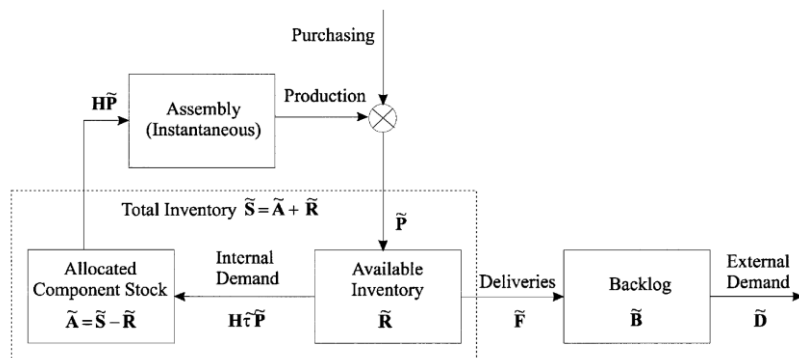
Figure 2: Material flows in the MRP processes including backlogs provide basis of fundamental MRP equations [27].

The symposium in *Storlien* in 1997 and proceedings [12] summarize the previous work on the development of the MRP theory and set the pillars for the modern development of the theory. The conference gave and incentive for further wider development of the theory, which today is intensively pursued in different directions of scientific study. Summarized presentation of all MRP equations and in-depth explanations of how input-output analysis (describing structures) and Laplace transforms (describing time delays) are used to calculate NPV of the system (to evaluate it). The overview of this basic theory is given in [28] and [37] and application to the material science and technology in [8].

2 MRP THEORY EXTENDED TO SUPPLY CHAIN MODELLING

Grubbström and Thorstenson [38] published their results obtained from applying the annuity stream principle for evaluating capital tied up in inventory and work-in-progress in 1986.

Segerstedt [48] introduced capacity constraints in the MRP. Capacities are included as n integral part of an input matrix, which consequently loses its otherwise quadratic form. Segerstedt's work is upgraded by Grubbström and Wang [39], and newer contribution in this area is the work of Grubbström and Huynh [30]. The constraints in capacities opened a new direction to the strategic decision making on investments in production economics [11], [45]. Study of availability of human resources [10] by introduction of production functions and actuarial mathematics into the MRP theory was provided by the research group at University of Ljubljana and MEDIFAS.

Since the problems of the real world are often not deterministic, the next great challenge of the theory was a transition towards stochastic processes. The first researched where level of stocks is linked to demand which occurs at independent time intervals of a stochastic nature was presented in [25]. Demand is met by production processes which have different volumes (contingents) of products that appear in different, stochastically distributed time moments. The relatively high complexity of such problems requires the use of numerical examples in finding optimal solutions. Stochastic processes are further in depth analyzed by Linköping School with series of papers written by Tang [49], [50], [51], [52], [53].

Possibilities for comprehensive research of inventories in stochastic systems are clearly evident from the works which introduce safety stocks into the MRP theory model [26], [27], [36]. These researches determine the size of contingents at lower levels of the batch and their impact on the behavior of the entire MRP system [35]. Slovenian research group contributed to further study of uncertainty in the production and distribution systems by studying vulnerability of such systems and exposure of such systems to risk [1], [3], and the connection with the theory of decisions and games [13], [40]. Uncertainties in production and closed loop supply chains have also been studied using fuzzy approach [19], [45]. Slovenian researchers also contributed to the development of the MRP with inclusion of the spatial component [2], [9], [14], [15].

Probably the most noticeable recent development of the MRP theory is the extension from its classical production and inventory field to the observation of a closed-loop global supply chains [18], [29], [41]. This transition extends the boundaries of the applicability of the theory from the microeconomic to the macroeconomic level, and trade-off between economy and environment. The MRP theory becomes a tool for adopting global, strategic decisions, while still allowing microanalysis of a single activity cell at the lowest level of the system. Transportation lead times play crucial role inside global supply chain systems [16]. Allocations of all elements of the system at any stage can be captured within transportation input-output matrices $\tilde{\Pi}_H(\rho)$ and $\tilde{\Pi}_G^{14}(\rho)$, which coincide with input-output matrices **H** and **G** [17]. Such extended MRP model allows for in depth research of global supply chains and their behavior, including locational analysis [42], or even energy needs and recycling activities in the closed loop systems [43]. In the research field of energy infrastructure fuzzy approach was implemented to soften otherwise exact and deterministic MRP theory model for investments evaluation [45]. Extended MRP theory was first used in practical application where Spanish baby food company's processes were evaluated [44]. Today's research of practical applications is directed to further improvements of global supply chains of perishable goods [5], [6], [7], [8] and towards the problems of ageing human resources. Impact of supporting robots on NPV of supply chains and other ergonomic solutions is analyzed, and possibilities to redirect a part of income of supply chains to pension funds for early retirement are studied. Both problems are studied through sophisticated embedded production functions and robotization in the extended MRP models [2].

3 CONCLUSION

50 years after the first introduction of Laplace transforms to production economics, 20 years after the first symposium on MRP in Storlien, and 10 years after the first location problems have been introduced to extend the MRP theory to model global supply chains, we can realize that the theory has become mature enough to be used for solving applied problems at nowadays opened questions on how to manage and control European and global supply chains.

References

[1] Bogataj, D., Aver, B., Bogataj, M. (2016). Supply chain risk at simultaneous robust perturbations. *Int.j.Prod.Econ* 181, part A: 68-78.

[2] Bogataj, D., Battini, D., Calzavara, M., Persona, A. (2017). Investments in workplace ergonomics for older workers from the supply chain approach. Proceedings of 24th International conference on production research, DEStech Publications, In Press.

[3] Bogataj, D., Bogataj, M. (2007). Measuring the supply chain risk and vulnerability in frequency space. *Int.j.Prod.Econ* 108(1-2): 291-301.

[4] Bogataj, D., Bogataj, M. (2011). The role of free economic zones in global supply chains - a case of reverse logistics. *Int.j.Prod.Econ* 131(1): 365-371.

[5] Bogataj, D., Bogataj, M. (2015). Floating points of a cold supply chain in an environment of the changing economic growth. In: ZADNIK STIRN, L. (Ed.), et al. SOR '15 proceedings, 13th International Symposium on Operational Research in Slovenia, Bled, Slovenia, September 23-25, 2015. Ljubljana: SDI_SOR, pp. 47-52.

[6] Bogataj, D., Bogataj, M., Drobne, D., Ros McDonnell, L., Rudolf, R., Hudoklin, D (2015). Investments in smart nano-control systems in cold supply chains. In: ZADNIK STIRN, L. (Ed.), et al. SOR '15 proceedings, 13th International Symposium on Operational Research in Slovenia, Bled, Slovenia, September 23-25, 2015. Ljubljana: SDI_SOR, pp. 53-59.

[7] Bogataj, D., Bogataj, M., Hudoklin, D. (2017). Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model. *Int.j.Prod.Econ* 193: 51–62.

[8] Bogataj, D., Drobne, D. (2017). Control of Perishable Goods in Cold Logistic Chains by Bionanosensors. In: Mehdi Khosrow-Pour, DBA, Materials Science and Engineering: Concepts, Methodologies, Tools, and Applications, Chapter 19, 471-497, IGI Global, Information Resources Management Association, USA. DOI: 10.4018/978-1-5225-1798-6

[9] Bogataj, D., Ros-McDonnell, D., Bogataj, M. (2016). Management, financing and taxation of housing stock in the shrinking cities of aging societies. *Int.j.Prod.Econ* 181(A): 2-13.

[10] Bogataj, D., Vodopivec, R., Bogataj, M. (2013). The extended MRP model for the evaluation and financing of superannuation schemes in a supply chain. *Technological and economic development of economy*, 19(S1): S119-S133.

[11] Bogataj, L., Bogataj, M. (2007). The study of optimal additional investments in capacities for reduction of delays in value chain. *Int.j.Prod.Econ* 108: 281-290.

[12] Bogataj, L., Grubbström, R.W (Eds). (1998). Input-output analysis and Laplace transforms in material requirements planning. *Symposium, Storlien, 1997*. ISBN 961-6044-30-3. Portorož: UL-Faculty of Maritime Studies and Transport.

[13] Bogataj, L., Horvat, L. (1996). Stochastic considerations of Grubbström-Molinder model of MRP, input-output and multi-echelon inventory systems. *Int.j.Prod.Econ* 45(1-3): 329-336.

[14] Bogataj, M., Bogataj, L. (2001). Supply chain coordination in spatial games. *Int.j.Prod.Econ* 71(1): 277-285.

[15] Bogataj, M., Bogataj, L. (2004). On the compact presentation of the lead times perturbations in distribution networks. *Int.j.Prod.Econ* 88(2): 145–155.

[16] Bogataj, M., Grubbström, R.W. (2012). On the Representation of Timing for Different Structures within MRP Theory. *Int.j.Prod.Econ* 140(2): 749–755.

[17] Bogataj, M., Grubbström, R.W. (2013). Transportation delays in reverse logistics. *Int.j.Prod.Econ* 143(2): 395–402.

[18] Bogataj, M., Grubbström, R.W., Bogataj, L. (2011). Efficient location of industrial activity cells in a global supply chain. *Int.j.Prod.Econ* 133(1): 243–250.

[19] Bogataj, M., Usenik, J. (2005). Fuzzy approach to the spatial games in the total market area. *Int.j.Prod.Econ* 93-94: 493-503.

[20] Grubbström, R.W. (1967). On the Application of the Laplace Transform to Certain Economic Problems. *Management Science 13*(7): 558-567.

[21] Grubbström, R.W. (1980). A Principle for Determining the Correct Capital Costs of Work-In-Progress and Inventory. *International Journal of Production Research 18*(2): 259-271.

[22] Grubbström, R.W. (1990). The distribution of an additive in a chemical process - an application of input-output theory. Engineering Costs and Production Economics 19(1-3): 333–340.

[23] Grubbström, R.W. (1991). A Closed-Form Expression for the Net Present Value of a Time-Power Cash Flow Function. *Managerial and Decision Economics 12*(5): 377-381.

[24] Grubbström, R.W. (1995). Modelling production opportunities - a historical overview. *Int.j.Prod.Econ 41*(1-3): 1-14.

[25] Grubbström, R.W. (1996). Stochastic relationships of a multi-period inventory process with planned production using transform methodology. *Int.j.Prod.Econ 45*(1-3): 407-419.

[26] Grubbström, R.W. (1998). A net present value approach to safety stocks in a planned production. *Int.j.Prod.Econ 56-57*(1): 213-229.

[27] Grubbström, R.W. (1999). A net present value approach to safety stocks in a multi-level MRP system. *Int.j.Prod.Econ* 59(1-3): 361-375.

[28] Grubbström, R.W. (2007). Transform Methodology Applied to Some Inventory Problems. *Zeitschrift für Betriebswirtschaft* 77(3): 297–324.

[29] Grubbström, R.W., Bogataj, M., Bogataj, L. (2007). A compact representation of distribution and reverse logistics in the value chain. V L.B. Ros McDonnell & L. Bogataj (Ed.), MEORL, Serial No. 5. Ljubljana: UL_EF, KMOR.

[30] Grubbström, R.W., Huynh, T.T.T (2006). Multi-level, multi-stage capacity-constrained production–inventory systems in discrete time with non-zero lead times using MRP theory. *Int.j.Prod.Econ 101*(1): 53–62.

[31] Grubbström, R.W., Jiang, Y. (1989). A Survey and Analysis of the Application of the Laplace Transform to Present Value Problems. *Revista di matematica per le scienze economiche e sociale 12*(1): 43-62.

[32] Grubbström, R.W., Lundquist, J. (1977). The Axsäter integrated production-inventory system interpreted in terms of the theory of relatively closed systems. *J. of Cybernetics 7*(1-2): 49-67.

[33] Grubbström, R.W., Molinder, A. (1994). Further theoretical considerations on the relationship between MRP, input-output analysis and multi-echelon inventory systems. *Int.j.Prod.Econ 35*(1-3): 299-311.

[34] Grubbström, R.W., Ovrin, P. (1992). Intertemporal generalization of the relationship between material requirements planning and input-output analysis. *Int.j.Prod.Econ 26*(1-3): 311-318.

[35] Grubbström, R.W., Tang, O. (1998). Modelling rescheduling activities in a multi-period production-inverntory system. V *10th International Working Seminar on Production Economics, Igls/Innsbruck, 16-20 Feb. 1998, Pre-prints, Vol. 2* (str. 67-84). Innsbruck: Igls.

[36] Grubbström, R.W., Tang, O. (1999). Further developments on safety stocks in an MRP system applying Laplace transforms and input-output analysis. *Int.j.Prod.Econ 60-61*(1): 381-387.

[37] Grubbström, R.W., Tang, O. (2000). An Overview of Input-Output Analysis Applied to Production-Inventory Systems. *Economic Systems Review* 12: 3-25.

[38] Grubbstrom, R.W., Thorstenson, A. (1986). Evaluation of Capital Costs in a Multilevel Inventory System by Means of the Annuity Stream Principle. *European Journal of Operational Research* 24(1): 136-145.

[39] Grubbström, R.W., Wang, Z. (2000). Introducing capacity limitations into multi-level, multi-stage production-inventory systems applying the input-output/Laplace transform approach. *International Journal of Production Research, 38*(17), 4227-4234.

[40] Horvat, L., Bogataj, L. (1999). A market game with the characteristic function according to the MRP and input-output analysis model. *Int.j.Prod.Econ 59*(1-3): 281-288.

[41] Kovačić, D., Bogataj, L. (2011). Multistage reverse logistics of assembly systems in extended MRP Theory consisting of all material flows. *Central European Journal of Operations Research* 19(3): 337–357.

[42] Kovačić, D., Bogataj, M. (2013). Reverse logistics facility location using cyclical model of extended MRP theory. *Central European Journal of Operations Research* 21(1): 41–57.

[43] Kovačić, D., Bogataj, M. (2015). Net present value evaluation of energy production and consumption in repeated reverse logistics. *Technological and economic development of economy* 23(6): 877-894.

[44] Kovačić, D., Hontoria, E., Ros-McDonnell, L., Bogataj, M. (2015). Location and Lead-Time Perturbations in Multi-Level Assembly Systems of Perishable Goods In Spanish Baby Food Logistics. *Central European Journal of Operations Research* 23(3): 607–623.

[45] Kovačić, D., Usenik, J., Bogataj, M. (2017). Optimal decisions on investments in urban energy cogeneration plants - extended MRP and fuzzy approach to the stochastic systems. *Int.j.Prod.Econ* 183(B): 583-595.

[46] Leontief, W.W. (1928). Die Wirtschaft als Kreislauf. *Archiv für Sozialwissenschaft und Sozialpolitik* 60(3): 577 623.

[47] Orlicky, J.A. (1975). Material Requirements Planning. McGraw-Hill, New York.

[48] Segerstedt, A. (1996). Formulas of MRP. *Int.j.Prod.Econ* 46-47(1): 127-136.

[49] Tang, O. (2000). Modelling stochastic lead times in a production-inventory system based on the Laplace transform method. *International Journal of Production Research* 38(17): 4217-4226.

[50] Tang, O., Grubbström, R.W. (2002). Planning and replanning the master production schedule under demand uncertainty. *Int.j.Prod.Econ* 78(3): 323-334.

[51] Tang, O., Grubbström, R.W. (2003). The detailed coordination problem in a two-level assembly system with stochastic lead times. *Int.j.Prod.Econ* 81-82(1): 415-429.

[52] Tang, O., Grubbström, R.W. (2005). Considering stochastic lead times in a manufacturing/remanufacturing system with deterministic demands and returns. *Int.j.Prod.Econ* 93-94(1): 285-300.

[53] Tang, O., Grubbström, R.W. (2006). On using higher-order moments for stochastic inventory systems. *Int.j.Prod.Econ* 104(2): 454-461.

[54] Vazsonyi, A. (1954). The Use of Mathematics in Production and Inventory Control. *Management Science* 1(1): 70-85.

[55] Vazsonyi, A. (1955). The Use of Mathematics in Production and Inventory Control II. *Management Science* 1(3-4): 207-223.

CONTRIBUTIONS TO THE LONG-TERM CARE INSURANCE FUND FOR WORKERS WHO HOLD PHYSICALLY DEMANDING AND LABOUR-INTENSIVE JOBS IN SUPPLY CHAINS

Valerija Rogelj and **Marta Kavšek**
Fakulteta za organizacijske študije, Ulica talcev 3, Novo mesto, Slovenia
E-mail: valerijarogelj@gmail.com

Abstract: In national pension schemes, also the retirement age of industrial workers who do physically demanding and labour-intensive work is increasing. In our research, we discovered that functional capacities of these workers are decreasing on average much earlier than functional capacities of other workers and they are becoming dependent on the help of others and require Long-Term-Care (LTC) on average 4 years earlier. Therefore, we should put in place not only supplementary occupational pension schemes, which would finance the early retirement of this workers but also supplementary LTC schemes which would finance prolonged period when they are dependent on the help of others. From employees' gross income, employers should consider also contributions to supplementary LTC insurance fund for industrial workers who do physically demanding and labour-intensive work. What percentage of gross income employers have to contribute to LTC insurance fund to cover these expenditures for earlier LTC is calculated in this paper. For this purpose, we have developed an actuarial model of the net present value of the expenditures of LTC regarding differences in trajectories of multiple decrements of the elderly who performed physically demanding work and those whose work was not so physically demanding. We found that the *NPV* of the expenditures for LTC for those who performed physically demanding work is more than 47 % higher than the *NPV* of the expenditure of LTC for others.

Keywords: physically demanding working places, Long-Term-Care, annuity stream, retirement age, supply chain, human resources

1 INTRODUCTION

The European working population is ageing [3, 13, 14]. This influence reduction in working population in urban agglomerations [7, 8, 10] and quality of production in supply chains [4], [5, 6]. It also could influence disruptions of the chains [8]. As considered in [1, 2, 5] the early retirement age, as determined in an occupational pension scheme, could decrease by increasing the contribution rate from increased gross earnings $(1+\alpha)c_L$ to the eerily retirement occupational pension schemes in the amount of αc_L. Here c_L is the gross earning which already includes the social contribution to normal pension and other social security funds, but no extra funding is assumed for those who hold physically demanding and labour-intensive jobs. Thus, if the labour expenditure would increase from c_L to $(1+\alpha)c_L$, where αc_L is factored into the additional occupational pension scheme, the retirement age of such workers can be lowered, which could increase the quality and quantity of production in the workplace and reduce the lead time in a supply chain. In our research, we also discovered that functional capacities of these workers who hold physically demanding and labour-intensive jobs are decreasing on average much earlier than of other workers, and they are becoming dependent on the help of others, and become entitled to the so called Long-Term-Care (LTC) on average 4 years earlier [16, 17]. Today, these expenditures are partly covered by payments from social funds, but mostly from the savings of workers and their families or from reverse mortgages [4, 11, 12]. Therefore, we should put in place not only supplementary occupational pension schemes, which would finance the early retirement of this workers but also supplementary LTC schemes which would finance prolonged period when they are dependent on the help of others.

The benefit to the total supply chain could be evaluated through Grubbström's MRP theory using the *NPV* approach, as presented in [6]. In a paper of Battini et al. [1] following the idea

of [2] we have seen that also investments in improved ergonomics which often include supporting robots, could be modelled as increased labour expenditures. Therefore we shall structure αc_L so as to include the parts for earlier retirement α_{RCL}, investments in supporting robots to improve ergonomics [1, 2] in the amount of α_{ECL} and α_{LTCCL} as contribution for LTC insurance. The system could be sustainable if inequality (16) derived in [7] is achieved in the supply chain. From this condition, it follow

$$
\sum_{i=1}^{n} c_i (\alpha_{R,i} + \alpha_{E,i} + \alpha_{LTC,i}) L_i \tilde{x}_i (\rho - \omega) \leq \sum_{i=1}^{n} \pi_i \tilde{x}_i (\rho - \omega) - \sum_{i=1}^{n} \pi_i \tilde{x}_i (\rho) +
$$
$$
+ \mathbf{K} \left\{ \tilde{\mathbf{t}}(\rho - \omega) \tilde{\mathbf{T}}(\rho - \omega) - \tilde{\mathbf{t}}(\rho) \tilde{\mathbf{T}}(\rho) \right\},
$$

(1)

where the notation is the following:

$\tilde{x}_i$	- Intensity of net production	$\mathbf{K}$	- Vector of setup and other fix expenditures per cycle
ρ	- Continuous interest rate	L_i	- Number of workers needed for net intensity
$e^{\omega \cdot t}$	- Growth of production	$\tilde{\mathbf{t}}$	- Vector of timing of individual activity cells
π_i	- Price of item i	n	- Number of activity cells in a supply chain
$\tilde{\mathbf{T}}$	- Length of cycle		

What is the difference between contribution rate α_{LTC} for the workplaces where physically intensive labour is required, and for others is the main question of this paper. Our analysis wishes to contribute the model of dataset for *NPV* evaluation on the bases of *MRP theory,* where production functions and functional decline of aging workers are included in the labour cost as the subject of an investigation regarding their impact on *NPV of profit.* Existing national health insurance does not provide sufficient funds to cover expenditures for long term care when people become dependent on the help of other, as defined by the Blue Book of standards and norms in nursing and midwifery nursing and care [18]. The consequence of this is a lack of trust in the national institutions that provide insurance for LTC related expenditures. It is necessary to re-establish trust, and this requires the establishment of a new, transparent, secure and a long-term sustainable system of LTC, which will provide the coverage of the actual needs of people, who are dependent on the assistance of others. Such needs were also expressed by the European Commission in its publication *Adequate social protection for long-term care needs in an aging society* [12, 13, 14].

The basic aim of this paper is to develop a comprehensive and transparent model to support decision making, which will show to the stakeholders in the system of LTC, the impact of changes in the variables of the model on *NPV* of total supply chain related to contributions for LTC insurance. Here we distinguish between those, who have worked in physically demanding and work intensive workplaces, and all other elderly.

2 METHOD

We have interviewed the elderly from three different nursing homes in Slovenia (sample of 100 respondents from the population of 20,000 nursing homes residents in Slovenia) about their lifestyle and the age when entering the institutional care. Respondents were divided into two groups: the ones who hold physically demanding and labour-intensive jobs and those who did not. The actuarial present value of annual payments of annual amounts of care – AAC,

which is paid to cover LTC related expenditures of the insured person aged *x* years from the moment of obtaining the right to the LTC to his/her death, denoted by *NPV LTC_x*, is written and calculated by the following formula [15]:

$$NPV LTC_x = \sum_{j=0}^{100-x} {}_j p_x \cdot v^j \cdot AAC_{x+j} \qquad (2)$$

where: ${}_j p_x$ is the probability that a person aged *x* years survive *j* years, and *v* is a discount factor and AAC_{x+j} is the annual amount of care at the age of *x+j* years. Services to be performed under the LTC, are divided into basic care, social care, and nursing. Basic services of LTC include: accommodation, preparation of food, technical support and transportation. Basic daily nursing activities include (ADL - *Basic Activities of Daily Living*): bathing, dressing, feeding, positions in bed and getting up from it, movement, using the lavatory and interventions by doctor orders - such as: wound wrapping, pain relief therapy, a distribution of medication, monitoring vital functions, and more. Supportive daily tasks (IADL - *Instrumental Activities of Daily Living*) are primarily food preparation, laundry, transportation, and cleaning. The methodology for determining the general elements of pricing for the social services at today's living and organizational forms is defined by the Regulation of methodology for pricing welfare services [19]. This regulation determines all expenditures that can be considered as elements of the pricing.

3 RESULTS

3.1 The lifetime expenditure on LTC for beneficiary who did not hold physically demanding and labour-intensive job (in average)

This beneficiary is - according to our statistics – on average, 72 years old at the time of the occurrence of dependence on the help of others. The average beneficiary receives a old age retirement pension of 600.00 € per month. The first three years, from the age of 72 until the age of 75, he is receiving social care and assistance at home. Two years, from the age of 72 up to the age of 74, he needs help for one hour per day – that costs him 8.00 €, the additional 8.00 € is paid by the municipality of his residency. In the third year, from the age of 74 to the age of 75, he needs 2 hours of care per day. This costs him 16.00 €, and the municipality of his residency adds another 16.00 €. Community Nursing Service visits the beneficiary as a chronic patient 4 times per month, costing 93.60 € per month. We assume that expenditure for food on the average is 200.00 € per month. The expenditure of basic, social and health care in the first years amounts to 773.60 € per month. In the third year, when the beneficiary needs more of direct support, expenditures increase to 1,253.60 € per month. In the next two years, from the age of 75 to 77, the beneficiary moves to a Serviced apartment, where his expenditures increase for the expenditure of the rent – that is 185,00 € per month. This increases the monthly expenditure to 1,438.60 €. In the fifth year, when the beneficiary reaches 77 years of age, he/she moves into a Nursing home.

The monthly expenditure of basic care is 485.40 €, the monthly expenditure of social services is 0 € and the expenditure for nursing in the category 1 is 725.10 € per month. Of this amount, the expenditure of nursing care covered by Health Insurance Institute is 239.70 € per month. The beneficiary is entitled to an allowance for care and help according to the Law on Pension and Disability Insurance in the amount of 146.06 €. When he reaches 76 year of age, due to declining functional abilities, the needs for services of nursing are increased. That is why the patient is assigned the category Nursing 2, where the nursing expenditure 350.10 € per month, the expenditure of social care is 138.30 € per month, and the expenditures of basic care are 485.40 €, which increases the total expenditure to 973.80 €.

At the age of 81 years old, the in-care faces the additional drop of functional capabilities, so that he is allocated in the category Nursing 3, where the expenditure of nursing is 491.10 €, social care expenditure is 141.34 € and the expenditure of basic care is 485.40 € per month, which increases the total expenditures for care to a 1,263.90 € per month. This situation lasts for up to 89 years of age. At the 89 years of age, the in-cares functional capabilities drop to the extent that it is completely dependent on the assistance of others. So the basic care costs 485.40 € per month, social care is 417.90 €, of which the DPP is 146.06 € and the rest of the amount covers the beneficiary – that is 271.84 € per month; the nursing in category 3 is 491.10 €, which represents a monthly expenditure of care of 1,394.40 € per month.

3.2 Summary – Calculation of *NPV* for beneficiary who did not hold physically demanding and labour-intensive job

The *NPV* of total expenditures of LTC for the average elderly person who has not worked in particularly difficult professions, at the rate of 0.62 % (as was the market interest rate for 10-year Slovenian government's bonds on 30. 9. 2016) is **127,147.83 €**. Of this amount the *NPV* for rent of non-profit sheltered housing is 3,226.99 €, *NPV* of primary care is 42,954.09 €, *NPV* for social care is 49,132.38 € and the *NPV* for nursing is 31,834.38 € (see Table 2 in the paper of Kavšek and Bogataj [17]).

3.3 The lifetime expenditure on LTC for beneficiary who hold physically demanding and labour-intensive job and *NPV* of these expenditures (in average)

It is expected, that the man, who performed heavy physical labour in his working age and has retired at the age of 65, when he secured a monthly pension benefit of 600 €, will be able to live independently in his own home, where he lived before retirement. It can be expected that after three years; his functional capacity will decrease so that he will need home care. He will need 30 hours of home care per month in the first two years and 60 hours per month in the third year. Later, he will move to the sheltered housing and lived there for two years. After these two years, he will have to move into institutional care in the home for the elderly. The life expectancy and the associated cash flows are calculated in accordance with the complete tables of mortality 2007 for the Republic of Slovenia [20].

3.4 Summary – Calculation of *NPV* for beneficiary who hold physically demanding and labour-intensive job

According to Table 1 in [17], it follows that the *NPV* of the total expenditure of the LTC for the average elderly person, who hold physically demanding and labour-intensive job, at the interest rate of 0.62 % (as was the market interest rate for 10-year Slovenian government's bonds on 30. 9. 2016) is **181,555.36 €**. The structure of this amount of *NPV* is the following: for rent of non-profit sheltered housing is 3,920.39 €, *NPV* of basic care is 62,597.49 €, *NPV* for social care is 66,123.70 € and the *NPV* for nursing is 48,913.79 €.

The beneficiary who hold physically demanding and labour-intensive job, came into care 4 years earlier (in average), which is to be expected according to our tests of differences in entry age. Assuming that, while life expectancy for both beneficiaries by the age of 65 is the same (as our statistics do not perceive differences in life expectancy [20]) it is expected, that the total net present value of expenditures of care for elderly people, who hold physically demanding and labour-intensive job, will be 43 % higher than for those, who had not performed hard work. This amount should be part of participation in gross earnings to special professional LTC fund.

4 CALCULATION OF CONTRIBUTION $\alpha_{LTC,i}c_i$ FOR LTC INSURANCE WITH PAYMENT PERIOD 40 YEARS FOR MAN BETWEEN AGE 25 TO 65 WHO HOLD PHYSICALLY DEMANDING AND LABOUR-INTENSIVE JOB WHICH COVER EXPENDITURE FOR THE MINIMUM STANDARDS OF LTC

The fundamental question that we have set and should be answered is how this difference would be covered from gross salaries during the working period of a person who holds physically demanding and labour-intensive job. From the inequality (1) in the structure c_i ($\alpha_{R,i} + \alpha_{E,i} + \alpha_{LTC,i}$) we need to calculate how much should be $\alpha_{LTC,i}c_i$ payable for 40 years for those between age 25 to 65, who are paying a part of their salary to additional LTC fund, because they hold physically demanding and labour-intensive jobs. They need to accumulate the difference 181,555.36 € - 127,147.83 € = 54,407.53 € in LTC fund when they reach age 65 years to cover their higher LTC expenditures according to the prescribed minimal standards.

Based on the complete tables of mortality for 2007 in the Republic of Slovenia [20] we can calculate the amount of premium $\alpha_{LTC,i}c_i$ for a worker who starts to work and contribute to the additional LTC fund when he reach 25 years, who hold physically demanding and labour-intensive jobs and is paying premium for 40 years until his retirement at 65. He is contributing the premium to this fund for all 40 years between age 25 and 65, and his LTC benefits start when he is 68 years old which means 43 years after employment ($_{43}p_{25}$= 0.7491; $\ddot{a}_{25:\overline{65|}}$=28.255); based on the following other data:
1. Technical interest rate 0.0175;
2. Operating expenditures of fund management 5 %;

$$Pr\left(25,M,0.0175\right)=\alpha_{LTC,i}c_i=\frac{0.95\cdot v^{43}\cdot{}_{43}p_{25}\cdot \Delta NPV\, LTC}{1.05\cdot \ddot{a}_{25:\overline{65|}}}=\frac{0.95\cdot 0.47426\cdot 0.7491\cdot 54,407.53}{1.05\cdot 28.255}$$

$$=\frac{18,362.92}{28.255}=626.61€$$

Therefore, the monthly contribution to the additional LTC fund should be

$$\frac{\alpha_{LTC,i}c_i}{12}=52.22€$$

5 CONCLUSION

Most of the health care including nursing for older people is funded by the state, but the users themselves, i.e., households, finance the majority of social care also in the context of LTC. But for those who do physically demanding and labour-intensive work there should be incentive for managers and owners of the supply chain to retire them earlier and pay for them not only additional pension because of their physically demanding work, but also the additional expenditures for higher LTC, which are only results of this physically more demanding work activities, because they are entering LTC earlier and cost more. The second option for supply chain managers and owners is investment in improvement of ergonomic environment, especially supporting robots [1, 2], which would contribute to less physically demanding and less labour-intensive jobs.

References

[1] Battini D., Calzavara, M., Sgarbossa F., Persona A., 2017, MRP theory supporting trade-off between investments in collaborative robots and production in foreign countries for a water pumps supply chains, In: ZADNIK STIRN, L. (ed.), et al. *SOR '17 proceedings*, Ljubljana: SDI-SOR

[2] Bogataj, D., Battini D., Calzavara, M., Persona A. 2017. Investments in workplace ergonomics from the supply chain approach. *Proceedings 24th International Conference on Production Research*, DEStech Publications, In Press

[3] Bogataj, D., Bogataj, M. 2011. The role of free economic zones in global supply chains - a case of reverse logistics. *International journal of production economics*, 131(1): 365-371

[4] Bogataj, D. 2013. Pensions and home ownership in the welfare mix for older persons. In: ZADNIK STIRN, L. (ed.), et al. *SOR '13 proceedings*, Ljubljana: SDI-SOR, 2013, pp. 281-286

[5] Bogataj, D., Bogataj, M., Hudoklin, D. 2017. Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model. *International journal of production economics*, Vol.193: 51-62.

[6] Bogataj, D., Vodopivec, R., Bogataj, M., 2013. The extended MRP model for the evaluation and financing of superannuation schemes in a supply chain. *TEDE*, 19(S1): S119-S133

[7] Bogataj, D., Ros-McDonnell, D., Bogataj, M. 2015. Reverse mortgage schemes financing urban dynamics using the multiple decrement approach. *Springer Proceedings in Mathematics & Statistics*, Ser.No.135, Cham: Springer: 27-47

[8] Bogataj, D., Ros-McDonnell, D., Bogataj, M. 2016a, Management, financing and taxation of housing stock in the shrinking cities of aging societies. *Int..J. Prod. Econ.,* 181(A):. 2-13

[9] Bogataj, D., Aver, B., Bogataj, M., 2016b. Supply chain risk at simultaneous robust perturbations. *International journal of production economics,* 181(A): 68-78

[10] Bogataj, D., Ros McDonnell, D., Temeljotov Salaj, A., Bogataj, M. 2015. Sustainable Urban Growth in Ageing Regions: Delivering a Value to the Community. *Lecture Notes in Management and Industrial Engineering*. Cam. Springer; 215-224.

[11] Bogataj, D., Bogataj, M. 2015. Housing equity withdrawal in the portfolio choice for deferred pension income. *International journal of social sciences and humanities invention*, 2(7): 1459-1473

[12] Costa, F., J. 2011. *Reforming Long-term Care in Europe*. Blackwell Publishing Ltd, 3-14.

[13] European Commission. 2015. *The 2015 Ageing Report*. EE+ 3/2015. Brussels: EC-DG ECFIN.

[14] European Commission, Social Protection Committee. 2014. *Adequate social protection for long-term care needs in an ageing society*. Report. Brussels: European Commission services.

[15] Gerber, H. 1996. *Matematika življenjskih zavarovanj*, Ljubljana, DMFAS

[16] Kavšek, M., Bogataj, D. 2017a. Towards new quality standards of long-term care in Slovenia. *Revija za univerzalno odličnost,* 6(1): 11-24

[17] Kavšek, M., Bogataj, D. 2017b. Vpliv težkega fizičnega dela v aktivni dobi na neto sedanjo vrednost izdatkov za dolgotrajno oskrbo. *Revija za univerzalno odličnost*, 6(2): 98 -111

[18] Razširjeni strokovni kolegij za zdravstveno nego. 2013. *Modra knjiga standardov in normativov v zdravstveni in babiški negi ter oskrbi*. Ljubljana: Zbornica zdravstvene in babiške nege Slovenije – Zveza strokovnih društev medicinskih sester, babic in zdravstvenih tehnikov Slovenije in Sindikat delavcev v zdravstveni negi Slovenije.

[19] Pravilnik o metodologiji za oblikovanje cen socialno varstvenih storitev. 2006. *Uradni list RS*, št. 87/06

[20] SURS. 2007. *Popolne tablice umrljivosti prebivalstva Slovenije*. Ljubljana: Statistični urad RS

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Special Session 7:
Systems Optimization and Control with Applications

MULTIOBJECTIVE OPTIMIZATION FOR JOB SCHEDULING AT AUTOMATED CONTAINER TERMINALS

Anita Gudelj
University of Split/Faculty of Maritime Studies
Ruđera Boškovića 37, 21 000 Split, Croatia
E-mail: anita@pfst.hr

Maja Krčum
University of Split/Faculty of Maritime Studies
Ruđera Boškovića 37, 21 000 Split, Croatia
E-mail: mkrcum@pfst.hr

Mirko Čorić
University of Split/Faculty of Maritime Studies
Ruđera Boškovića 37, 21 000 Split, Croatia
E-mail: mcoric@pfst.hr

Abstract: The objective of this paper is traffic control and optimization for job scheduling of an Automated Guided Vehicle System which is embedded in a container terminal. The aim of AGV scheduling is not only reduces the cost of terminal operation but also maximizes the system performance. This study first formulates mathematical model which is focused on the optimization of job scheduling. The model considers two objectives (i.e., AGV traveling time by minimization vehicle waiting times for quay cranes involved) and their weighted sum is investigated as the representative example. In addition, the study is extended to seek optimal schedule in AGV system using multi-objective genetic algorithm which yields improvements in system throughput along with a decrease in the numbers of AGVs. The algorithm deals with multi-constrained scheduling problem with shared resources. The developed model is verified by a computer simulation using MATLAB environment. Results of simulation and experiments applied on real container terminal are presented.

Keywords: job scheduling, AGVs, multi-objective optimization problem

1 INTRODUCTION

Frequency of container carrying ships arriving at seaport container terminals has increased steadily during the past several decades [3]. The efficiency of container transfers must be as high as possible so as to reduce delays, transport time, and consequently, reduce the costs to terminal management. This requires efficient terminal equipment.
Automating container terminals are a trend in container terminal operations. Automated guided vehicles (AGVs) have become important transportation means in container transhipment systems. Increasing automation of yard vehicles has the potential to increase the productivity and the efficiency of container transport [4].

The moving of automated guided vehicles can described as discrete event system (DES) which includes set of discrete states and events. Conflicts and deadlocks are undesirable events in the system (even dangerous). It is necessary to implement supervisory policy which has to ensure that the process does not get into any of forbidden states (AGV collision or deadlock of vehicles). In [2] the authors modelled the container transportation system as a MRF1 class of Petri net with disjoint sets of resource and job places. Each container job involves the loading of a container onto the AGV, the movement of the AGV to the destination of the container, and the unloading of the container from the AGV. An AGV can be assigned one job at a time. MRF1PN has the matrix formulation which together with the PN marking transition equation provides the rigorous framework needed for analysis and simulation of DES. In this paper, we use that matrix formulation as input parameters in our

multi-objective genetic algorithm to seek optimal schedule in AGV system. Using matrices from MRF1PN, each chromosome is decoded in the priorities of the jobs, delay and release times. After that, set of chromosomes shall be forwarded to the schedule generator procedure which constructs parameterized schedules. Schedules are defined from generated firing sequence which has to be checked for conflict and deadlock existence. After a schedule is obtained, the corresponding measure of quality is a feedback to the genetic algorithm. The final goal is related to optimization of processing time by minimization vehicle waiting times for quay cranes. The determination the appropriate number of AGVs involved while maintaining the system throughput must be solved before this goal can be achieved. The algorithm was developed and tested using Matlab programming language.

2 OVERVIEW OF THE PROBLEM

Fig. 1 shows a layout of the seaport container terminal showing berth, quay cranes, AGV paths, AGV import area and nodes in the yard. This diagram is developed to model the actual container terminal of Port Koper (Slovenia). To design a highly efficient automated container terminal we proposed AGV system which can be deployed to transport containers within the terminal area. To increase the efficiency, two circular paths of the moving AGVs are introduced. The path A would be from the quay cranes to the stock crane on the yard area, to the rail station and return to the quay area. The second, an alternative route, path B, would be the from the quay cranes to the storage area and back to the quay side. The paths consist of several segments. Along the segment A, vehicles transport containers from the quay crane to a specific storage slot where the container is disposed of. Thereafter, the AGV drives along the segment B to a position where they will load a container from a stack and along the segment C it transported the container to the rail station. After unloading the container, empty vehicle goes along the segment D to the node where it waits for the next job. If the AGV vehicle doesn't transport a container to the railway, after unloading containers at the warehouse, an empty vehicle goes along the segment E to the parking slot. We only deal with a static transportation which assumes that the number of containers to be moved between two locations is determined at the beginning of the planning horizon, and travel times between locations as well as loading and unloading times are deterministic and known in advance. Each job involves the loading of a container onto the AGV, the movement of the AGV to the destination of the container, and the unloading of the container from the AGV. An AGV can be assigned one job at a time. After completing a job, an AGV can start another job.
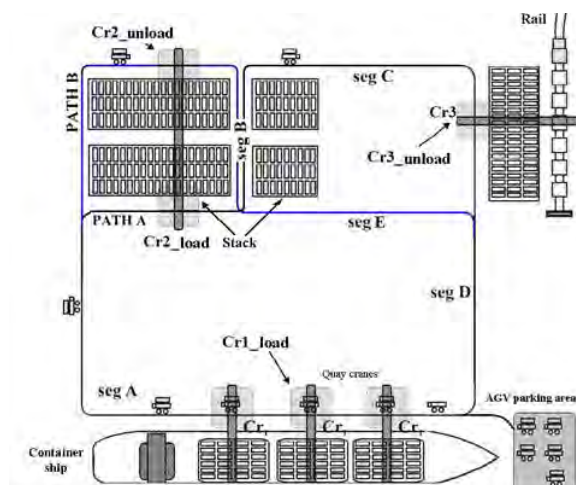
Figure 1: Layout diagram of the static container environment

2.1 Problem formulation

In the resource constrained multi-project scheduling problem, we consider a multi-project made up of a set of single projects $P = \{P_1, P_2, \cdots, P_m\}$. A single project $P_p \in P$ consists of a set of jobs $J_p \in \{j_0^p, j_1^p, \cdots, j_n^p, j_{n-1}^p\}$, including n real jobs and two dummy jobs as the start and finish jobs of the project, numbered from 0 to $n + 1$. A job j with duration p_j must not be interrupted once it has been started. While being processed, job $j \in J$ requires $r_{j,k}$ units of resource type resource $k \in \mathcal{R}$ in every time unit of its on-preemptible duration. Let $i_j \in J$ be the job performing by the resource $k \in \mathcal{R}$ just before performing the job j. It is necessary to take into the consideration the set-up time $s_{ij,k} \geq 0$ needed to prepare a resource $k \in \mathcal{R}$ for the job j, proceeded by the job i. Let V and K be the set of AGVs and the set of quay cranes QC, respectively. Set of resources $\mathcal{R}$ involves all cranes and segments. Resource type k has a limited capacity of $R(k)$ at any point in time. $A(t)$ is set of jobs in progress during the time period t.

Let α be travel cost per unit time of an AGV and β be penalty cost per unit time for the waiting time AGV for cranes. It is assumed that $\alpha \ll \beta$. Mathematically, the problem is formulated as follows:

$$min\left(\alpha \sum_{j \in J} F_j + \beta \sum_{l \in V} \sum_{k \in K}(y_j^k - e_{j_l}^l)\right) \tag{1}$$

subject to

$$\sum_{i=1}^{NcA} x_{ij} = 1, \qquad \forall j \in V \tag{2}$$

$$\sum_{j \in V} x_{ij} = 1, \qquad \forall i = 1, 2, \ldots, NcA \tag{3}$$

$$F_{i_j} + s_{i_j,k} + p_j \leq F_j, \qquad \forall j \in J, k \in \mathcal{R} \tag{4}$$

$$e_{i_j} + s_{i_j,k} + p_j \leq F_j, \qquad \forall j \in J, k \in \mathcal{R} \tag{5}$$

$$\sum_{j \in A(t)} r_{j,k} \leq RD(k), \qquad r_k \in \mathcal{R}, \ t \geq 0, \tag{6}$$

$$F_j \geq 0, \ RD(k) \geq 0, \ \forall j = 1, 2, \ldots, n, \quad k \in K, k \in \mathcal{R} \tag{7}$$

Where: y_i^k – the event time shows the beginning of a pickup a container from an AGV for a task related to the j-th job of k-th crane; NcA – the number of containers that can be loaded from the ship; F_j - finish time for a job j; x_{ij}^k - 1 if i-th container is assigned to j-th AGV by the crane k, else is 0.

The objective function (1) seeks to minimize the performance measure. During the iterative procedure of this algorithm, it is attempted to minimize the total travel time of AGVs, which is the first term of objective functions (1), and to minimize the sum of waiting times of the AGVs at the quay crane. The problem is to find solution vector $x_{i,j}$, $i=1,\ldots,$ $NcA, \forall j \in V$ and job scheduling such that constraint (2) - (8) are met. Constraints (2-3) ensure that only one vehicle is assigned to each unloaded container to transport it from berth side to the yard slot. Constraint (4) defines the precedence constraints for each pair of jobs (i, j) and $i \in P_j$ (a set of direct predecessors for each job j). The relation indicates that job i must be finished before the start of job j. Constraint (5) indicates that no job may be started before its predecessor is finished and the required setup time for preparing the corresponding of the

resource $k \in \mathcal{R}$ has elapsed. Constraint (6) takes care of the renewable resource limitations. Finally, constraint set (7) ensures that the job start times, finish times and resource limitations assume nonnegative integer values.

3   EXPERIMENTAL RESULTS

In this paper, the objective of the AGV schedule is to allocate a set of AGV vehicles to enhance the overall productivity of the system and reduce delays in a number of loading/unloading jobs using the multi-objective genetic algorithm. The final goal is normally related to optimization of processing time by minimization vehicle waiting times for quay cranes. In addition, the second aim is to reduce the total computational effort. In this sense, we consider AGV-container scheduling problem as a multi-project scheduling problem. In this case, each path (A or B) along which AGV can drive, will be considered as each project. Each project involves the set of jobs which have to be executed in predefined order.

3.1 Evolution strategy

Genetic Algorithm (GA) is adaptive search and optimization algorithm based on the principles of "survival of the fittest", where weak individuals die before reproducing, while stronger ones survive and bear many offspring and breed children. Given a current population, the reproduction, crossover and mutation operators are performed to obtain the next generation. The Fig. 2 depicts applied evolutionary strategy with selection, crossover and mutation operators developed in [1].

Reproduction is accomplished by first copying 20% of the best individuals from one generation to the next, in what is called an elitist strategy. The advantage of an elitist strategy over traditional probabilistic reproduction is that the best solution is monotonically improving from one generation to the next. The potential downside is population convergence to a local minimum. This can, however, be overcome by high mutation rates. The mutation consists of replacing 30% the worst parents of the existing population by newly generated chromosomes. The rest of the new population is created by the crossover operator which consists of exchanging the information of two parents of the existing population to produce two new chromosomes. The parameterized uniform crossover is employed with a probability equal to 0.75. Fig. 3 shows the comparative total computational effort, run times and optimal results for the problem *with different population sizes* using exactly the same operators for crossover and mutation and the same parameter settings [2]. Clearly, larger population produces a shorter makespan but also a longer runtime. As it can see on Fig. 3 a difference cannot be neglected because the runtime is 2434.344 sec which is twice faster than the runtime for *ga*50 algorithm. So in this paper, the population size is set to 20 individuals.
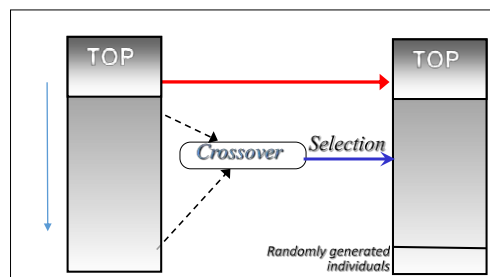
Figure 2: The reproduction process
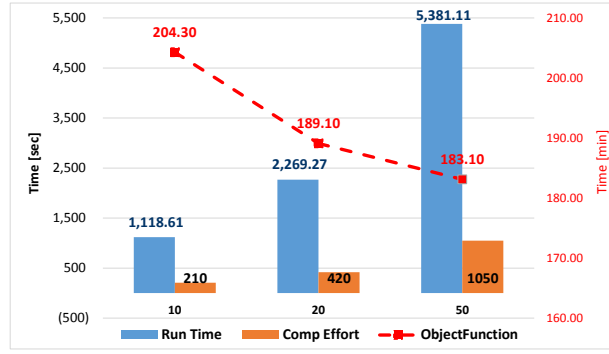
Figure 3: Effect of population size on the optimal solution

3.2 Chromosome Representation and Decoding

As in [2], we have used the indirect encoding, because the direct use of schedules as chromosomes was too complicated to represent and manipulate.

Encoding: A candidate solution can be encoded as a string consists of $m+2n+x$ real random numbers between 0 and 1 to form a chromosome, where n=14 is the number of jobs, m=2 the number of projects and x=8 the number of resources. The expression (8) shows one example of the chromosome.

$$\left(\underbrace{0.32, 0.77,}_{m \text{ genes}} \ \underbrace{0.52,\dots,0.66,}_{n \text{ genes}} \ \underbrace{0.42,\dots,0.68,}_{m \text{ genes}} \underbrace{0.91,\dots,0.85}_{x \text{ genes}} \right) \qquad (8)$$

Decoding: A feasible solution can be decoded from the given chromosome.

The first m genes are used to determine the priorities of each project. The genes between $m+1$ and $m+n$ are used to determine the delay time used at each of the n iterations of scheduling procedure. For this problem, the delay times are calculated by the following expression [1]:

$$DelayGen_g = gene_{m+g} \times MaxDur \times 1.5, \quad \forall g = 1,\dots, n \qquad (9)$$

where *MaxDur* is the maximum duration among all job durations.

The genes between $m+n+1$ and $2m+n$ are used to determine the release time of each project. The times are determined by

$$rt_i = ERD_i + gene_{m+n+i} \times \left(DD_i - ERD_i \right) \qquad (10)$$

where ERD_i is the earliest release date for each project i. The last x genes are used to determine the release dates of each resource:

$$RT_i = gene_{2m+n+i}, \quad \forall i = 1,\dots, x. \qquad (11)$$

3.3 Summary of the results

Initially, the authors assume that all containers have the same length of 40 feet. We have modelled the process of moving 30 containers in direction A, and 30 containers in direction B simultaneously. The average unloading time of quay cranes is 66 sec. The speed of a full AGV is 4 m/s and the speed of an empty AGV is 5.5 m/s. The weights of the objective function have been set to α = 0.9 and β= 0.1 which should be a reasonable choice in practice [5]. In case of conflict, random dispatching is used. The number of AGVs needed to minimize the total time for unloading all containers from vessels is limited to N vehicles due to the costs for operating as well as space on the terminal. In this study, we use nearest-vehicle-first dispatching rules (i.e. a free AGV at the smallest distance is dispatched to the

267

quay crane needing an AGV). Algorithm has been tested 10 times and the best solution is represented in Fig. 4.

As α<<β from Fig. 4 it is apparent that the waiting time has been minimized first because that objective function faster converges to the optimum than makespan as the second objective.

.
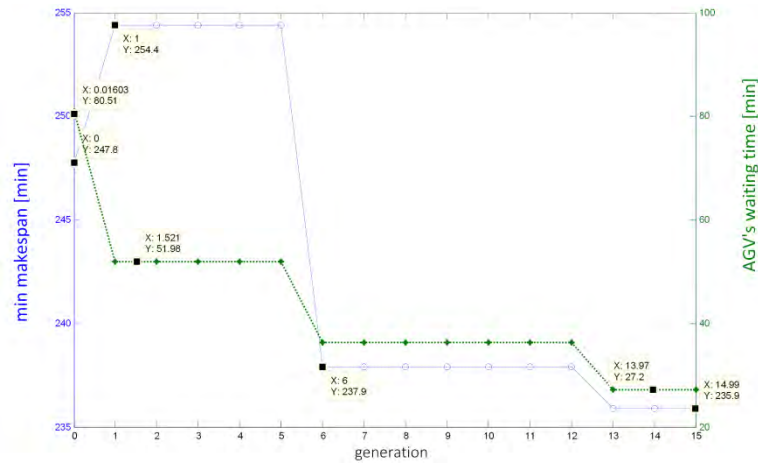
Figure 4: Relationship between total waiting time and total traveling time of AGVs per generations

The best (optimal) solution is 48.07 minutes, that is, 235.9 minutes the total time for finishing all jobs and total waiting time of AGVs 27.2 minutes. Certainly, this makespan isn't the shortest, but now the user has to decide what matters to him; maximize the use of resources and thus reduce the cost of resources, or reduce the time spent performing jobs business and thus allow the ship to leave the port as soon as possible. Also, the algorithm found that 11 AGVs are needed / required to fulfil all jobs.

Conclusion

In this paper, we have studied the problem of job scheduling problem for container transport. The main contribution of this paper is the development of rule-based methods for the AGV dispatching problem in seaport container terminals and their evaluation. In our multi-objective genetic algorithm MRF1PN matrix model of the system, developed in [2] as input parameters. Therefore there is no need to mathematically define any constraints, variables or supervisor which results in a substantial reduction of problem formulation complexity. Developed algorithm can be the main component of a decision support system for terminal management.

References

[1] Goncavles, J.F., Mendes, J.M., Resende, M.G.C. 2008. A Genetic Algorithm for the Resource Constrained Multi-Project Scheduling Problem. European Journal of Operational Research, 189: 1171–1190.

[2] Gudelj, A., Kezić, D., Vidačić, S. 2012. Planning and Optimization of AGV Jobs by Petri Net and Genetic Algorithm. JIOS, 36(2): 99-122.

[3] Skinner, B., Yuan, S., Huang, S., Liu, D., Cai, B., Dissanayake, G., Lau, H., Bott, A., Pagac, D. 2013. Optimisation for job scheduling at automated container terminals using genetic algorithm. Computers & Industrial Engineering, 64(1): 511-523.

[4] Vis, I.F.A. 2006. Survey of research in the design and control of automated guided vehicle systems. European Journal Operation Research, 170, 677–709.

IMPROVING EMERGENCY SYSTEM USING SIMULATION AND OPTIMIZATION

Ľudmila Jánošíková
University of Žilina, Faculty of Management Science and Informatics
Univerzitná 1, 010 26 Žilina, Slovak Republic
E-mail: Ludmila.Janosikova@fri.uniza.sk

Peter Jankovič
University of Žilina, Faculty of Management Science and Informatics
Univerzitná 1, 010 26 Žilina, Slovak Republic
E-mail: Peter.Jankovic@fri.uniza.sk

Marek Kvet
University of Žilina, Faculty of Management Science and Informatics
Univerzitná 1, 010 26 Žilina, Slovak Republic
E-mail: Marek.Kvet@fri.uniza.sk

Abstract: The paper deals with the optimal location of emergency stations where ambulances are kept. It describes a detailed computer simulation model that is used to evaluate indicators of the system performance assuming the station location is given. The model was calibrated using real data from a health care provider in the Slovak Republic. The weighted p-median model was used to relocate current number of ambulances. In spite of its simplicity, the model was able to reduce the average response time, to increase the percentage of calls responded to within a predetermined time limit, and to achieve a more balanced workload of ambulances in comparison to the current distribution of the stations.

Keywords: emergency medical service, computer simulation, ambulance location, p-median problem, discrete optimization

1 INTRODUCTION

The main role of an emergency medical service (EMS) system is to provide quick medical aid to people whose health conditions suddenly have got worse. Apart from the quality of treatment, response time, i.e. the time interval from the emergency call is received in the dispatch centre to arrival of the ambulance on scene, is a crucial factor influencing patients' outcome. Response time depends mainly on the number and location of ambulances operating in a given region but also on the management of a patient during pre-hospital care. The goal of this paper is to investigate the role of simulation and optimization techniques in improving EMS technology. The motivation came from Falck Záchranná, a.s. that is a commercial EMS provider in the Slovak Republic. They had some ideas on improving the EMS system and asked the university for their quantitative evaluation with the aim to obtain practical recommendations for decision makers. In this paper we focus on the strategic level of the urgent health care management and aim at determining the optimal location of emergency stations where ambulances are kept.

The paper is organised as follows. In Section 2 we describe the computer simulation model that was used for evaluation of the proposed changes in the EMS technology. Section 3 presents a mathematical programming model for ambulance relocation. Both models are applicable in a wide-spread region containing both urban and rural areas. Section 4 contains numerical results and their discussion. Section 5 provides concluding remarks.

2 SIMULATION MODEL

A detailed computer simulation model was built to evaluate the proposed changes in the current EMS technology. The model was calibrated using four sources of data:

1. publicly available statistics published by the Operation centre of the EMS of the Slovak Republic;
2. a sample of patients data from Falck Záchranná a.s.;
3. OpenStreetMap data about transportation network;
4. LandScan data about population distribution.

The Operation centre of the EMS of the Slovak Republic publish annual statistics about interventions of the ambulances and emergency helicopters. We used the latest statistics available for year 2015 [http://www.emergency-slovakia.sk/index.php/2015]. From this report we learn how many emergency calls the system received and served, and also how many patients were transported by ambulances between hospitals. Inter-hospital transports are important because they reduce the availability of ambulances.

Falck Záchranná a.s. supplied us with the data about 149 474 patients served in year 2015. This dataset contains information about the time and date of each incident, the location of the patient, the initial medical diagnosis, and time stamps of the whole EMS trip. Important knowledge can be extracted from the data. First of all, time distribution of patients can be revealed. As regards year seasons and weekdays, we did not observe statistically significant differences in the number of calls. Daily shares of calls ranges from 13.67% to 14.97%. This finding is in compliance with the situation in Mecklenburg County, USA published in the latest study [7]. The authors analyse 62 092 calls received in year 2004 and report approximately even distribution of calls during a week with daily shares ranging from 13.73% to 14.88%. A similar situation is reported by Marek et al. in [6] for the Olomouc region in the Czech Republic. The distribution of 45 037 calls (in 2009) ranges from 12.81% to 16.66% per day. In Slovakia and the Czech Republic, the demand is highest on Saturday.

However, call rate changes significantly during a day (see Fig. 1). We can observe two peaks, one between 9 and 11 am and the other between 5 and 9 pm. We have proven that the arrival of calls can be modelled as a non-homogeneous Poisson process with arrival rate varying by the time of day. The shortest inter-arrival time between calls is from 7 to 8 pm (46 seconds) and the longest inter-arrival time is from 4 to 5 am (145 seconds).
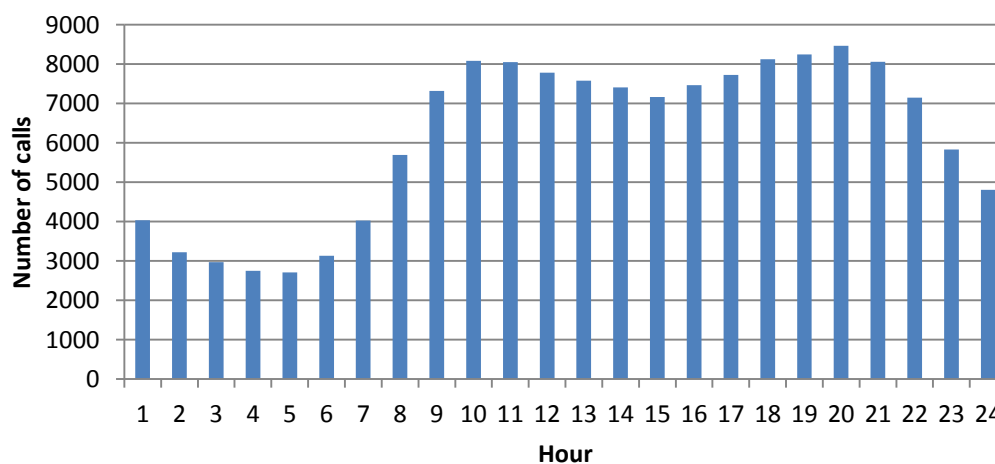
Figure 1: Demand distribution during a day

	Q1	Q2	Q3	Q4	Q5	Others
Number	9500	2348	2260	18472	7041	109853
%	6.36	1.57	1.51	12.36	4.71	73.49

The spatial distribution of patients is modelled using the LandScan database [http://web.ornl.gov/sci/landscan]. LandScan data represent an ambient population (average presence of people over 24 hours) on the area of 30″×30″. The territory of the Slovak Republic is covered by 70 324 grid elements. Grid zones are assigned to previously generated calls with a probability that is proportional to their population.

The patients are divided into 6 groups according to the initial diagnosis (see Table 1). The first five diagnoses are the most severe conditions in which immediate treatment is crucial. They are denoted as First Hour Quintet (FHQ) and include: chest pain, severe trauma, stroke, difficulties in breathing, and cardiac arrest. Although FHQ does not contain unconsciousness, it belongs to life-threatening conditions, therefore we include it into group Q5 in the statistics. The diagnosis is important to determine the priority of the call and to model the on-scene time. The probability distribution that is used to model the time of patient's treatment by the ambulance crew differs for every diagnostic group and the type of crew (crew with a paramedic or a physician). Analysis of real data revealed that the gamma distribution fits well when a physician is on scene. The on-scene time for a paramedic can be modelled by the gamma distribution in case of Q1 conditions and non-critical patients, the Erlang distribution for Q2, Q4 and Q5, and the Weibull distribution for Q3 conditions.

When an emergency call is received by a dispatch centre, the call is allocated to the nearest available ambulance. In case of FHQ patient, if the closest ambulance is a paramedic ambulance, the closest ambulance with a physician is dispatched simultaneously. The paramedic ambulance waits for the physician at the site of incident and then it becomes idle. When the patient is not transported to a hospital, the ambulance becomes idle immediately after emergency care on-scene was provided. When the ambulance is idle, it is available for next dispatch while travelling back to its original station.

The real data show that 77% of the patients treated by a paramedic crew and 51% of the patients treated by a physician are transported to a hospital. Currently patients are transported to 74 hospitals. The time the crew spends in a hospital passing the patient to the hospital staff is modelled separately for every hospital. Most often, the Erlang distribution fits well. The ambulance becomes idle after the patient was taken over by the hospital staff.

The call handling time in the dispatch centre is ignored in the current version of the simulation model because the relevant data were not provided by the Operation centre.

In the simulation model, ambulances travel along the road network. The digital road network was adopted from the OpenStreetMap database [https://www.openstreetmap.org] that is a freely available source of geographical data.

To calculate the shortest travel time the type of the road and speed limits recorded in the OpenStreetMap data are taken into account. If the road does not have a speed limit, an average speed is used obtained from the analysis of real trips.

Computer simulation enables to evaluate:
- average and median response time at different levels (municipality, county, region, state);
- service accessibility that is quantified as the percentage of rescue calls responded to within specified time limits;
- statistics about utilization of ambulances.

4 MATHEMATICAL PROGRAMMING MODEL FOR STATION RELOCATION

In the Slovak Republic, the number and locations of the EMS stations are defined by the Regulations of the Ministry of Health of the Slovak Republic No. 10552/2009-OL and 11378/2010-OL. The Regulations define 273 stations uniformly distributed on the whole state territory. The criterion for the distribution was to reach 95% of patients in 15 min or less regardless of their diagnoses. The stations are currently deployed in 211 towns and villages. They serve 2 928 municipalities with a total population of 5 410 827 (Dec 2012).

In this section, a mathematical programming model for EMS station location will be introduced. The model preserves the current number of stations and looks for a new station location.

Our previous research suggested that the weighted p-median problem could be a good model for ambulance location [5]. The goal is to find the location of a fixed number of p stations in order to minimize the average travel time of ambulances to potential patients. The average travel time is proportional to the total travel time to all demands and can be computed by dividing the total travel time by the number of all demands. Therefore instead of minimizing the average travel time one can minimize the total travel time needed to reach all potential patients. The objective function is a surrogate of the efficiency criterion which aims at providing the best possible level of service to as many people as possible with a limited number of resources. Besides efficiency, equity (or fairness) is a core performance dimension in a health care system. Fairness is achieved when each customer receives the service of required and/or acceptable quality. Although an equity principle is usually favoured by medical experts, efficiency is often applied by policy makers and Felder and Brinkmann [2] prove that it maximizes the number of lives saved.

Because stations are supposed to be deployed in a large-scale area, a macroscopic view must be applied. It means that demands must be aggregated and the whole village or city is considered as one demand zone. In accordance with Felder and Brinkmann [2], we suppose that the number of calls in a municipality is proportional to the number of its inhabitants. A municipality is not only a customer but at the same time it is a candidate for station location. Let us denote the set of candidate locations by I. Based on a previous study of geographical and demographical characteristics of the service areas we can conclude which stations can be relocated. Because we do not have information about workload of individual ambulances we suppose that the only factor affecting if the station can be relocated is the number of inhabitants in its current residence. It is supposed that an ambulance is able to serve an area with the population of $Q = 25\ 000$ [1]. If we denote the number of inhabitants in municipality j by symbol b_j and the current number of stations by symbol r_j, then three situations may occur:

1. If $b_j > r_jQ$, the stations cannot be relocated. The demand in municipality j will be $c_j = b_j - r_jQ$.
2. If $b_j > Q$ and at the same time $b_j < r_jQ$, b_j div Q stations must remain in the municipality, the others may be relocated. The demand in municipality j will be $c_j = b_j$ mod Q.
3. If $b_j < Q$ and $r > 1$, one station must remain in the town, the others may be relocated. The demand in municipality j will be 0.

After application of these rules, we have $p = 207$ stations that can be relocated. 1 511 039 inhabitants (28% of the whole population) are covered by an ambulance in their residence, the rest of population (3 899 788 people) represent the uncovered demand and input the model. Municipalities with uncovered demand constitute the set J of customers in the following mathematical programming model. Symbol t_{ij} states for the shortest travel time of an ambulance from node $i \in I$ to node $j \in J$.

The decision on opening a station must be done for each candidate location $i \in I$. This decision can be modelled by the binary variable y_i, which takes the value 1 if a station is located in node i, otherwise it takes the value 0. The assignment of municipality j to the station located in node i is modelled by binary variables x_{ij}. Variable x_{ij} takes value 1, if municipality j will be served by an ambulance located in node i, otherwise $x_{ij} = 0$. The model of the weighted p-median problem can be written as:

$$minimize \qquad \sum_{i \in I} \sum_{j \in J} t_{ij} c_j x_{ij} \qquad\qquad (1)$$

$$subject\ to \qquad \sum_{i \in I} x_{ij} = 1 \qquad for\ j \in J \qquad (2)$$

$$x_{ij} \leq y_i \qquad for\ i \in I,\ j \in J \qquad (3)$$

$$\sum_{i \in I} y_i = p \qquad\qquad (4)$$

$$x_{ij}, y_i \in \{0,1\} \qquad for\ i \in I,\ j \in J \qquad (5)$$

In the model constraints (2) ensure that every municipality j with uncovered demand will be assigned to exactly one station i. Constraints (3) ensure that if a municipality j is assigned to a node i, then a station will be open in the node i. Constraint (4) limits the total number of stations that can be sited. The remaining obligatory constraints (5) specify the definition domains of the variables.

The model was solved by radial approach [4] using commercial solver FICO Xpress 8.0.

4 COMPUTATIONAL EXPERIMENTS

The mathematical programming model relocates 105 (38%) stations and results in the average travel time of 3.09 min. This can be regarded as a lower bound of a real response time. However, it is very rough estimation, because the mathematical programming model is based on an implicit assumption that there is always an available ambulance to respond to a call. Moreover, the model concerns only the uncovered demand, and it is supposed that the travel time for people who have an ambulance in their residence is zero. Therefore computer simulation should be used to estimate performance characteristics of the proposed system more realistically. The results of simulation experiments are summarised in Table 2. The table contains selected quality indicators.

Table 2: Performance characteristics from simulation

	Current location (Dec 2016)	p-median	Difference
Average response time [min]	11:43	11:05	-0:38
% of calls responded within 15 min	74.1	77.9	3.8
Average ambulance workload [%]	35.15	35.01	-0.14
Coefficient of variation of ambulance workload [-]	0.26	0.21	-0.05

We can observe that the relocation of stations improves the quality of the system. It reduces the average response time by 38 seconds. The second row of the table indicates that currently the EMS system does not achieve the target to reach 95% of patients within 15 min. Only 74.1% of population have a good access to urgent health care. The coverage could certainly be improved by adding new ambulances. Supposing the number of stations is given, their

relocation using the *p*-median model can increase the number of calls reached in 15 min by 3.8%. Better location of the stations has an impact on the distances travelled by ambulances that result in lower utilization of the ambulances and more uniform workload. The coefficient of variation (CV), which is the ratio of the standard deviation of how busy ambulances are and the average busy probability, is used to calculate workload balance. Workload balance can be viewed as equity among crew members. The *p*-median model results in lower CV, i.e. more equitable distribution of workload among ambulances.

5 CONCLUSIONS

In the paper, the problem of EMS ambulance location is formulated as the weighted *p*-median problem. The model reflects preferences applied by policy makers in many European countries who prefer efficiency in EMS provision over equal access. Despite its simplicity, the model has proven to improve the EMS performance. The performance indicators were evaluated using a precise and detailed simulation model. The simulation model itself is not able to propose the best station location, however it is useful in the sensitivity analysis to answer the "what if" questions. It means it enables to experiment with changing system parameters such as the number and location of the EMS stations, demographic characteristics that influence the rate of calls in some regions and so on. Mathematical as well as simulation models can serve as decision supporting tools for an authority responsible for the emergency system design.

References
[1] Bahelka, M. (2008) Analysis of the emergency medical system after reform (Analýza systému záchrannej zdravotnej služby po reforme) [in Slovak] [online]. Bratislava: Health Policy Institute. Available from http://www.hpi.sk/hpi/sk/view/3795/analyza-systemu-zachrannej-zdravotnej-sluzby-po-reforme.html. [Accessed 3 March 2014].

[2] Felder, S., Brinkmann, H. (2002) Spatial allocation of emergency medical services: minimising the death rate or providing equal access? Regional Science and Urban Economics, 32, 27-45.

[3] Janáček, J., Kvet, M. (2015) Min-max optimization of emergency service system by exposing constraints. Komunikacie, 17(2), 15-22.

[4] Janáček, J., Kvet, M. (2016) Sequential approximate approach to the p-median problem. Computers & Industrial Engineering, 94, 83-92.

[5] Jánošíková, Ľ., Žarnay, M., Márton, P., Kvet, M. (2013) Models for location of emergency medical service stations and their comparison using computer simulation (Modely pre umiestnenie staníc záchrannej zdravotnej služby a ich porovnanie pomocou počítačovej simulácie) [in Slovak]. In Proceedings of the seminar Úlohy diskrétní optimalizace v dopravní praxi 2013. Pardubice, Czech Republic, Oct 28–29, pp. 52-61.

[6] Marek, J., Slováková, J., Hanousek, J. (2015) Number of EMS operations (Počet výjezdů IZS) [in Czech]. Forum Statisticum Slovacum, XI(4), 113-119.

[7] Zaffar, M.A., Rajagopalan, H.K., Saydam, C., Mayorga, M., Sharer, E. (2016) Coverage, survivability or response time: A comparative study of performance statistics used in ambulance location models via simulation-optimization. Operations Research for Health Care, 11, 1-12

PRODUCTION SCHEDULING OPTIMIZATION IN THE STEEL INDUSTRY USING GENETIC ALGORITHMS

Davorin Kofjač
University of Maribor, Faculty of Organizational Sciences
Kidričeva cesta 55a, Kranj, Slovenia
E-mail: davorin.kofjac@fov.uni-mb.si

Robert Rupnik
SIJ Acroni d.o.o.
Cesta Borisa Kidriča 44, Jesenice, Slovenia
E-mail: rupnik642@gmail.com

Alenka Brezavšček
University of Maribor, Faculty of Organizational Sciences
Kidričeva cesta 55a, Kranj, Slovenia
E-mail: alenka.brezavscek@fov.uni-mb.si

Abstract: This paper represents a practical application of the genetic algorithms (GA) to production scheduling optimization in steel industry. The thermal treatment process was analysed and used as a base for the development of the discrete event simulation model. The model was validated with the actual data. The model was then used as a fitness function for GAs, which were implemented in a form suitable for combinatorial optimization. Several combinations of GA operators were tested to find the (near) optimal schedule of the panels at the input of the thermal process to ensure minimal production flow time. Results have shown that the current process can be improved considerably by strictly adhering to the process norms. Furthermore, additional time savings can be achieved by GAs utilization.

Keywords: steel industry, production scheduling, optimization, genetic algorithms

1   INTRODUCTION

Scheduling problems arise frequently in industry [1]. The scheduling is one of the functions of the production planning and control system, where each task is placed with indication, in the time, of the work rank that will execute it [2]. The scheduling depends on the sequencing of the production that specifies the order in which the tasks must be executed. Some rules exist to define priorities at the moment of the sequencing, such as [2]: shorter time of processing, more urgent due date, shorter recess, critical reason, lowest set-up cost, physical restrictions etc. Due to many different and often contradicting constraints which must be taken into account while defining a feasible and, possibly, optimal schedule of the steel production process, scheduling in the steel industry has been recognized as one of the most difficult industrial scheduling problems [3].

As cited in [2], many techniques are used to get the best sequence of operations in different industries. Silva and Morabito [4] have developed a heuristic approach combined with the classic problem of the knapsack for the scheduling of production in a steel foundry. Toso and Morabito [5] have used linear programming to the production lots sequencing and sizing in a plant of rations, whereas Landmann and Erdmann [6] used a heuristic diffuse approach for the scheduling of the production in iron foundries.

In this paper we will focus on optimization of production process in a steel production company in Slovenia which is one of the leading producers of flat rolled steel products. The paper is organized as follows: the optimization problem under consideration is described in Section 2, the fundamentals of genetic algorithms which are used to search for (near) optimal solution are given in Section 3, the optimization results are presented and discussed in Section 4, while some concluding remarks are given in Section 5.

2   THE PROBLEM

The focus of our optimization is the process of thermal treatments on metal panels which is illustrated on Figure 1. It can be seen from Figure 1 the metal panels for thermal treatment are entering the process from the previous department. The panels enter the process randomly, irrespective of their thickness as well as thermal treatment needed. Batch of the panels (queue) is located in front of the process line which consists of one transport wagon, four furnaces for performing thermal treatments and one cooling chamber.
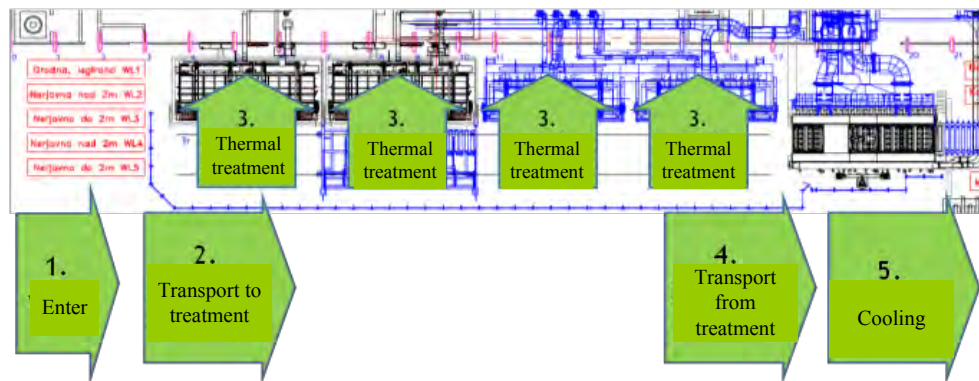
Figure 1: Production process of metal panel thermal treatment [7]

The duration of the thermal treatment is a random variable which depends on the panel chemical structure defining the treatment type (steeling, extinguishing, etc.), and on the panel thickness. These parameters also define the cooling time in the cooling chamber which is the final phase of the production process. Although the thermal treatments can be performed on all of four furnaces simultaneously, the problem appears due to a single transport wagon and a single cooling chamber. Unavailability of either wagon or cooling chamber results in an unexpected delay in production process leading to higher production cost. The aim of optimization is to find such schedule (combination) of panels waiting in the queue for the thermal treatment that would ensure minimal flow time of the production process.

3   GENETIC ALGORITHMS

To search for (near) optimal solutions of the panel scheduling problem, we will use the genetic algorithms (GA). The so-called combinatorial optimization with different schedules of the panels waiting in the queue for the thermal treatment will be used.

The genome of the GA will be formed as a series of panels waiting in queue for the thermal treatment. Genome therefore represents one of the potential schedules of panels. The length of the genome is equal to the number of panels that are in the input queue for the thermal treatment. An example of the genome is presented in Figure 2. Each number in the genome represents the identification (ID) of a particular panel. The example in
Figure 2 shows that the thermal treatment will first be conducted on the panel with ID 5, followed by the panel with ID 3, then ID 2 etc. The last panel to enter the thermal treatment will be the one with ID 23.

5	3	2	8	9	12	4	6	...	23

Figure 2: Genome representation where each number represents the panel ID

Performance assessment, i.e. fitness, of the individual genome will be carried out using a simulation model (Figure 3), which will result in the production flow time - the simulation model represents the fitness function. GAs will use operators which are adapted for combinatorial optimization:
- Roulette selection,
- Elitism,
- One-point permutational crossover,
- Swap mutation.
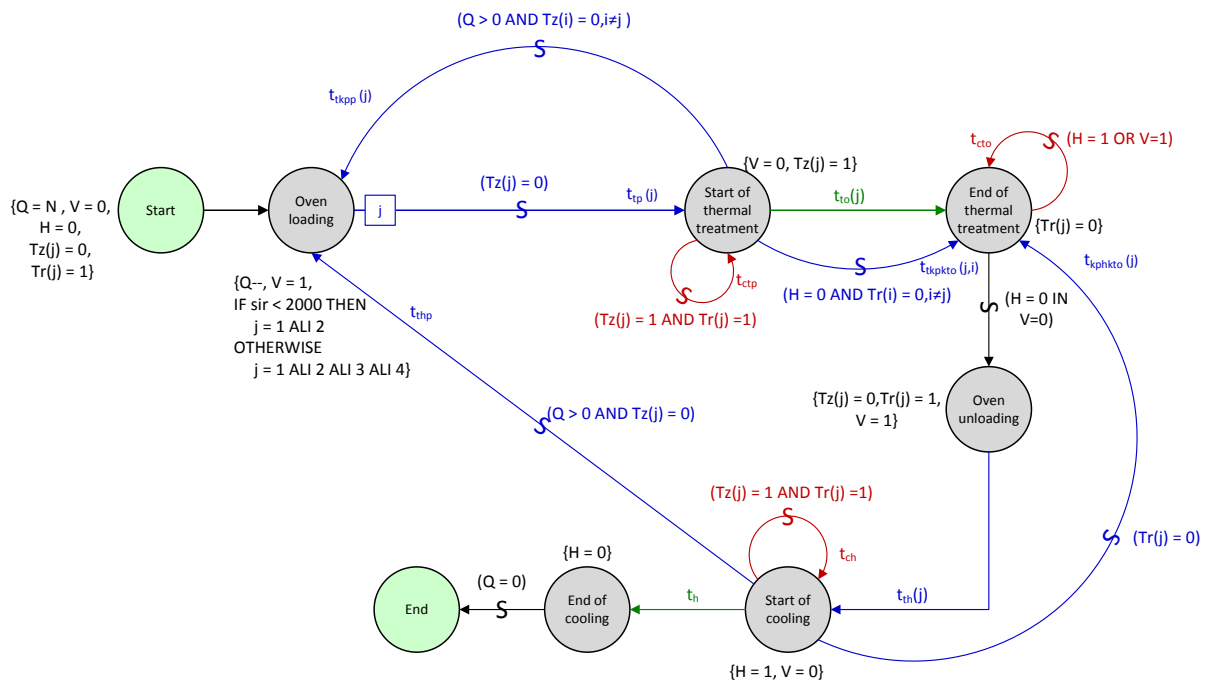
Figure 3: Simulation model of the process of metal panel thermal treatment [8]

4 RESULTS

The simulation model presented in Figure 3 was tested on actual data from the production for three consecutive months. The test data were obtained from the company.

We have presumed that situations where furnaces were filled with two panels at the same time can be treated as being filled with one panel, because two panels enter and leave the furnace at the same time.

For each test month, we performed tests on the data for the whole month, followed by tests where monthly data were separated into four parts (i.e., to yield approximately weekly data). Each test was repeated 10 times because of the stochastic nature of GA.

The simulation model was validated by comparing different parameters, namely:
- Actual flow time – this time was evaluated based on the actual data from the production. We calculated it as the difference between the time that the first panel in the present manufacturing series is loaded into the furnace and the time when the last panel of this series is unloaded from the cooling chamber.
- Simulation of actual flow time – this time was obtained by simulation, where the actual thermal treatment and the cooling times were considered on the actual schedule of panels. The difference between the actual flow time and its simulation time represents the so-called idle time in the process due to various interruption factors (pauses of employees, ...).

277

- Optimization of actual flow time – this time was obtained by simulation, where the actual thermal treatment and the cooling times were considered, while the schedule of panels was generated using GAs.
- Simulation of expected flow time – this time was obtained by simulation, where the expected thermal treatment and the cooling times were considered on the actual schedule of panels.
- Optimization of expected flow time – this time was obtained by simulation, where the expected thermal treatment and the cooling times were considered, while the schedule of panels was generated using GAs.

All the tests were carried out on a computer with an Intel Core2 Quad 2.4GHz with 8GB of memory, 256GB solid state drive with a 64-bit Windows 10 Pro.

4.1 Preliminary tests

Since we wanted to make quick tests on a relatively small sample of data, the preliminary tests were conducted on the data of the first quarter of the first month. Here the chromosome length was 673 genes. We have found that the most reasonable GA configuration is the one with 10% elitism, 0.75 crossover rate and 0.05 mutation rate.

We compared the simulation/optimization results to the actual flow time (Table 1). The results were normalized to preserve the anonymity of data and are presented as a ratio between the achieved simulation/optimization flow time and the actual flow time.

It can be seen from Table 1 that the time achieved by simulation of the actual flow time differs slightly in comparison to the actual flow time (only by 0.08%). This result indicates that the model is valid. The time achieved by optimization of the actual flow time again differs slightly in comparison to the actual flow time, probably due to the already optimized actual production schedule. Further results indicate possible time savings with simulation and optimizations of the expected flow time. It can be seen that different simulation/optimization scenarios (P:20, G:100), (P:20, G:50) and (P:20, G:20)[1] indicate possible time savings of 7.1%, 8.9%, 7.9% and 7.6%, respectively.

Table 1: Ratio between the simulation/optimization flow time and the actual flow time for different simulation/optimization scenarios [9]

Simulation/optimization	Ratio of simulation/optimization time and actual flow time
Simulation of actual flow time	1.008
Optimization of actual flow time	1.008
Simulation of expected flow time	0.929
Optimization of expected flow time (P:20, G:100)	0.911
Optimization of expected flow time (P:20, G:50)	0.921
Optimization of expected flow time (P:20, G:20)	0.924

The results of comparison of the simulation duration and time savings for different simulation/optimization scenarios are shown in **Error! Not a valid bookmark self-reference.**. The results of each simulation run were normalized in the form of the ratio between the achieved simulation time and the minimal simulation time achieved among all iterations within a particular scenario. As expected, the scenario (P:20, G:100) gives the best overall results, while under this scenario the time GA needed to find the solution is the longest among all

---

[1] P – number of individuals, G – number of generations

testing scenarios, namely 100 generations. Scenario (P:20, G:100) provides, compared to the (P:20, G:20), better result by 0.7%. If we compare the scenario (P: 20, G: 100) with respect to the actual production flow time, the time saving is almost 8.2%.

Table 2: Comparison of time savings for different optimization scenarios [9]

	P:20, G:100	P:20, G:50	P:20, G:20
Average simulation time [s]	3665	1687	645
Normalized average flow time [s]	1.005	1.008	1.012
St. deviation of normalized average flow time	0.004	0.003	0.003
Variability coefficient [%]	0.35	0.29	0.31
Savings vs. P:20, G:20 [%]	0.7	0.4	-
Savings vs. actual flow time [%]	8.2	7.9	7.5

Comparison of simulation results within different simulation/optimization scenarios are shown in Figure 2. The expected times achieved by simulation runs for a particular scenario were arranged in ascending order. It can be observed that the scenario (P:20, G:100) achieves the shortest flow time for all iterations, and from this viewpoint, is the most robust. However, the increased number of generations improved the obtained results, but it also increased the duration of the simulation. We believe that for the purposes of model validation, satisfactory results can be achieved with the scenario (P:20, G:50), wherein the duration of the simulation is about the half of that for the scenario (P:20, G:100) (see Table 2). Therefore, all further tests were performed considering the scenario (P:20, G:50).
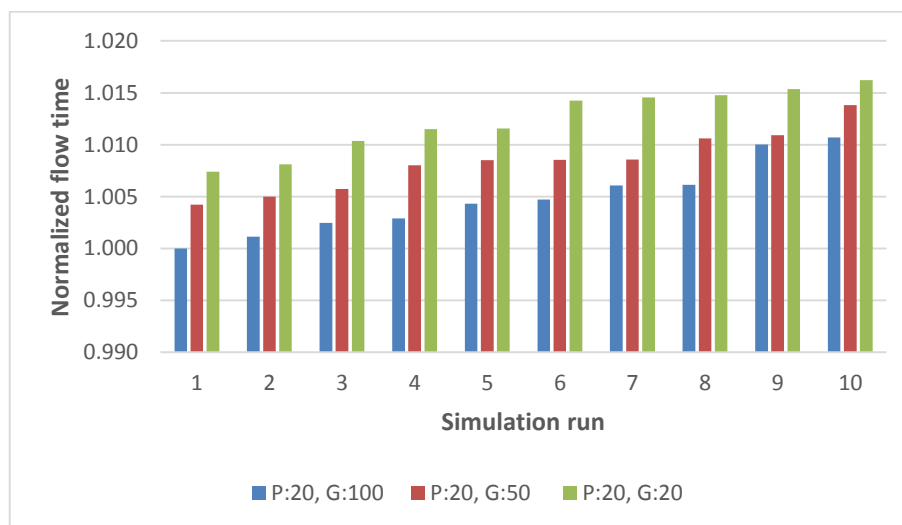
Figure 2: Comparison of 10 simulation runs for different simulation/optimization scenarios

4.2 Overall tests

Further tests were performed on a dataset of three consecutive months. The results are given in Table 3. The values in Table 3 represent the ratio between the simulation/optimization flow time and the actual flow time for three consecutive months. Here the chromosome length was 2694, 2655 and 2902 genes for months 1 to 3, respectively. Simulation of actual flow time shows the time difference of 5.5%, 3% and 1% regarding the actual flow time. Such little time differences indicate again the validity of the model. Optimizations of the actual flow time show the time difference of 5.6%, 3% and 2%. These time differences are again similar to those of the actual flow time simulations, probably due to the already optimized production schedule.

Higher time differences occur in simulation of the expected flow times (12.4%, 13% and 7%). These results indicate again that workers should strictly adhere to the pre-calculated processing times and avoid longer pauses causing long idle flow times. Further time savings could be achieved by optimizing the schedule with GAs (13.2%, 13% and 8%). However, optimization with GAs produces only slightly better results than simulation of expected flow time.

Table 3: Ratio between the simulation/optimization flow time and the actual flow for time three consecutive months (scenario P:20, G:50)

	Month 1	Month 2	Month 3
Simulation of actual flow time	0.945	0.97	0.99
Optimization of actual flow time	0.944	0.97	0.98
Simulation of expected flow time	0.876	0.87	0.93
Optimization of expected flow time	0.868	0.87	0.92

5   CONCLUSION

In the paper, the discrete event simulation model for optimization of scheduling of metal panels for thermal treatment in steel industry is presented. To find (near) optimal solution the GAs have been used. Validation of the model has been performed with the actual data for three consecutive months. Several simulation runs considering actual and expected times of thermal and cooling operations have been performed. Validation of the model has confirmed its stability and reliability. Furthermore, optimization results have shown that considerable improvements can be achieved even by strictly adhering to the process norms. Therefore, it would be worthwhile to reorganize the production process to run smoothly without unnecessary idle times. Additional time savings can be reached by GAs utilization. Since the production costs in steel industry are very high, in our opinion the effort of GAs implementation would be justifiable as well as cost effective.

References

[1] Ali, M.M., Kaelo, P., Ackerman, J. (2004). Scheduling of material through a steel plant. Mathematics in industry, Informative service. http://www.maths-in-industry.org/ miis/61/1/ MISGtplate.pdf [Accessed 12/June/2017].

[2] Landmann, R., et al. (2007). Production Scheduling Optimization in the Casting Industry using Genetic Algorithms. POMS 18th Annual Conference, Dallas, Texas, USA, May 4 to May 7, 2007.

[3] Biondi, M., Saliba, S., Harjunkoski, I. (2011). Production Optimization and Scheduling in a Steel Plant: Hot Rolling Mill. Preprints of the 18th IFAC World Congress Milano (Italy) August 28 - September 2, 2011.

[4] Silva, R. de J., Morabito, R. (2004). Otimização da programação de cargas de forno em uma fábrica de fundição em aço-inox. Revista Gestão e Produção. Vol(11): 135-151.

[5] Toso, E.A.V., Morabito, R. (2005). Otimização no dimensionamento e sequenciamento de lotes de produção: Estudo de caso numa fábrica de rações. Revista Gestão e Produção. Vol(12): 203-217.

[6] Landmann, R., Erdmann, R.H. (2006). A Heuristic Model for Production Scheduling in the Foundry Industry with the Employment of Fuzzy Logic. 17th Annual Conference of the Production and Operations Management Society. Boston, Massachusetts, USA, Apr 28 to May 1, 2006.

[7] Brezavšček, A., Kofjač, D., Rupnik, R. (2016). Optimiranje izkoriščenosti peči Wellman v obratu Predelava debele pločevine, Internal report of research project, 1st phase, University of Maribor, Faculty of Organizational Sciences, Kranj.

[8] Brezavšček, A., Kofjač, D., Rupnik, R. (2016). Optimiranje izkoriščenosti peči Wellman v obratu Predelava debele pločevine, Internal report of research project, 2nd phase, University of Maribor, Faculty of Organizational Sciences, Kranj.

[9] Brezavšček, A., Kofjač, D., Rupnik, R. (2016). Optimiranje izkoriščenosti peči Wellman v obratu Predelava debele pločevine, Internal report of research project, 3rd phase, University of Maribor, Faculty of Organizational Sciences, Kranj.

MULTI CRITERIA ASESSMENT OF APPLE CULTIVARS

Črtomir Rozman
University of Maribor Faculty of agriculture and life science
Pivola 11, 2311 Hoče, Slovenia
E-mail: crt.rozman@um.si

Tatjana Unuk
University of Maribor Faculty of agriculture and life science
Pivola 11, 2311 Hoče, Slovenia
E-mail: tatjana.unuk@um.si

Karmen Pažek
University of Maribor Faculty of agriculture and life science
Pivola 11, 2311 Hoče, Slovenia
E-mail: karmen.pazek@um.si

Stanislav Tojnko
University of Maribor Faculty of agriculture and life science
Pivola 11, 2311 Hoče, Slovenia
E-mail: Stanislav.tojnko@um.si

Mario Lešnik
University of Maribor Faculty of agriculture and life science
Pivola 11, 2311 Hoče, Slovenia
E-mail: Mario.lesnik@um.si

Abstract: Selection of a proper cultivar is one of the most important management decisions when investing into apple orchard. This paper presents a methodology for evaluation of suitability for cultivation of some apple cultivars by using a multi-criteria model based on analytical hierarchical process. The DEX method was used to support growers in making decisions about which new apple cultivar to grow. The model was applied to 13 cultivars with data derived from questionnaires completed by group of experts of different fields (orchardists, experts for fruit production economics and fruit marketing, cultivar breeders, experts of state service and commission for introduction of new cultivars, experts for fruit storage and fruit quality assessment, plant protection and fruit growing advisers). The results are shown as priority for individual cultivar. The potential of the model for assessing apple cultivar is demonstrated with the aim of providing a comprehensive explanation and justification of the assessment technique. It also indicates strong and weak points (in market potential, fruit estimation, production demands and tree estimation) of each assessed cultivar.

Keywords: apple cultivar, multicriteria model, expert group, DEX

1 INTRODUCTION

Establishing an apple orchard requires numerous management decisions. Since the orchard is a long-term investment, the selection of cultivars defines the difference between breaking even and making a profit (Lauer, 1995). Plant breeders test thousands of new cultivars for several years at many locations over a range of plant populations under different management practices. The results of cultivar evaluations are appreciated among producers. But the amount of information available can be daunting to producers and make their cultivar selection decision difficult. The choice of cultivars is thus one of the most important decisions leading to success or failure (Way, 1979).

Since cultivar selection is related to the assessment of different and, in some cases, conflicting attributes, MCDA emerges as a possible methodological solution (Rozman et al., 2009). The MCDA for cultivar assessment was first introduced by Srdjevic et al. (2009), who

developed a group multicriteria model for walnut cultivar assessment based on the analytic hierarchical process methodology, which resulted in quantitative cultivar assessment. A similar approach was used by Rozman et al. (2015) for the assessment of apple cultivars. In contrast, Pavlovic et al. (2011) proposed the DEX method, which uses discrete attribute values and utility functions in the form of if-then decision rules.

2 METHODOLOGY

We used standard multi criteria DEX methodology for model development (see Pavlovič et al., 2011). Following multicriteria hierarchy was used for the development of DEXi model.
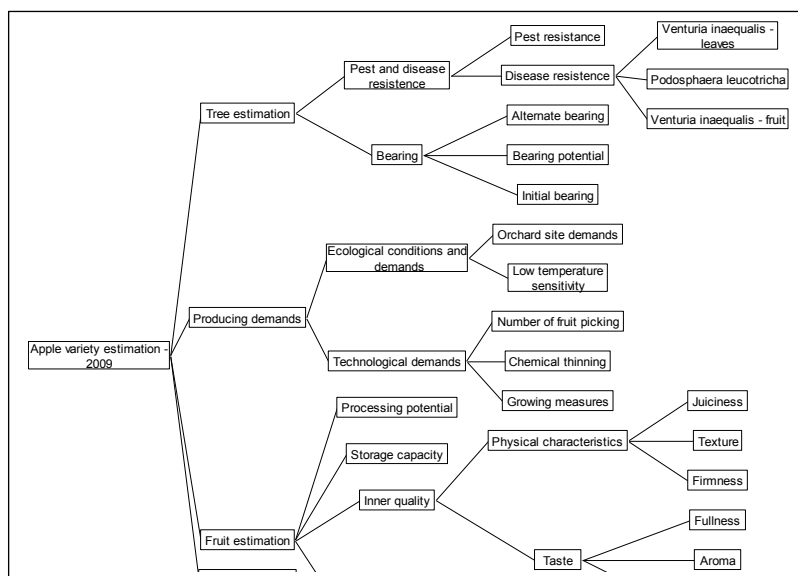
Figure 1: Hierarchy of apple cultivar assessment model

The assessment of basic attributes for each cultivar was conducted by the expert group, consisting of sixteen experts: Eleven experts had a scientific research background, and five had a production background, coming from major apple-producing companies (farms) in Slovenia. Each expert assessed the basic attributes for all thirteen varieties using the standard questionnaire. The assessment was conducted after a presentation of the DEXi methodology and principles. The scales of each basic attribute were then given their ordinal numbers. For instance, for the storage capacity attribute, the scales (figure 4.2) were defined as X={bad, good, excellent}. Thus, the corresponding ordinal values are as follows: ord (bad)=1, ord (good)=2, and ord (excellent)=3.

For each basic attribute, the average value was calculated in a spreadsheet as follows:

$$a = \frac{1}{n} \sum_{i=1}^{n} ord(x_i)$$

(1)

Where:
a - average ordinal value; x_i - ordinal value given by the expert I; n - number of experts (n=16).

The Excel function ROUND (with zero number of digits) was used in order to obtain integer values. Finally, the average ordinal value was transformed back to its corresponding discrete value, and the last was used as input in the DEXi model as the aggregated value of the group assessment. The data in figure 2 was transformed into a tab delimited file and imported into the DEXi model. A similar procedure was used for weights determination. Each expert was asked to fill out the questionnaire about attribute importance using a one through five assessment scale.

Figure 2: Calculated averages of ordinals

The described procedure was conducted in 2009 and 2015 using the same expert group and same varieties (cultivars).

3 RESULTS AND DISCUSSION

Fig. 3 and fig. 4 show assessment results for all varieties in 2013 and 2014.

The Fuji, Braeburn, Mairac, and Cameo 2015 assessments decreased to "unacceptable" in comparison to the 2009 assessments. Figure 4.19 shows that, after the experts experienced six years of growing those four cultivars, the cultivars' producing demands proved to be more difficult than was expected in 2009. Similar results were achieved for Greenstar. Its assessment decreased from "medium perspective" in 2009 to unacceptable in 2015. This can be attributed to a drop in the assessment of producing demands. For the Gala, Dalinbel, and Pinova varieties we can observe a decrease in the assessments (figure 4.21). As in the previous case, this can also be attributed to a decrease in the assessments of producing demands. Obviously, after the experts experienced those varieties for six years, it was clear that production is more demanding than was originally assumed. The assessment remained unchanged for the Golden delicious, Topaz, and Kanzi cultivars, although some changes in the assessment of subattributes did occur. The Opal cultivar is the only cultivar with an improved assessment (from "perspective" to "excellent") in 2015.
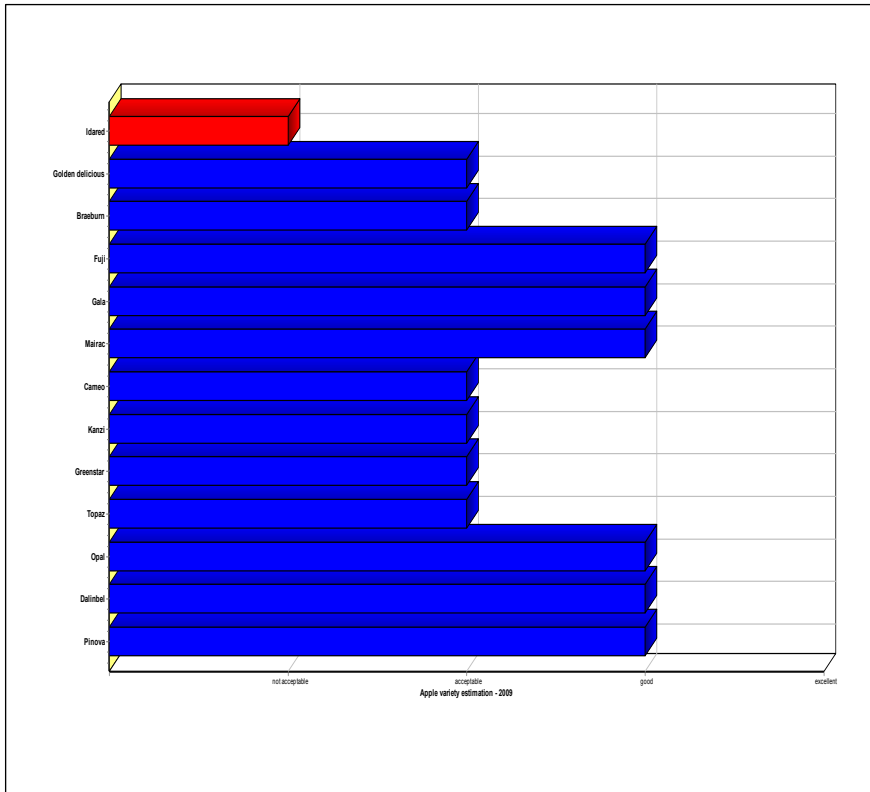
Figure 3: Assesment results for 2009
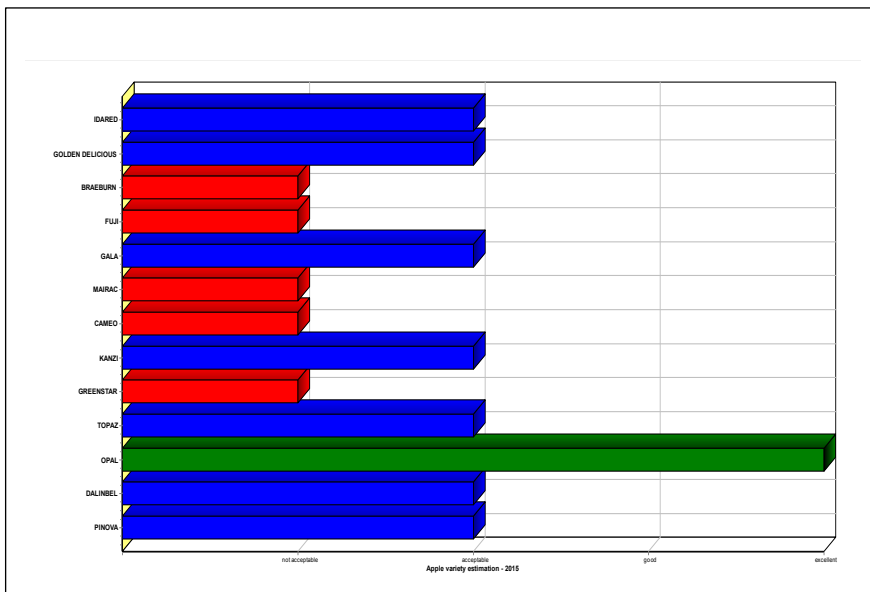
Figure 4: Assesment results for 2015

4 CONCLUSION

Selection of a proper cultivar is one of the most important management decisions when investing in an apple orchard. The software tool DEXi was used to support growers in making decisions about which new apple cultivar to grow. The model was applied to thirteen cultivars with data derived from questionnaires completed by a group of experts in different fields (orchardists, experts in fruit production economics and fruit marketing, cultivar

breeders, experts in state service and commission for the introduction of new cultivars, experts in fruit storage and fruit quality assessment, and plant protection and fruit growing advisers). The results are shown as suitability ratings for individual cultivars. The detailed analysis and explanations will be provided in upcoming book Decision support systems in Fruit production by Verlag dr. Kovač.

Acknowledgement

References

[1] Lauer, J. G. 1995. "Crop Variety Selection Software for Microcomputers." J. Prod. Agric., 8(3), pp. 327-437.

[2] Pavlovič, M., Čerenak, A., Pavlovič, V., Rozman, Č., Pažek, K., and Bohanec, M. 2011. "Development of DEX-HOP Multi-Attribute Decision Model for Preliminary Hop Hybrids Assessment." Computers and Electronics in Agriculture, 75(1), pp. 181-189.

[3] Rozman, Č., Hühner, M., Kolenko, M., Tojnko, S., Unuk, T., and Pažek, K. 2015. "Apple Variety Assessment with Analytical Hierarchy Process." Erwerbs-Obstbau, 57(2), pp. 97-104.

[4] Srdjevič, Z., Srdjević, B., and Suvocarev, K. 2009. "Objective Evaluation of Walnut Cultivars by the Analytic Hierarchy Process." In: 7th World Congress on Computers in Agriculture Conference Proceedings, American Society of Agricultural and Biological Engineers, p. 1.

[5] Way, R. D. 1979. "Apple Varieties Grown in New York State." New York's Food and Life Sciences Bulletin, 78, pp. 1-14.

SOLVING A PASSENGER FERRY FLEET ASSIGNMENT PROBLEM

Maja Škurić
University of Montenegro, Maritime Faculty Kotor
Dobrota 36, 85330 Kotor, Montenegro
E-mail: majaskuric@gmail.com

Vladislav Maraš
University of Belgrade, Faculty of Transport and Traffic Engineering
Vojvode Stepe 305, 11000 Belgrade, Serbia
E-mail: v.maras@sf.bg.ac.rs

Abstract: This paper deals with passenger ferry fleet assignment problem that takes into account ferry service operations inside the Boka Kotorska Bay, Montenegro. We focus on determining the optimal assignment of available and new proposed ferry fleet. An integer linear programming is applied in the case of owned and chartered fleet. Finally, we perform a numerical analysis based on test examples in order to evaluate the model. The obtained results show that with the adequate assignment of passenger ferry fleet inside the Boka Kotorska Bay, the costs of ferry service can be minimized.

Keywords: ferry fleet assignment, integer programming, numerical analysis, Boka Kotorska Bay

1 INTRODUCTION

The local maritime transport of passengers has been recognized as a very important mode of transportation especially in the regions that do not dispose with the adequate road infrastructure. Since the cruise industry rapidly increased in the Adriatic Sea and hundreds of thousand passengers are visiting the attractive places, in some of them is offered the local ferry transport in order to overcome the crowds and the bottlenecks in the road transport. This is only feasible by using passenger ferry service with the appropriate size of passenger ferry fleet, i.e. solving a ferry fleet assignment problem. In that manner, passengers from cruise ships would have the opportunity to use ferries and visit local places during their stay in port. In area such as Kotor cruise port located in Boka Kotorska Bay, Montenegro, passenger ferry service is in the phase of the concrete development since the demand for this mode of transport is increasing. The illustration of the study area is given in Fig. 1 where the cruise port is positioned as well as the other small ports at which ships can be serviced.

Figure 1: Passenger ferry service study area (Boka Kotorska Bay)

Simulation modelling of traffic in the Kotor cruise port for performance evaluation and optimization is evaluated in [8]. The authors proposed a scenario of an extended main berth in order to minimize congestion in Kotor bay and to maximize the port's revenue. This has been evaluated in [5]. On the other hand, ship emissions and their externalities with an emphasis to Dubrovnik and Kotor cruise ports have been presented in [4]. On the other hand, since the trend of increased traffic is noticeable, different assumptions of the study area has been investigated. According to the studies [13-15], the authors dealt with the potential for the development of ferry service inside the Boka Kotorska Bay having in mind the statistical data of the cruise ships throughput and demand for the passenger transportation. For example, the average number of passengers per cruise ship call in 2006 was 235; in 2011 it was 600, while in 2015 it was 1075.

Having in mind that passenger ferry fleet assignment problem can be formulated for a typical day, we have paid attention to two similar studies that have been investigated in [10, 11]. In these papers the aim is related to assign most appropriate fleet type to flights while minimizing the cost and determine optimal number of aircraft grounded overnight at each airport. Also, a review paper that explains industrial aspects of combined fleet composition and routing in maritime and road-based transportation is provided in [6]. The number of ferries needed to be chosen in the framework of ferry network design is studied in [9]. On the other hand, in [2] the authors stated the importance of ferry traffic to cruise ports. An and Lo [1] applied the methods to ferry service network design in Hong Kong. An integer programming model for the ferry scheduling problem has been done in [7]. Ceder and Sarvi [3] formulated the design of passenger ferry routes. A multi-fleet ferry routing and scheduling problem that takes into account ferry services with different operation characteristics and passengers with different preferred arrival time windows has been investigated in [16]. Yan et al. [18] developed several coordinated scheduling models that combine ferry company alliances, ferry fleet routing and timetable setting, ferry fleet size and related cost data. Hence, Winebrake et al. [17] explored technical solutions to reduce pollution from passenger ferries operating in the New York–New Jersey Harbor.

The structure of the paper is as follows. Proposed ferry fleet assignment model is presented in Section 2. Experimental analysis in the case of Boka Kotorska Bay is reported in Section 3 wile the next one contains results and discussion of the analysis. Section 5 gives concluding remarks.

2 PROBLEM STATEMENT

2.1 Model assumptions, sets, parameters and variables

In this paper we concentrate on solving passenger ferry fleet assignment problem in the reported bay using integer linear programming. The main goal is to determine the appropriate assignment of ferry fleet in order to minimize the total costs of the fleet. Since the assignment problems have been investigated a lot in the case of air transport, the related model has been adopted and modified from [10, 11]. For integer linear programming model, we use the following sets:

F - set of ferry trip,
R - set of ferry fleet type including owned and chartered ships,
P - number of ports in the bay network,
C - set of last port, representing all ports with ferry ships berthed overnight.

Indexes that are following the sets, parameters and decision variables are:
i - index of ferry trip,
j - index for ferry fleet type,

p - index for port.

Parameters included in integer linear programming formulation are:

c_{ij} - transportation cost of assigning ferry fleet type j to a designated trip i,

N_j - number of available ferry ships in fleet type j,

D_j - total capacity if ferry ships of type j,

d_{ij} - demand for trip i of ferry fleet type j,

S_{ip} - is equal to +1 if ferry trip i is an arrival at port p; is equal to -1 otherwise.

The decision variables are:

x_{ij} - is equal to 1 if ferry trip i is assigned to owned ferry fleet type j; otherwise is 0.

y_{ij} - is equal to 1 if ferry trip i is assigned to chartered ferry fleet type j; otherwise is 0.

W_{pj} - integer decision variable as a number of ferry ships of fleet type j at docks at port p.

The model formulation for calculating transportation costs c_{ij} is as follows:

$$\min \sum_{i \in F} \sum_{j \in R} c_{ij}(x_{ij} + y_{ij}) \tag{1}$$

s.t.

$$\sum_{j \in R} x_{ij} = 1, \quad \forall i \in F \tag{2}$$

$$\sum_{j \in R} y_{ij} = 1, \quad \forall i \in F \tag{3}$$

$$\sum_{i \in F} (x_{ij} + y_{ij}) \geq 1, \quad \forall j \in R \tag{4}$$

$$W_{p-1,j} + \sum_{i \in R} S_{ip}(x_{ij} + y_{ij}) = W_{pj}, \quad \forall p \in P \text{ and } \forall j \in R \tag{5}$$

$$\sum_{p \in C} W_{pj} \leq N_j, \quad \forall j \in R \tag{6}$$

$$\sum_{j \in R} d_{ij}(x_{ij} + y_{ij}) \leq D_j, \quad \forall i \in F \tag{7}$$

$$x_{ij}, y_{ij} \in \{0,1\}, \qquad \forall i \in F \text{ and } \forall j \in R \tag{8}$$

$$W_{pj} \in Z^+, \qquad \forall p \in P \text{ and } \forall j \in R \tag{9}$$

The objective function in (1) seeks to minimize the total costs of various ferry fleet types' assigned to all ferry trips inside the bay. Constraints (2 and 3) ensure that each ferry ship of owned and chartered fleet is to be assigned to exactly one ferry trip, respectively. Constraint (4) indicates that each trip is to be performed by one or more ferry ships (of owned or chartered fleet). Constraint (5) is a ferry balance constraint. It ensures that ferry ship of owned or chartered fleet of the right fleet type will be available. Constraint (6) represents the fleet size constraint that is a number of ferry fleet type which is less or equal to available ferry fleet type. Constraint (7) ensures that the total capacity of owned and chartered ferry fleet is enough for the demand of the passengers. Constraints (8) and (9) represent the binary and integer status of the decision variables (Z^+ is the set of positive integer numbers).

2.2 Specification of ferry costs

Here we divide transportation costs in two categories for the case of owned fleet; trip and operating costs. In the case of chartering ships, the company is paying the charter fee. Trip costs reported in this analysis are: wharfage fee, w_{ij}, costs for water supply, electricity and

waste removal, wew_{ij}, fuel costs, fc_{ij}, respectively, while operating costs are: crew costs, cc_{ij}, lubricant costs, lc_{ij}, repairs and maintenance costs, rm_{ij}, insurance costs, ic_{ij} and administrative costs $ad_{c_{ij}}$, respectively. Finally, the transportation costs are formulated as

$$c_{ij} = w_{ij} + wew_{ij} + fc_{ij} + cc_{ij} + lc_{ij} + rm_{ij} + ic_{ij} + ad_{c_{ij}} \tag{10}$$

3 EXPERIMENTAL ANALYSIS

In this analysis we propose three ferry trips in different days with the following data for test examples as specified in Tab. 1. The input data have been provided from [12] and obtained from the pilots and captains in Kotor cruise port. In our case of fleet assignment of the company engaged for the transport, there is one ship of 400 passenger capacity (FFT1), one with 200 seats for passengers (FFT2), one for 150 passengers (FFT3) and 18 smaller ships with total capacity of 900 passengers (FFT4-21) of which 14 are chartered, as it is real case in Boka Kotorska Bay. Also, ferry fleet can have a maximum one turnover in this test example.

Table 1: Test examples

Day	Trip no.	Demand for transport (Pax)	Description of trips
Day 1	1*,2*,3*	589	Trip 1*: Kotor cruise port – Perast – Risan – Tivat – Herceg-Novi – Kotor cruise port
Day 2	1*,2*,3*	1267	Trip 2*: Anchorage 1 – Perast – Risan – Tivat – Herceg-Novi – Anchorage 1
Day 3	1*,2*,3*	963	Trip 3*: Tivat's anchorage – Tivat - Kotor - Perast – Risan – Herceg-Novi – Tivat's anchorage
Day 4	1*,2*,3*	1647	

Applying the fleet assignment model, we provide the results of the real case situations where the ferry transport of passengers is realized with the different combinations of available owned and chartered ships (binary variable 0 for not employed ship and binary variable 1 for employed ship) (see Tab. 2). The calculation of total costs is also provided. As it can be noticed from the results of Tab. 2, the average costs per trip in the case of the first demand is 1545.67 €; in the case of second demand it is 3775.66 €; for third demand it is 2815.67 € and finally for the fourth demand it is 4960.00 €.

Table 2: Results of the fleet assignment model (the real case)

	Assignment							Costs/trip in €
	Trip no. 1*							
Demand	FFT1	FFT2	FFT3	FFT4-8	FFT9-13	FFT14-18	FFT19-21	
589 pax	1	1	0	00000	00000	00000	000	1573
1267 pax	1	1	0	11111	11111	11110	000	3813
963 pax	1	1	0	11111	11100	00000	000	2853
1647 pax	1	1	1	11111	11111	11111	111	4949
	Trip no. 2*							Costs/trip in €
Demand	FFT1	FFT2	FFT3	FFT4-8	FFT9-13	FFT14-18	FFT19-21	
589 pax	1	1	0	00000	00000	00000	000	1422
1267 pax	1	1	1	00000	11100	11111	111	3632
963 pax	1	1	1	10000	10000	10000	110	2672
1647 pax	1	1	1	11111	11111	11111	111	4752
	Trip no. 3*							Costs/trip in €
Demand	FFT1	FFT2	FFT3	FFT4-8	FFT9-13	FFT14-18	FFT19-21	
589 pax	1	0	0	11110	00000	00000	000	1642
1267 pax	1	0	0	11111	11111	11111	111	3882
963 pax	1	0	0	11100	11010	00111	111	2922
1647 pax	1	1	1	11111	11111	11111	111	5179

4 RESULTS OF NEW ASSIGNMENT AND DISCUSSION

The integer linear programming is solved in MATLAB 7.12.0. All tests were performed on i7 processor at 2.20 GHz with 8 GB RAM. The model has been validated and verified. The results of new ferry fleet assignment are given in Tab. 3 in the case with the new proposed fleet. We propose the combination of owned fleet (one ship of 400 passenger capacity (FT1); one ship of 250 passenger capacity (FT2); one ship of 200 passenger capacity (FT3); one ship of 150 passenger capacity (FT4); one ship of 140 passenger capacity (FT5)) and chartered fleet (one ship of 345 passenger capacity (FT6); one ship of 215 passenger capacity (FT7); one ship of 125 passenger capacity (FT8)). The chartered fleet is also including the charter fee.

Table 3: Results of the newly proposed ferry fleet configuration

Demand	Assignment Trip no. 1*								Costs/trip in €
	FT1	FT2	FT3	FT4	FT5	FT6	FT7	FT8	
589 pax	0	1	0	0	0	1	0	0	1559
1267 pax	1	1	0	1	0	1	0	1	3625
963 pax	1	1	0	0	0	1	0	0	2561
1647 pax	1	1	1	0	1	1	1	1	4815
Demand	Trip no. 2*								Costs/trip in €
	FT1	FT2	FT3	FT4	FT5	FT6	FT7	FT8	
589 pax	0	1	0	0	0	1	0	0	1309
1267 pax	1	1	0	1	1	1	0	0	3044
963 pax	1	1	0	0	0	1	0	0	2161
1647 pax	1	1	1	1	1	1	1	0	4194
Demand	Trip no. 3*								Costs/trip in €
	FT1	FT2	FT3	FT4	FT5	FT6	FT7	FT8	
589 pax	0	1	0	0	0	1	0	0	1464
1267 pax	1	1	0	0	0	1	1	1	3450
963 pax	1	1	0	0	0	1	0	0	2466
1647 pax	1	1	1	1	0	1	1	1	4747

The results are indicating that the average costs for the elaborated demand are less than those provided in the case of real situation. The optimization of the fleet assignment is provided because the costs are reduced by approximately 10% against the previous examples. The total revenue per each demand is supposed to be 4712 €, 10136 €, 7704 € and 19764 €, respectively. Obviously, the profit of a company is higher in the case of reduced ferry costs.

5 CONCLUSION

In this paper passenger ferry fleet assignment problem is investigated. The proposed integer linear programming model is applied on ferry service operations inside the Boka Kotorska Bay. We determine the assignment of ferry fleet types in order to minimize the costs of transportation. In experimental analysis, we proposed three ferry trips with a predefined passenger demand. The results indicated that the longest trip route and the biggest level of passenger demand, the highest costs per trip per passenger have been reported. But, on the other side, the employment of ferry fleet type is maximal which means that there is no loss due to unrealized transport of each passenger.

The main contribution of this investigation lies in the new approach of the ferry service employment in the mentioned area. Moreover, there is a lot of space for further investigations. These have to be directed to the detail analysis of transportation costs including the routing performances of ferry ships. Also, prediction analysis of new maritime

traffic is to be evaluated. Finally, the emission estimation of air pollutants considering ecological issues is not to be forgettable, especially because the bay is included in the natural, cultural and historical region of Kotor that is under the protection of UNESCO.

References

[1] An, K., Lo H.K. 2014. Ferry service network design with stochastic demand under use equilibrium flows. Transportation Research Part B, Vol 66: 70-89.

[2] Castillo-Manzano, J.I., Fageda, X., Gonzalez-Laxe, F. 2014. An analysis of the determinants of cruise traffic: An empirical application to the Spanish port system. Transportation Research Part E, Vol 66: 115-125.

[3] Ceder, A., Sarvi, M. 2007. Design and Evaluation of Passenger Ferry Routes. Journal of Public Transportation, Vol 10(1): 59-79.

[4] Dragović, B., Tzannatos, E., Tselentis, V., Meštrović, R., Škurić, M., 2015. Ship emissions and their externalities in cruise ports. Transportation Research Part D, dx.doi.org/10.1016/j.trd.2015.11.007

[5] Dragović, B., Škurić, M., Kofjač, D. 2014. A proposed simulation based operational policy for cruise ships in the Port of Kotor. Maritime Policy & Management, Vol 41(6): 560-588.

[6] Hoff, A., Andersson, H., Christiansen, M., Hasle, G., Lokketangen, A. 2010. Industrial aspects and literature survey: Fleet composition and routing. Computers & Operations Research, Vol 37: 2041-2061.

[7] Karapetyan, D., Punnen, A.P. 2013. A reduced integer programming model for the ferry scheduling problem. Public Transport, Vol 4(3): 151-163.

[8] Kofjač, D., Škurić, M., Dragović, B., Škraba, A. 2013. Traffic modelling and performance evaluation in Cruise Port of Kotor. Strojniški vestnik - Journal of Mechanical Engineering, Vol 59(9): 526-535.

[9] Lai, M.F., Lo, H.K. 2004. Ferry service network design: optimal fleet size, routing, and scheduling. Transportation Research Part A, Vol 38: 305-328.

[10] Ozdemir, Y., Basligil, H., Sarsenov, B. 2012a. A large scale integer linear programming to the daily fleet assignment problem: a case study in Turkey. Procedia - Social and Behavioral Sciences, Vol 62: 849-853.

[11] Ozdemir, Y., Basligil, H., Nalbant, K.G. 2012b. Optimization of Fleet Assignment: A Case Study in Turkey. An International Journal of Optimization and Control: Theories & Applications, Vol 2(1): 59-71.

[12] PoK (Port of Kotor). 2016. Port of Kotor's Business Report for 2015. Kotor: PoK.

[13] Škurić, M., Maraš, V. 2017. Some results of nautical risk assessment in port. International Maritime Science Conference. IMSC 2017, 20-21 April, Solin, Croatia, 335-339, ISSN 1847-1498.

[14] Škurić, M., Maraš, V. 2016a. Determining the size of ferry fleet: Fuzzy logic approach. In Maritime Technology and Engineering 3. Guedes Soares & Santos (Eds), Taylor & Francis Group, London, 117-121, ISBN 978-1-138-03000-8.

[15] Škurić, M., Maraš, V. 2016b. Analysis of the nautical risk assessment in passenger ports. International Conference on Traffic and Transport Engineering. ICTTE 2016, 24-25 November, Belgrade, Serbia, 333-338, ISBN 978-86-916153-3-8.

[16] Wang, D. Z.W., Lo, H. 2008. Multi-fleet ferry service network design with passenger preferences for differential services. Transportation Research Part B, 42: 798-822.

[17] Winebrake, J.J., Corbett, J.J., Wang, C., Farrell, A.E., Woods, P. 2005. Optimal Fleetwide Emissions Reductions for Passenger Ferries: An Application of a Mixed-Integer Nonlinear Programming Model for the New York–New Jersey Harbor. Journal of the Air & Waste Management Association, Vol 55(4): 458-466.

[18] Yan, S., Chen, C-H., Chen, H-Y., Lou, T-C. 2007. Optimal scheduling models for ferry companies under alliances. Journal of Marine Science and Technology, Vol 15(1): 53-66.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 1:
Econometric Models and Statistics

PURCHASING POWER PARITY IN CENTRAL AND EASTERN EUROPEAN COUNTRIES: AN ANALYSIS BASED ON NONLINEAR ROLLING KSS UNIT ROOT TEST

Jani Bekő

Department of Political Economy, Faculty of Economics and Business, University of Maribor
Razlagova 14, 2000 Maribor, Slovenia,
E-mail: jani.beko@um.si

Alenka Kavkler

Department of Quantitative Economic Analysis, Faculty of Economics and Business, University of Maribor
Razlagova 14, 2000 Maribor, Slovenia
E-mail: alenka.kavkler@um.si

Abstract: This paper investigates the Purchasing Power Parity (PPP) theory for a group of ten Central Eastern European countries covering the period from January 2001 to December 2016. We employ a rolling window nonlinear unit root test based on the exponential smooth transition autoregressive model. The results of unit root tests for the real exchange rate series indicate that PPP is not valid for the majority of subsamples of the selected set of countries. Additional empirical work is demanded in order to detect the factors that cause the violation of the PPP proposition in Central Eastern European countries.

Keywords: purchasing power parity, rolling KSS unit root test, CEE countries.

1 INTRODUCTION

The theory of Purchasing Power Parity (PPP) belongs to one of the most intensively investigated topics in international empirical economics. PPP basically suggests that shifts in exchange rates are primarily influenced by differences between foreign and domestic prices; under sufficiently competitive market conditions equalization of price levels should emerge. The collection of papers testing the PPP is incredibly opulent, though the validity of PPP remains subject of disputes in scientific arena. The empirics of PPP is usually focused on developed market economies (see, for example, Christidou and Panagiotidis, 2010; Huang and Yang, 2015; Kutan and Zhou, 2015), although the number of studies scrutinizing the PPP for transition countries is steadily increasing (see, for example, Yilanci, 2012; He *et al.*, 2013; Bahmani-Oskooee *et al.*, 2015; Jiang *et al.*, 2016). In general, the PPP tests for transition countries produced mixed results. He and Chang (2013) and He *et al.* (2013), using a new approach to panel unit root testing, provided mainly favourable evidence about the long-run PPP for this group of economies. Scrutinizing the nonlinear behaviour of real exchange rates, Jiang *et al.* (2016) are also able to support the PPP proposition for seven out of ten Central Eastern European countries. Contrary to these outcomes, even after considering sharp breaks and smooth shifts in real exchange rate dynamics of eight transition economies Bahmani-Oskooee *et al.* (2015) confirmed the PPP only in two cases.

In this paper, we empirically elaborate the concept of PPP by employing a rolling window nonlinear unit root test for the group of Central and Eastern European (CEE) countries: Bulgaria, Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Romania, Slovakia and Slovenia. The paper is divided into following five sections. After the introduction, section 2 gives a short description of selected macroeconomic indicators of Central and Eastern European economies. The econometric methodology and data used in the study are presented in section 3. The empirical results are given in section 4. Section 5 concludes the paper.

2 MACROECONOMIC DEVELOPMENT IN CEE COUNTRIES

The fundamental political and economic transformation of Central and Eastern European countries goes back in the early 1990s. The economic transformation of these countries comprised macroeconomic stabilization, privatization and microeconomic restructuring, foreign trade liberalization and the beginning of institutional and legal integration with European Union. The macroeconomic stabilization envisaged primarily inflation control as well as monetary and currency reforms. After initial devaluation of their currencies, the European transition countries experimented with various exchange rate regimes. The combination of employing different exchange rate arrangements with a diverse pace of disinflation caused significant macroeconomic adjustments costs in individual countries, but resulted in the targeted price stability. The increasing competition from foreign trade and the deregulation of domestic prices coupled with supply side type restructuring of companies and labour markets triggered the much-needed economic growth in this region. The process of catching-up with developed market economies remains paramount for Central and Eastern European countries, although the figures in Table 1 testify that the economic development of these economies appears to be gradual at best. The levels of GDP per capita of ten Central and Eastern European countries are placed below the EU average. The GDP per capita of the Czech Republic – country with the highest reference figure – reaches 88% of the EU level, whereas Bulgaria's level of development is more than 50% below the EU average.

The average GDP growth rates in the 2003–2007 period were at least around 5% in the observed sample of countries. The notable exception is Hungary with the average yearly GDP growth amounting to 3.5% (Table 1). Although all ten countries exhibit positive output growth in 2016, the post-crisis growth rates are lower compared with the GDP dynamics in the 2003–2007 period.

Table 1 indicates that inflation has been considerably higher in 2003–2007 than after 2014. In fact, Central and Eastern European countries faced extremely low inflation rates since the outbreak of Great Recession, even years of deflation are evidenced in Bulgaria, Croatia, Lithuania, Romania, Slovakia and Slovenia.

Table 1: GDP and inflation development in CEE countries

Country	GDP p.c. (EU28=100) 2016	GDP growth (in %)				HICP inflation (in %)			
		2003-2007 average	2014	2015	2016	2003-2007 average	2014	2015	2016
Bulgaria	48	6.6	1.3	3.6	3.4	5.9	-1.6	-1.1	-1.3
Croatia	59	4.7	-0.5	1.6	2.9	2.7	0.2	-0.3	-0.6
Czech Republic	88	5.5	2.7	4.5	2.4	1.8	0.4	0.3	0.6
Estonia	74	8.2	2.8	1.4	1.6	3.9	0.5	0.1	0.8
Hungary	67	3.5	4.0	3.1	2.0	5.4	0.0	0.1	0.4
Latvia	65	9.9	2.1	2.7	2.0	6.5	0.7	0.2	0.1
Lithuania	75	8.6	3.5	1.8	2.3	2.5	0.2	-0.7	0.7
Romania	59	6.6	3.1	3.9	4.8	9.5	1.4	-0.4	-1.1
Slovakia	77	7.3	2.6	3.8	3.3	5.0	-0.1	-0.3	-0.5
Slovenia	83	4.7	3.1	2.3	2.5	3.6	0.4	-0.8	-0.2

Source: Eurostat 2017; EC 2017.

From Table 2 it can be seen that with the exception of Slovenia all the remaining European transition countries recorded substantial current account deficits in the 2003–2007 period. In the post-crisis years, reduction of domestic consumption, improvements in export

competitiveness and stronger external demand caused a massive rebalancing in the foreign trade positions of these economies. As a result, in 2016, only Lithuania and Romania reported a slight current account deficit.

Statistical data in Table 2 also reveal that fiscal consolidation is under way in the majority of European transition countries producing sizeable reductions of general government deficit in Bulgaria, Croatia, the Czech Republic, Hungary and Slovenia (Table 2). Improvement of general government balances can be attributed to the restrictive fiscal policy in individual countries and to a more vigorous output growth.

Table 2: External and fiscal balance in CEE countries

Country	Current account balance (% of GDP)				General government balance (% of GDP)			
	2003-2007 average	2014	2015	2016	2003-2007 average	2014	2015	2016
Bulgaria	-12.8	0.0	0.4	4.2	1.1	-5.5	-1.6	0.0
Croatia	-5.9	1.1	5.0	2.6	-3.9	-5.4	-3.4	-0.8
Czech Republic	-3.9	-1.2	-1.2	0.3	-3.0	-1.9	-0.6	0.6
Estonia	-12.6	1.0	2.1	2.0	2.2	0.7	0.1	0.3
Hungary	-8.3	2.0	3.1	5.0	-7.1	-2.1	-1.6	-1.8
Latvia	-15.3	-2.0	-0.8	1.9	-0.8	-1.6	-1.3	0.0
Lithuania	-9.3	3.8	-2.2	-1.1	-0.8	-0.7	-0.2	0.3
Romania	-9.2	-0.1	-0.6	-2.4	-1.7	-1.4	-0.8	-3.0
Slovakia	-7.1	0.6	0.1	0.2	-2.7	-2.7	-2.7	-1.7
Slovenia	-2.5	6.2	5.4	7.0	-1.4	-5.4	-2.9	-1.8

Source: Eurostat 2017; EC 2017.

3 PRESENTATION OF THE METHODOLOGY AND DATA

Kapetanios et al. (2003) developed a test for the null hypothesis of unit root against the alternative hypothesis of a nonlinear stationary smooth transition autoregressive (STAR) model. The researchers attempted to distinguish between the nonstationary linear processes and the stationary nonlinear ones. The motivation for the development of the new test lies in the persistent failure of the standard ADF test to reject the null of a unit root. Kapetanios et al. (2003) extended the ARMA framework by analyzing a particular kind of nonlinear dynamics under the alternative hypothesis, namely exponential smooth transition autoregressive (ESTAR) models. The smooth transition autoregressive (STAR) model of order 1 is given by the equation

$$y_t = \beta y_{t-1} + \beta^* y_{t-1} G(\gamma, c; y_{t-d}) + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \quad d \geq 1, \tag{1}$$

where β and β^* are unknown parameters and ε_t is a sequence of independent identically distributed errors. Initially, y_t is assumed to be a zero-mean process, but the framework can easily be extended to include more general processes with non-zero mean and time trend. G represents a continuous transition function bounded between 0 and 1. The slope parameter γ is an indicator of the speed of transition between 0 and 1, whereas the threshold parameter c points to where the transition takes place. y_{t-d} is the transition variable and stands for the variable y lagged d times. The most popular functional forms are the Logistic Smooth Transition Autoregressive (LSTAR) form with logistic transition function and ESTAR with exponential transition function. The LSTAR transition function is monotonously increasing,

while ESTAR is U-shaped around *c* and thus enables reswitching. The ESTAR functional form can be defined as

$$G(\gamma, c; y_{t-d}) = 1 - \exp\left(-\gamma(y_{t-d} - c)^2\right).$$ (2)

Kapetanios *et al.* (2003) applied the ESTAR transition function with *c* equal to zero. By substituting G in equation (1) with the ESTAR transition function from equation (2), we obtain the ESTAR model

$$y_t = \beta y_{t-1} + \beta^* y_{t-1}\left[1 - \exp\left(-\gamma \cdot y_{t-d}^2\right)\right] + \varepsilon_t.$$ (3)

The null hypothesis of unit root implies $\gamma = 0$, since $G(0; y_{t-d}) = 0$, and is tested against the alternative hypothesis of $\gamma > 0$. Kapetanios *et al.* (2003) derived the limiting nonstandard distribution of the test statistic that involves Brownian motion.

Our sample consisted of the following Central Eastern European countries: Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Romania, Slovakia and Slovenia. The monthly averages of nominal exchange rates and consumer price indices were obtained from the European Central Bank and from Eurostat. The Euro-based real exchange rates comprised the period from January 2001 to December 2016. For all countries in the sample, the consumer price indices referred to year 2015.

4 EMPIRICAL RESULTS

Following Yilanci (2012), we employed rolling KSS unit root test approach with a fixed length window of 60 observations (5 years). The first window thus spans from January 2001 till December 2005 and the second window from February 2001 till January 2006. Thus, 133 windows (subsamples) are obtained. The results are presented graphically in Figure 1. The x-axis shows the end month of the window and the y-axis the scaled KSS statistic, namely the test statistic divided by the 5 % critical value of -2.93. Thus, for the subsamples with scaled statistic over 1, the null hypothesis of unit root is rejected and PPP holds. As can be seen from Figure 1 below, the patterns of PPP validity differ significantly, with some similarities across countries. For example, for the 5-year subsamples ending in the years 2007 and 2008, when the financial crisis began, PPP does not hold in any of the countries. On the other hand, for most of the countries in our sample, there are subsamples ending in 2012 and/or 2013 with valid PPP.

298

Czech Republic

Estonia

Hungary

Latvia

Lithuania

Romania
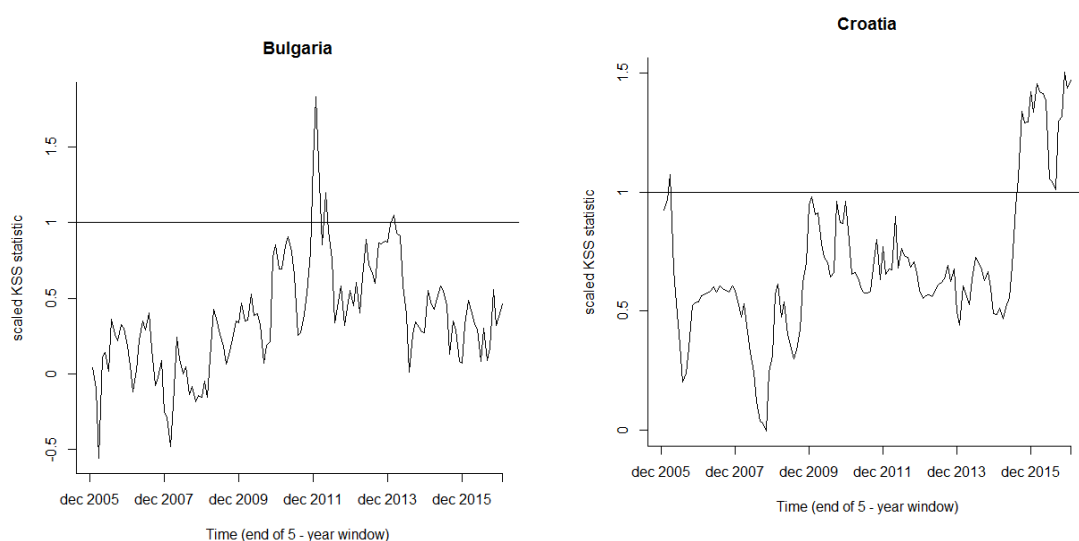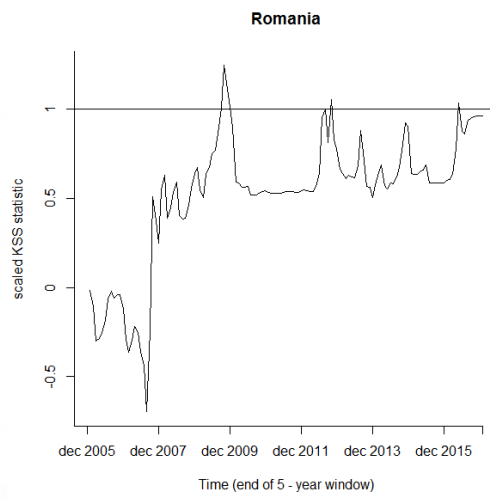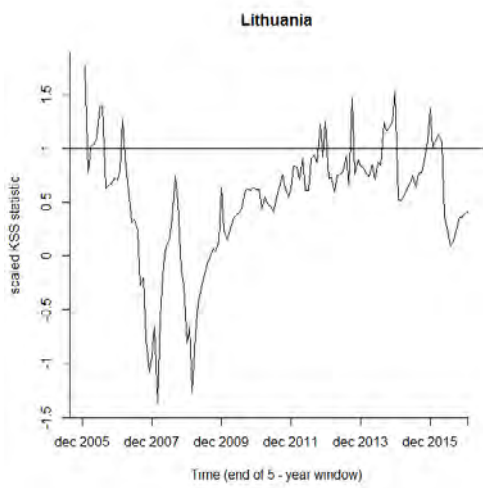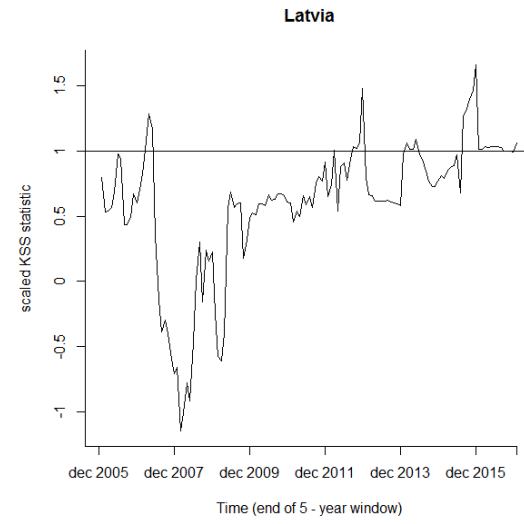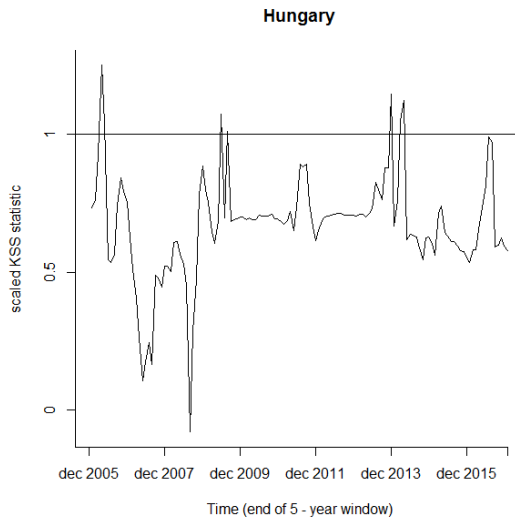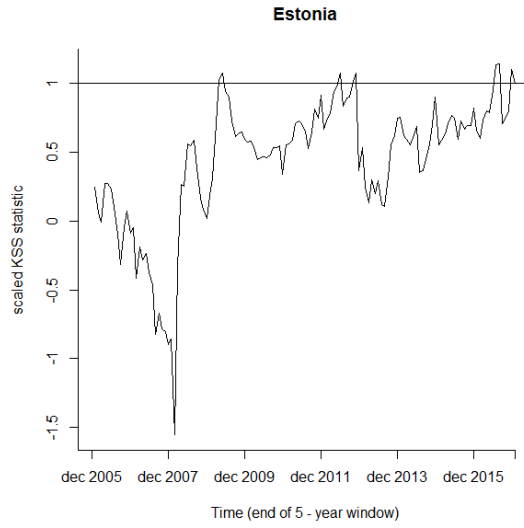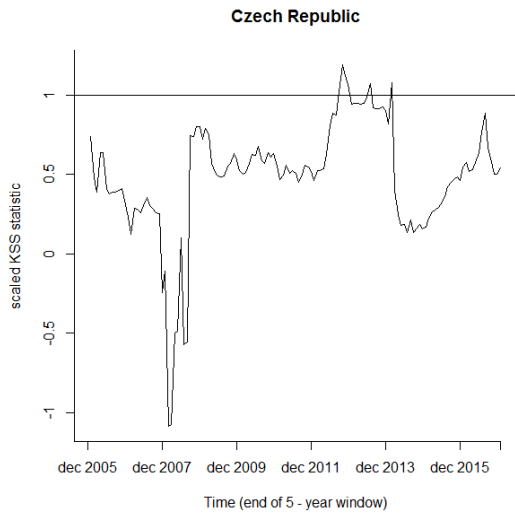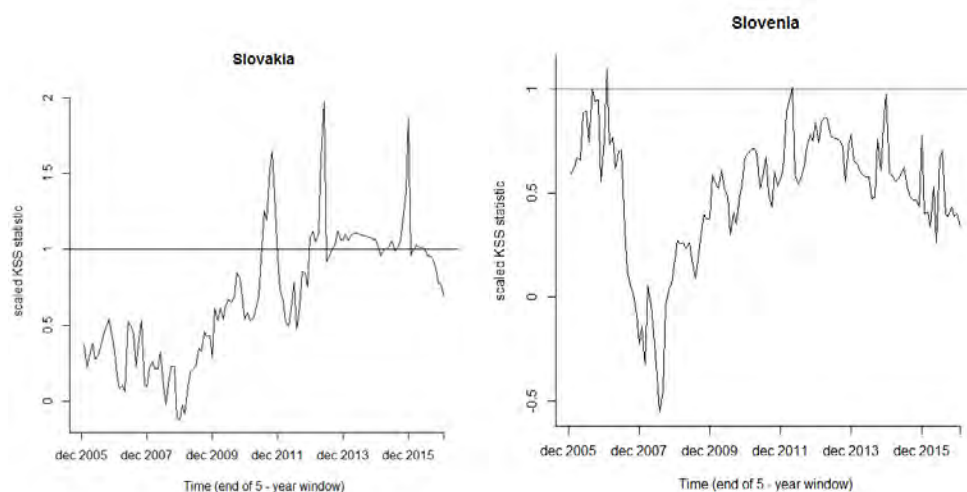
Figure 1: Rolling window KSS unit root test results

We also summarized the rolling KSS test results in Table 3 below, similarly as Yilanci (2012). For each country, the total number of samples, the number of stationary samples and the percentage of stationary samples is shown. The countries with the highest percentage of stationary samples (and PPP validity) are Slovakia (31.58%), Latvia (21.05%) and Lithuania (15.79%), whereas for other seven countries the ratio lies below 15%. The countries with less than 4% of stationary subsamples are Slovenia (1.50%) and Bulgaria (3.76%).

Table 3: Summary of rolling window test results

Country	Total no. of subsamples	No. of stationary subsamples	Percentage of stationary subsamples
Bulgaria	133	5	3.76
Croatia	133	19	14.29
Czech Republic	133	6	4.51
Estonia	133	9	6.77
Hungary	133	7	5.26
Latvia	133	28	21.05
Lithuania	133	21	15.79
Romania	133	6	4.51
Slovakia	133	42	31.58
Slovenia	133	2	1.50

5 CONCLUSION

In this paper, we examined the PPP concept by testing the stationarity properties of Euro-based real exchange rates for 10 Central Eastern European countries. The obtained results of the nonlinear rolling window KSS test clearly suggest that even after taking into account the nonlinear reversion of real exchange rates of selected countries we were not able to confirm the validity of PPP for the majority of subsamples of the observed countries. Similarly as Yilanci (2012), we found support for PPP only in sections of the observed time period, although countries with the highest percentage of stationary samples in our research are Slovakia, Latvia and Lithuania, and not Romania and Bulgaria, as it is reported in Yilanci

(2012). We can infer that additional empirical work is demanded in order to detect the factors that cause the violation of the PPP concept in Central Eastern European countries.

References

[1] Bahmani-Oskooee, M., Chang, T. and Wu, T-P. (2015). Purchasing power parity in transition countries: Panel stationary test with smooth and sharp breaks. *International Journal of Financial Studies*, 3, 153–161.

[2] Christidou, M. and Panagiotidis, T. (2010). Purchasing power parity and the European single currency: Some new evidence. *Economic Modelling,* 27, 1116–1123.

[3] He, H. and Chang, T. (2013). Purchasing power parity in transition countries: Sequential panel selection method. *Economic Modelling*, 35, 604–609.

[4] He, H., Ranjbar, O. and Chang, T. (2013). Purchasing power parity in transition countries: Old wine with new bottle. *Japan and the World Economy*, 28, 24–32.

[5] Huang, C. H. and Yang, C. Y. (2015). European exchange rate regimes and purchasing power parity: An empirical study on eleven eurozone countries. *International Review of Economics and Finance,* 35, 100–109.

[6] Eurostat (2017). http://ec.europa.eu/eurostat/web/products-datasets/-/tec00114.

[7] EC-European Commission (2017). European Economic Forecast. *Institutional Papers* 053. May.

[8] Jiang, C., Jian, N., Liu, T-Y. and Su, C-W. (2016). Purchasing power parity and real exchange rate in Central Eastern European countries. *International Review of Economics and Finance*, 44, 349–358.

[9] Kapetanios, G., Shin, Y. and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112, 359–379.

[10] Kutan, A. M. and Zhou, S. (2015). PPP may hold better than you think: Smooth breaks and non-linear mean reversion in real effective exchange rates. *Economic Systems*, 39, 358–366.

[11] Yilanci, V. (2012). The validity of purchasing power parity in Central and Eastern European countries: A rolling nonlinear unit root test. *Economic Research*, 25, 973–986.

TECHNOLOGY COMPETENCY ASSESSMENT OF ENTERPRISES BY USING DIFFERENT TYPES OF CLUSTERING

Ufuk Bolukbas
Yildiz Technical University, Vocational School
Gaziosmanpaşa, Istanbul, Turkey
E-mail: <bolukbas@yildiz.edu.tr>

Ali Fuat Guneri
Yildiz Technical University, Department of Industrial Engineering
Yıldız, 34349 Istanbul, Turkey
E-mail: <guneri@yildiz.edu.tr >

Abstract: In this study, technology management, technology competency and innovation management issues are examined by taking into account literature and manufacturing enterprises. The technology literature is investigated and reviewed for obtaining an evaluation framework. The proposed framework is structured based on the main criteria and decision variables based on literature review. The aim of the study is evaluating the technology management performances of SMEs to make the comparisons between firms from different clusters. K-means cluster analysis is applied on the survey data to analyse and evaluate the performances of manufacturing enterprises in Istanbul, Turkey. As a result, technology competency levels of the enterprises are determined in three, four and five clusters types.

Keywords: technology management, technology competency, statistical analysis, cluster analysis, small and medium-sized enterprises, SMEs.

1 INTRODUCTION

Technology management (TM), technological skills and technological competencies are critical factors that play a considerable role in firm's ability to achieve competitive advantages associated with technology. Technology is very important not only for competitive advantages of firms and sectors but also for the competency of countries, thus, it plays a decisive role on development and underdevelopment level in terms of the effects it creates. The manufacturing industry is one of the main drivers of the Turkish economy. A number of manufacturing sub-sectors in Turkey have been growing in recent years [13, 14]. Small to medium sized enterprises (SMEs) are considered to be the backbone of any economy as they play a major role in the economic development of a country [2].

SMEs and barriers to eco-innovation in the European Union is researched by exploring different firm profiles in the six clusters [17]. Brunswicker and Vanhaverbeke [5] use cluster analysis to sort firms applying similar innovation sourcing strategies into homogenous groups based on the open innovation in Small and Medium-Sized Enterprises (SMEs) by external knowledge sourcing strategies and internal organizational facilitators. A study of enterprise evaluation clusters SMEs based on their degree of openness for the case of manufacturing industries [21]. A research study explores the international strategy and performance by clustering the strategic types of SMEs [10].

Sila and Dobni [24] conduct an online survey of North American SMEs and obtained 229 responses. The study utilizes several statistical methods, including cluster analysis and profile analysis, to test five hypotheses for identifying the Business to Business e-commerce (B2BEC) usage patterns of SMEs in their supply chains. A study suggests the questionnaire administered to various business owners within South Africa and a total of 105 usable responses are received to determine whether SMEs develop the product strategies [6].

Tseng [25] emphasizes that marketing information, infrastructure capability, process capability, marketing capability, R&D capability and innovation decision capability are

302

measured qualitatively and quantitatively. The paper suggests that R&D capability is also related to innovation decision capability, which is required for knowledge management innovation and for reducing uncertainty and risk activities. Tseng et al. [26] propose a hybrid method to improve selection decision making in service innovation based on infrastructure capability, knowledge capability, process capability, market capability, R&D capability, innovation capability and technology capability criteria.

Ten experts evaluated the model to define the technology management (TM) levels of enterprises based on the six dimensions; *Process management, Product competitiveness, Information and communication technologies (ICT), Marketing strategies, Innovation and entrepreneurship activity and Research and development (R&D) activity*. Technology and management issues and related survey questions are used as input values for determining capabilities of enterprises.

2 TECHNOLOGY EVALUATION FRAMEWORK

The proposed list of dimensions (decision criteria) and variables used in technology competency analysis are based on TM literature and expert evaluation shown in Table 1. Six dimensions for the technology activities of enterprises are presented to evaluate the technology competency levels;

- Process Management: It considers the economic and ecological efficiency, the presence of technology management process, quality assurance, working culture, productivity or its contribution to corporate or business strategies, objectives, goals and interests.
- Product Competitiveness: It considers innovative and technological products development capability, advertising and promotional activities and the product potential to compete with its competitors in terms of technical performance or in any other dimension seen as important by customers.
- Information and Communication Technologies (ICT): It considers the computerized technologies, technology investments, barrier to ICT, enterprise software applications, expertise, ICT budgets and database usage, hardware and software infrastructure and collaborations for infrastructure projects.
- Marketing Strategies: It considers project cooperation, gaining a competitive advantage in the market and the probability of the commercialization model and of the product benefits to reach market requirements as commercial risks, marketing and positioning strategies.
- Innovation and Entrepreneurship: It considers technological innovations and news, management activities, technological developments, qualified personnel and expert staff, knowledge acquisition, and new product development.
- Research and Development: It considers research collaboration, financial capabilities, the average annual budget allocated to R&D activities, research and development projects, intellectual and industrial property rights and research activities for new products.

This framework is based on many articles using various technology evaluation infrastructures and frameworks so that multi-criteria decision making (MCDM) and statistical approaches are investigated to focus on technology competency issues. There are many methods, models and evaluation structures in the TM literature, but our framework is available and supportive for qualitative and quantitative variables together by transforming the data into a basic and useful database. Bolukbas and Guneri [4] research on the technology competency and proposed a framework which is based on the important studies from literature and presented in Table 1.

Table 1: Dimensions of technology competency analysis.

References/Dimensions	Year	1. Process Management	2. Product Competitiveness	3. ICT	4. Marketing Strategies	5. Innovation and Entrepreneurship	6. R&D
Cruz-González et al.[8]	2015	X			X	X	
Oliveira et al. [20]	2015		X				X
Ayerbe et al. [3]	2014		X		X	X	X
Krishnaswamy et al.[15]	2014	X	X		X	X	X
Dereli and Altun [9]	2013						X
Martín-Rojas et al. [18]	2013	X	X	X	X	X	X
Kukko [16]	2013		X	X	X	X	
Hameed et al. [11]	2012	X		X		X	X
Horwitch and Stohr [12]	2012	X	X	X	X	X	X
Mearns [19]	2012	X		X	X	X	X
Rampersad et al. [22]	2012		X	X	X	X	X
Shih [23]	2012		X	X		X	X
Xu et al. [27]	2012		X	X	X	X	X
Yang [28]	2012		X	X	X	X	X
Amadi-Echendu et al. [1]	2011	X	X		X		X

2.1. Cluster analysis

The k-means cluster analysis is a repetitive method so that this approach requires an initial value. Analyses are traditionally carried out based on the average of the initial values of the variables for most repetitive methods. Some of the cluster analysis run the algorithm by assigning initial values to zero. However, it is emphasized in the clustering literature that the assignment of initial values will not cause any difference. Theoretical explanation of k-means average method is summarized for clarifying the steps in cluster analyses which are applied to the enterprises;

Hierarchical clustering requires a distance or similarity matrix between all pairs of observations. It is necessary to calculate all possible distances as working with large data (n> 250). This situation is quite troublesome because it requires a series of operations so that instead of using the hierarchical clustering, the k-means technique as a clustering method which does not need to calculate all the distances would be appropriate. This method presents various advantages and is different from the hierarchical clustering method. In the k-means technique, the number of clusters must be known previously or have been already determined by expert or researcher. K-means method divide observations into the k number of clusters to minimize the sum of squares in-groups. When considered as a point in multidimensional x space, $term, x_1, x_2, \dots, x_n$, is the observation vector within n number of variables. For each group observation, if we define the cluster centres $as, a_{1n}, o_{2n}, \dots, a_{kn}$, the method assigns observations to the nearest cluster by the equation below;

$$W_n = n^{-1} \sum_{i=1}^{n} min|x_i - a_{kn}|^2 \tag{1}$$

W_n= variable's (n.) average difference value between the cluster centre (a_{kn}) and the observation (x_i), x_n: the number of observation and a_{kn}: the center value of the n. observation in the cluster k.

The first step of this method is to specify cluster centres then each observation is assigned to the nearest cluster depending on its distance to cluster centres. The process continues until

no observations switch clusters. Final cluster centres explain the average values of the variables for the different clusters. The final cluster centre shows that which variables are important for which clusters After all observations have been assigned to the clusters, the cluster centres are recalculated. Assignments for observations are established again by taking into account the new cluster centres. This algorithm is repeated until no significant change is observed in cluster centres [5]. In the method, confidence intervals, 95% are taken as critical values basis (0.05) to analyse the variables' explorative situation by the analysis of variance (Anova).

The developed technology competency model is based on the studies in which frameworks and infrastructures are examined together. Questions of survey research are prepared in compatible with the dimensions and criteria used in the model. The survey responses of the companies from different manufacturing sectors are converted into numerical scale (1-5 scale) to obtain a technology competency database. Some of the questions are consisted of the sub-divisions and they are represented as a variable by using the sub-questions shown in examples;

Table 2: Decision variables and structures for clustering

No	Questions type	Variables	Usage	Status
1	Plural	B1	Beneficial	Max.
2	Single	B2	B/N	Max.
3	Single	B3	B/N	Max
4	Plural	B4	Beneficial	Max
5	Single	C1	B/N	Max
6	Single	C2	Beneficial	Max.
7	Single	C4	B/N	Max.
8	Plural	C5	Beneficial	Max.
9	Plural	D1	Beneficial	Max.
10	Plural	D2	Beneficial	Max.
11	Single	D3	B/N	Max.
12	Plural	D5	Beneficial	Max.
13	Single	D6	Beneficial	Max.
14	Plural	D7	Beneficial	Max.
15	Plural	D8	Non-beneficial	Max.
16	Plural	E2	Beneficial	Max.
17	Single	E4	B/N	Max.
18	Single	F1	Beneficial	Max.
19	Single	F3	Beneficial	Max.
20	Single	F4	Beneficial	Max.
21	Plural	F5	Beneficial	Max.
22	Single	G1	Beneficial	Max.
23	Plural	G4	Beneficial	Max.
24	Single	G5	B/N	Max.
25	Single	G7	B/N	Max.
26	Single	G9	Beneficial	Max.

B/N term is demonstrated as Beneficial/Non-beneficial types for some of the questions. The labels A, B, C, D, E, F and G symbolize different parts of the questionnaire. The label A is the introduction part in which basic statistical values of SMEs are collected in the survey so there is no statistical calculations for part A in the clustering. Carrasco et al. [7] use linguistic terms defined for a five-point Likert scale to generate information about hotel e-service quality for

aggregating heterogeneous questionnaires so that hotel experts and web users utilize the scale of the questions which is provided to select one of five responses; strongly agree, agree, neutral, disagree or strongly disagree. In our study, experts use a five point scale for survey questions as; 1 point for very low, 2 point for low, 3 point for normal, 4 point for good and 5 point for very good level of knowledge of information.

3 RESULTS

This paper demonstrates the analyses of an exploratory study carried out to learn about the use and impact of the technology infrastructures and management capabilities for manufacturing SMEs. The field research is basically completed on 23th May 2015. For the survey research, two hundred of the surveys were performed with managers of the enterprises by face to face and three hundred of the surveys were performed by telephone. The 450 of 500 surveys are valid for this research because of the missing value of some question. Firms are analysed with respect to technology evaluation surveys. As a result of examining the technology road maps, models and frameworks which are used in TM literature, a conceptual model is proposed to provide an interface for technology competency analysis. The case study examines the SMEs by the critical questions, therefore there is no limitations to claim the generalization of the findings for the study. The cluster analysis is studied on the two different classifying system which include metric and metric/non-metric variables. Final cluster centres are used to classify and define the best- and worst-performing firms that represent the different technology competency levels. The survey data obtained from responses of SMEs is used to represent the differences of clustering groups. Different cluster analysis ways as (3, 4 and 5 sets) represent approximately the same results for the firms' technology competency, depending on the performance clusters. The best and worst-performing firms are determined as the same on an average, %90 based on three different clustering types as 3-, 4-and 5-clusters. The results of MCDM approaches and cluster analysis can be used to make comparisons on the technology competency performances of the SMEs.

References
[1] Amadi-Echendu, J., Lephauphau, O., Maswanganyi, M. and Mkhize, M. 2011. Case studies of technology roadmapping in mining', Journal of Engineering and Technology Management, 28(1): 23-32.
[2] Ayanda, A.M., and Laraba, A.S. 2011. Small and Medium Scale Enterprises as A Survival Strategy for Employment Generation in Nigeria. Journal of Sustainable Development, 4(1): 200-206.
[3] Ayerbe, C., Lazaric, N., Callois, M. and Mitkova, L. 2014. The new challenges of organizing intellectual property in complex industries: a discussion based on the case of Thales. Technovation, 34(4): 232-241.
[4] Bolukbas, U. and Guneri, A. F. 2016. Technology Competency Evaluation of SMEs in the Machine Sub-sector by Multi Criteria Decision Making Approaches', In Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference, pp. 891-897.
[5] Brunswicker, S. and Vanhaverbeke, W. 2015. Open Innovation in Small and Medium-Sized Enterprises (SMEs): External Knowledge Sourcing Strategies and Internal Organizational Facilitators', Journal of Small Business Management, 53(4): 1241-1263.
[6] Cant, M.C., Wiid, J.A. and Kallier, S.M. 2015. Product Strategy: Factors That Influence Product Strategy Decisions Of SMEs in South Africa. Journal of Applied Business Research (JABR), 31(2): 621-630.

[7] Carrasco, R. A., Sánchez-Fernández, J., Muñoz-Leiva, F., Blasco, M. F. and Herrera-Viedma, E. 2017, Evaluation of the hotels e-services quality under the user's experience. Soft Computing, 21(4): 995-1011.

[8] Cruz-González, J., López-Sáez, P., Navas-López, J. E. and Delgado-Verde, M. 2015. Open search strategies and firm performance: The different moderating role of technological environmental dynamism. Technovation, 35: 32-45.

[9] Dereli, T. and Altun, K. 2013. Technology evaluation through the use of interval type-2 fuzzy sets and systems. Computers & Industrial Engineering, 65(4): 624-633.

[10] Hagen, B., Zucchella, A., Cerchiello, P. and De Giovanni, N. 2012. International strategy and performance-Clustering strategic types of SMEs. International Business Review, 21(3): 369-382.

[11] Hameed, M.A., Counsell, S. and Swift, S. 2012. A conceptual model for the process of IT innovation adoption in organizations. Journal of Engineering and Technology Management, 29(3): 358-390.

[12] Horwitch, M. and Stohr, E. A. 2012. Transforming technology management education: Value creation-learning in the early twenty-first century. Journal of Engineering and Technology Management, 29(4): 489-507.

[13] Istanbul Chamber of Industry (ICI), Committee of Quality and Technology. 2007. Management Approach for the SMEs, (Improved third edition), 2007-06, ISBN: 975-512-870-0, Istanbul.

[14] Istanbul Chamber of Industry (ICI), Committee of Quality and Technology. 2007. Product Development of the SMEs, (Improved third edition), 2007-08, ISBN: 975-512-866-2, Istanbul.

[15] Krishnaswamy, K. N., Mathirajan, M. and Subrahmanya, M. B. 2014. Technological innovations and its influence on the growth of auto component SMEs of Bangalore: A case study approach. Technology in Society, 38: 18-31.

[16] Kukko, M. 2013. Knowledge sharing barriers in organic growth: A case study from a software company. The Journal of High Technology Management Research, 24(1): 18-29.

[17] Marin, G., Marzucchi, A. and Zoboli, R. 2015. SMEs and barriers to Eco-innovation in the EU: exploring different firm profiles. Journal of Evolutionary Economics, 25(3): 671-705.

[18] Martín-Rojas, R., García-Morales, V. J. and Bolívar-Ramos, M. T. 2013. Influence of technological support, skills and competencies, and learning on corporate entrepreneurship in European technology firms', Technovation, 33(12): 417-430.

[19] Mearns, M. 2012. Knowing what knowledge to share: Collaboration for community, research and wildlife. Expert systems with Applications, 39(10): 9892-9898.

[20] Oliveira, M.G., Rozenfeld, H., Phaal, R. and Probert, D. 2015. Decision making at the front end of innovation: the hidden influence of knowledge and decision criteria. R & D Management, 45(2): 161-180.

[21] Othman, I.M., Amaraa, N. and Landrya, R. 2012. SMEs' degree of openness: the case of manufacturing industries. Journal of technology management and innovation, 7(1): 186-210.

[22] Rampersad, G., Plewa, C. and Troshani, I. 2012. Investigating the use of information technology in managing innovation: A case study from a university technology transfer office. Journal of Engineering and Technology Management, 29(1): 3-21.

[23] Shih, H. Y. 2012. The dynamics of local and interactive effects on innovation adoption: The case of electronic commerce. Journal of Engineering and Technology Management, 29(3): 434-452.

[24] Sila, I. and Dobni, D. 2012. Patterns of B2B e-commerce usage in SMEs. Industrial Management & Data Systems, 112(8): 1255-1271.

[25] Tseng, M. L. 2011. Using a hybrid MCDM model to evaluate firm environmental knowledge management in uncertainty. Applied Soft Computing, 11(1): 1340-1352.

[26] Tseng, M. L., Lin, Y. H., Lim, M. K., and Teehankee, B. L. 2015. Using a hybrid method to evaluate service innovation in the hotel industry. Applied Soft Computing, 28: 411-421.

[27] Xu, K., Huang, K.F. and Gao, S. 2012. Technology sourcing, appropriability regimes, and new product development. Journal of Engineering and Technology Management, 29(2): 265-280.

[28] Yang, J. 2012. Innovation capability and corporate growth: An empirical investigation in China. Journal of Engineering and Technology Management, 29(1), 34-46.

A CONCEPT OF $\mathcal{SM}$-MEASURE TO COMPARE HIERARCHICAL CLUSTERINGS

Samo Drobne and **Mitja Lakner**
University of Ljubljana, Faculty of Civil and Geodetic Engineering,
Jamova cesta 2, SI-1000 Ljubljana, Slovenia
{Samo.Drobne,Mitja.Lakner}@fgg.uni-lj.si

Abstract: In this paper, we suggested a new concept for clustering comparison measure to compare hierarchically aggregated functional regions. The here suggested measure is a metric measure and considers a single basic spatial unit as a functional region. However, its undesirable property – dependency on the number of functional regions – should be solved in the future.

Keywords: hierarchical functional regions, comparison, clustering measure, metric measure, $\mathcal{SM}$-measure, average maximum proportion.

1 INTRODUCTION

Clustering comparison measures are used to compare clusterings. Wagner and Wagner [39] identified three sections of clustering comparison measures: counting of pairs of elements, summation of set overlaps, and the use of the information-theoretical mutual information. There are several measures based on counting of pairs of elements that are defined in the same way in both clusterings; e.g. Chi Squared Coefficient suggested in 1900 by Pearson, General Rand Index [32], Adjusted Rand Index [11], Fowlkes–Mallows Index [8], Adjusted Fowlkes–Mallows Index [11, 40], Mirkin Metric [27], Jaccard Index [12], Partition Difference [21]. The measures based on summation of set overlaps try to match clusters that have a maximum absolute or relative overlap [39]; examples of such measures are: $\mathcal{F}$-Measure [20], Meila-Heckerman Measure [25], van Dongen Measure [36]. In the section of measures based on mutual information, there are normalized mutual information measures introduced by Strehl and Ghosh [33] and Fred and Jain [9], Variation of Information suggested by Meila [24], Adjusted Mutual Information proposed by Vinh [37] and Vinh et al. [38], and others.

In spatial sciences, the concept of functional regions (FRs) is one of the key concepts for analysing, modelling, monitoring, and predicting socio-economic structures. FRs can be described as reasonably functioning spatial entities composed of economically and socially connected areas, i.e. basic data/spatial units (BSUs) like census units, statistical units, statistical local areas, settlements, communities, municipalities, postal zones, etc. In the group of connected areas, many social and economic interactions, interdependence of commuting flows, flows of goods and services, communication flows, traffic flows, financial flows, etc., occur. Brown and Holmes [1] describe FRs as a combination of functionally complementing BSUs, which have more economic interactions with each other than with outside units. And, Johansson [14] and Karlsson and Olsson [15] define a FR as an area characterised by a high frequency of intra-regional economic interaction, such as labour commuting and intra-regional trade in goods and services, and an area of agglomeration of activities and transport infrastructure facilitating significant mobility of people, products, and information.

FR's organisation is based on horizontal relations in a space in a form of spatial flows or interactions between BSUs [35]. Functional regionalisation is the procedure of combining BSUs into FRs with the goal of generalising the functional flows and spatial interactions addressed. FRs are thus understood as generalised patterns of flows and interactions in a space [6].

The Intramax method [22] is a popular method for modelling functional regions; some recent examples are in [1–5, 7, 10, 13, 16–19, 26, 28–31, 41]. The hierarchical aggregation procedure, called Intramax procedure, seeks to maximise the proportion of the total interaction which takes place within the aggregations of basic data units, and thereby to minimise the proportion of cross-boundary movements in the system as a whole [22].

Comparison of (hierarchical) FRs is interesting when comparing (systems of) FRs of the same territory modelled by different methods, in different time horizons, by using different flows and/or different BSUs, etc. In many older applications, systems of FRs are compared solely visually; e.g. [23]. Nevertheless, there are a few papers that use one of the above-mentioned clustering comparison measures to compare different FR systems; e.g. Watts [42] used the AMI Index, i.e. the Adjusted Mutual Information Index [38]. However, when comparing the whole system of hierarchical FRs, an appropriate clustering comparison measure is needed. It means that the clustering comparison measure should consider a single BSU as a single FR.

In this paper, we analyse the use of selected clustering comparison measures to compare systems of hierarchical FRs modelled by the Intramax procedure. Further, we suggest a new concept of comparing FRs, which is easy to interpret, it is normalized, it considers each BSU as a single FR, but it should still be improved for the number of FRs.

2 METHODOLOGY

To analyse different clustering comparison measures, we compared systems of hierarchical FRs modelled by the Intramax procedure using the inter-municipal labour commuting flows for 2011 – but, for two different sets of municipalities. Namely, there were 192 municipalities in Slovenia in 2000, but their number increased to 210 in 2011. For the purpose of the analysis, which was done for the dimension of 210 municipalities in 2011, the database for 192 municipalities in 2000 was adequately expanded to the dimension of 210 municipalities. To model FRs as well as to calculate clustering comparison measures, we developed a programme code in *Mathematica 11.0*.

We analysed the measures based on counting of pairs of elements as described in [39], one measure based on summation of set overlaps, i.e. van Dongen Measure [36], and the most recently suggested measure based on mutual information, i.e. Adjusted Mutual Information Index [38]. Due to the length limitation of this paper, definitions of the tested measure are not explicitly provided, but the interested reader can find them in the original papers: General Rand Index in [32], Adjusted Rand Index in [11], Fowlkes–Mallows Index in [8], Adjusted Fowlkes–Mallows Index in [8, 40], Mirkin Metric in [27], Jaccard Index in [12], Partition Difference in [21], van Dongen Measure in [36], and Adjusted Mutual Information Index in [38].

Along the clustering comparison measures already introduced in the literature, we suggest a new concept of comparing FRs based on an average maximum proportion of matched municipalities in FRs. Let X denote the finite set of n BSUs (in our case municipalities) $\{X_1, ..., X_n\}$, where $|X| = n$, while FR denotes the system of N FRs $\{FR_1, ..., FR_N\}$, of power $|FR| = N$, which consists of disjoint subsets X, so that their union is X. For a system of functional regions FR, let us assume that in each FR_i there is at least one municipality. $FR'_j = \{FR'_1, ..., FR'_N\}$ denotes the second system of FRs consisting of the same n municipalities $\{X_1, ..., X_n\}$. $M = [m_{ij}]$ denotes the cross-matrix with dimensions $N \times N$ of pairs FR_i, FR'_j:

$$m_{ij} = \frac{\left|FR_i \cap FR'_j\right|}{max\{|FR_i|, |FR'_j|\}} , \quad i, j = 1, ..., N.$$

$\mathcal{SM}$-Measure, which is the average value of N maximum values of matrix elements M, expresses the average maximum proportion of matched municipalities in FRs. If both systems of FRs, FR_i and FR'_j, are equal, then $\mathcal{SM} = 1$.[1] In our experiment, FR_i were FRs modelled by inter-municipal labour commuting flows for 2011 using 210 municipalities in 2011, but FR'_j were FRs defined by 192 municipalities in 2000. In this way, we analysed the impact of new municipalities on modelling FRs.

3 RESULTS

Figures 1a–1h, 2, and 3 show the analysed clustering comparison measures for comparing the systems of hierarchical FRs modelled by the hierarchical aggregation procedure Intramax using the inter-municipal labour commuting flows for 2011 for two different sets of municipalities in 2000 and 2011.

Rand Index counts correctly classified pairs of elements [32]. However, Rand Index (see Fig. 1a) depends on both the number of clusters (FRs) and the number of elements (BSUs) [39]. Adjusted Rand Index is normalized for the number of clusters [11]. But, in some cases, it may result in negative values [25], it is hard to interpret [39], and it does not consider BSUs from the beginning of the aggregation procedure as FRs: Adjusted Rand Index at Fig. 1b starts from 0, but it should start from 1. Adjusted Fowlkes–Mallows Index [8, 40] (see Fig. 1d) and Partition Difference [21] (see Fig. 1g) consider BSUs from the beginning of the hierarchical aggregation procedure as FRs, but, as reported by Wagner and Wagner [39], the strong assumptions on the distribution of (Adjusted) Fowlkes–Mallows Index make it hard to interpret. Partition Difference is sensitive to cluster size and the number of clusters, and it is not normalized. The use of Mirkin Metric [27] (see Fig. 1e), Jaccard Index [12] (see Fig. 1f), and van Dongen Measure [36] (see Fig. 1h), show the same problems with sensitivity to cluster numbers, they are difficult to interpret, and/or not normalized. On the other hand, the AMI Index solves most of the aforementioned disadvantages, but, as reported by Romano et al. [34], it is still hard to interpret – and, it does not consider BSUs (i.e. municipalities in our experiment) as a single FR (see Fig. 1i). The last is solved by the $\mathcal{SM}$-Measure, but, it has an undesirable property, i.e. it is not independent of the number of FRs (see Fig. 1j).

4 CONCLUSIONS

In this paper, we suggested a new concept for comparing hierarchically aggregated functional regions. The $\mathcal{SM}$-Measure is a metric (easy to interpret) and normalized measure (using the nominal [0, 1] range). It considers a single basic spatial unit as a functional region – this is especially important for systems of functional regions at the beginning of the hierarchical aggregation procedure. On the other hand, the $\mathcal{SM}$-Measure is sensitive to the number of FRs in the system. Searching the solution to this problem will be the focus of our future research.

---

[1] The $\mathcal{SM}$-Measure was tested by simulating and comparing 500 randomly generated partitions, where the average value of $\mathcal{SM}$ was 0.5, while the minimum value of $\mathcal{SM}$ in this test was 0.2.
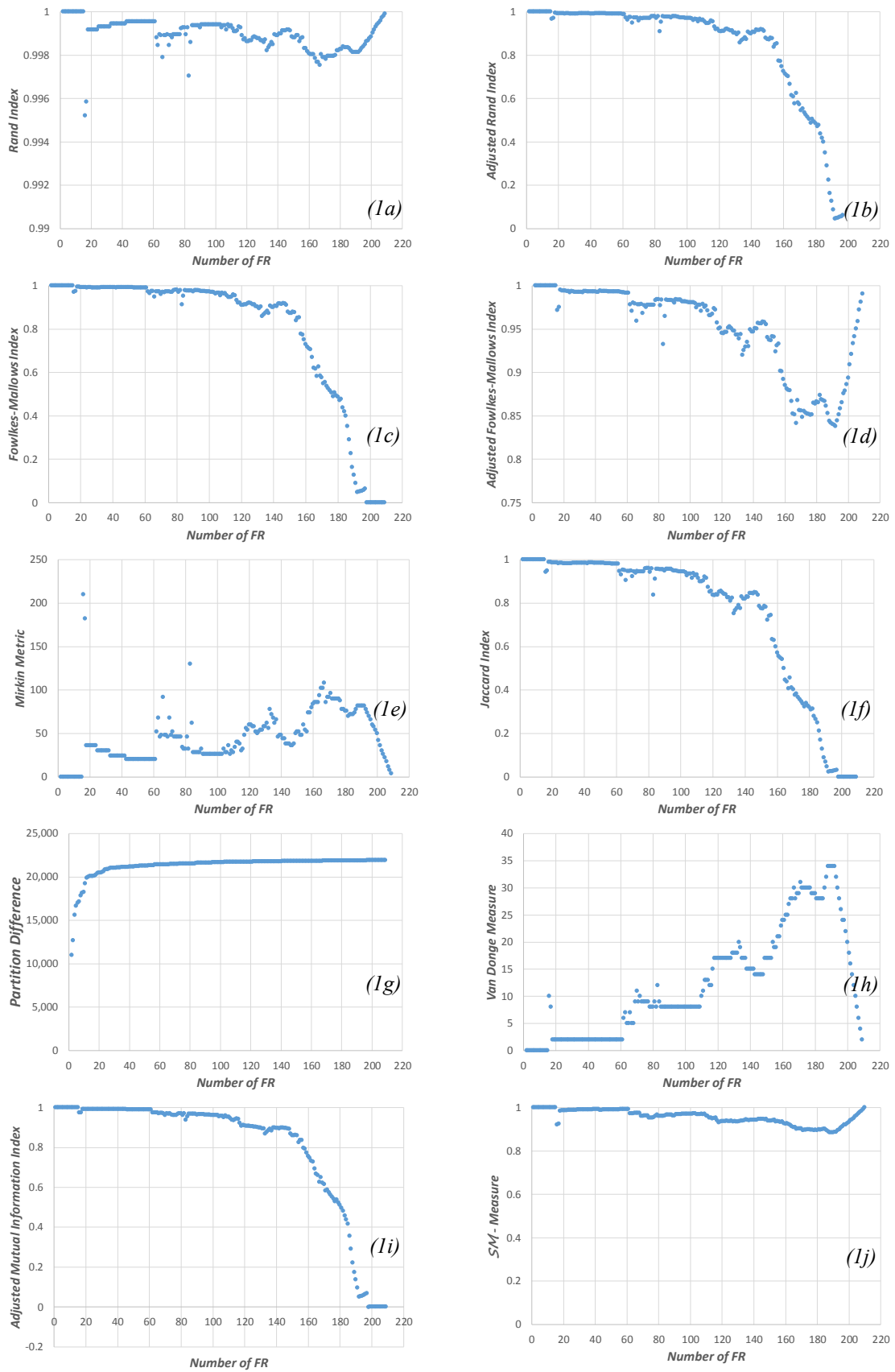
Figure 1: (1a) Rand Index [32], *(1b)* Adjusted Rand Index [11], *(1c)* Fowlkes–Mallows Index [8], *(1d)* Adjusted Fowlkes–Mallows Index [8, 40], *(1e)* Mirkin Metric [27], *(1f)* Jaccard Index [12], *(1g)* Partition Difference [21], *(1h)* van Dongen Measure [36], *(1i)* Adjusted Mutual Information Index [37, 38], *(1j)* $\mathcal{SM}$-Measure

References

[1] Brown, P.J.B. and Hincks, S. (2008). A Framework for Housing Market Area Delineation: Principles and Application, Urban Studies, 45(11): 2225–2247.

[2] Drobne, S. (2017). Functional regions and areas: literature review according to application fields. Geodetski vestnik, 61(1): 35–57.

[3] Drobne, S. and Bogataj, M. (2012). A method to define the number of functional regions: An application to NUTS 2 and NUTS 3 levels in Slovenia. Geodetski vestnik, 56(1): 105–150.

[4] Drobne, S. and Bogataj, M. (2014). Regions for Servicing Old People: Case study of Slovenia. Business systems research journal, 5(3): 19–36.

[5] Drobne, S. and Bogataj, M. (2015). Optimal allocation of public service centres in the central places of functional regions. In: 15th IFAC/IEEE/IFIP/IFORS Symposium on Information Control Problems in Manufacturing (pp. 2449–2454), INCOM 2015. Ottawa: IFAC.

[6] Drobne, S. and Lakner, M. (2016). Intramax and other objective functions. Moravian geographical reports, 24(2): 12–25.

[7] Feldman, O., Simmonds, D., Troll, N. and Tsang, F. (2006). Creation of a system of functional areas for England and Wales and for Scotland. In: European Transport Conference, 2005 Proceedings, Strasbourg, France, October 3–5, 2005, Association for European Transport. http://abstracts.aetransport.org/paper/index/id/2284/confid/11 (Accessed 14.11.2015.)

[8] Fowlkes, E.B. and Mallows, C.L. (1983). A Method for Comparing Two Hierarchical Clusterings. Journal of the American Statistical Association, 78(383):553–569.

[9] Fred, A.L.N. and Jain, A.K. (2003). Robust Data Clustering. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, (3):128–136.

[10] Goetgeluk, R., De Jong, T. (2007). What about the spatial dimension of subsidiarity in housing policy? ENHR 2007 International conference on Sustainable Urban Areas, The Netherlands, Rotterdam, June, 25–28 2007. https://www.yumpu.com/en/document/view/26672955/what-about-the-spatial-dimension-of-subsidiarity-in-housing-policy (Accessed 18.11.2015.).

[11] Hubert, L. and Arabie, P. (1985). Comparing partitions. Journal of Classification, 2:193–218.

[12] Jaccard, P. (1912). The distribution of the flora in the Alpine Zone. New Phytologist, 11(2): 37–50.

[13] Jaegal, Y. (2013). Delineating Housing Market Areas in the Seoul Metropolitan Area Using a Geo-Computational Approach. Journal of the Association of Korean Geographers, 2(1): 7–20.

[14] Johansson, B. (1998). Infrastructure, Market Potential and Endogenous Growth. Jönköping (Mimeo). Jönköping International Business School.

[15] Karlsson, C. and Olsson, M. (2006). The identification of functional regions: theory, methods, and applications. The Annals of Regional Science 40(1): 1–18.

[16] Kohl, T. and Brouver, A.E. (2014). The Development of Trade Blocs in an Era of Globalisation. Environment and Planning A, 46(7): 1535–1553.

[17] Koo, H. (2012). Improved Hierarchical Aggregation Methods for Functional Regionalization in the Seoul Metropolitan Area. Journal of the Korean Cartographic Association, 12(2): 25–35.

[18] Krygsman, S., De Jong, T. and Nel, J. (2009). Functional transport regions in South Africa: An examination of national commuter data. In: Proceedings of the 28th South African transport conference (SATC 2009), Pretoria, South Africa, June 6–9, 2009. Pretoria, Academic Press: 144–154.
http://repository.up.ac.za/bitstream/handle/2263/11952/Krygsman_Functional%282009%29.pdf (Accessed 18.11.2015.).

[19] Landré, M. and Håkansson, J. (2013). Rule versus Interaction Function: Evaluating Regional Aggregations of Commuting Flows in Sweden. European Journal of Transport and Infrastructure Research, 13(1): 1–19.

[20] Larsen, B. and Aone, C. (1999). Fast and Effective Text Mining Using Linear Time Document Clustering. Proceedings of the KDD, 16–29.

[21] Li, T., Ogihara, M. and M. Sheng (2004). On Combining Multiple Clusterings. Proceedings of the ACM Conference on Information and Knowledge Management, (13):294–303.

[22] Masser, I. and Brown, P.J.B. (1975). Hierarchical Aggregation Procedures for Interaction Data. Environment and Planning A, 7(5): 509–523.

[23] Masser, I. and Scheurwater, J. (1980). Functional Regionalisation of Spatial Interaction Data: An Evaluation of Some Suggested Strategies. Environment and Planning A, 12(12): 1357–1382.

[24] Meila, M. (2007). Comparing clusterings—an information based distance. Journal of Multivariate Analysis 98(5): 873 – 895.

[25] Meila, M., Heckerman, D. (1999). An Experimental Comparison of Model-based Clustering Methods. Proceedings of the Conference on Knowledge Discovery and Data Mining, 16–22.

[26] Meredith, D., Charlton, M., Foley, R. and Walsh, J. (2007). Identifying travel-to-work areas in Ireland: a hierarchical approach using GIS. Geographical Information Science Research Conference, NCG, NUI Maynooth: 11–13. http://www.geocomputation.org/2007/2B-Apps_Urban_Modelling_1/2B3.pdf (Accessed 15.8.2015.)

[27] Mirkin, B.G. (1996). Mathematical classification and clustering. Kluwer Academic Press.

[28] Mitchell, W. and Stimson, R. (2010). Creating a New Geography of Functional Economic Regions to Analyse Aspects of Labour Market Performance in Australia. In Dalziel, P. (Ed.). Innovation and Regions: Theory, Practice and Policy (pp. 178–220). Lincoln, New Zealand: AERU Research Unit.

[29] Mitchell, W. and Watts, M. (2010). Identifying Functional Regions in Australia Using Hierarchical Aggregation Techniques. Geographical Research, 48(1): 24–41.

[30] Mitchell, W., Baum, S., Flanagan, M. and Hannan, M. (2013). CofFEE Functional Economic Regions. AURIN project. Centre of Full Employment and Equity. Darwin, Australia. http://e1.newcastle.edu.au/coffee/functional_regions/ (Accessed 18.11.2015.).

[31] Nel, J.H., Krygsmany, S.C. and de Jong, T. (2008). The Identification of Possible Future Provincial Boundaries for South Africa Based on an Intramax Analysis of Journey-to-Work Data. Orion, 24(2): 131–156.

[32] Rand, W.M. (1971). Objective Criteria for the Evaluation of Clustering Methods. Journal of the American Statistical Association, 66(336):846–850.

[33] Romano, S., Vinh, N.X., Bailey, J. and K. Verspoor (2016). Adjusting for Chance Clustering Comparison Measures. Journal of Machine Learning Research 17: 1–32.

[34] Strehl, A. and Ghosh, J. (2002). Cluster Ensembles - A Knowledge Reuse Framework for Combining Multiple Partitions. Journal of Machine Learning Research, 3:583–617.

[35] Ullman, E. L. (1980). Geography as spatial interaction. Seattle and London, University of Washington Press.

[36] van Dongen, S. (2000). Performance Criteria for Graph Clustering and Markov Cluster Experiments. Technical Report INS–R0012, Centrum voor Wiskunde en Informatica.

[37] Vinh, N.X. (2010). Information theoretic methods for clustering with applications to microarray data. PhD thesis. Sydney. The University of New South Wales.

[38] Vinh, N.X., Epps, J. and Bailey, J. (2010). Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. The Journal of Machine Learning Research 11(oct): 2837–2854.

[39] Wagner, S. and Wagner, D. (2007). Comparing Clusterings - An Overview. Technical Report 2006-04, Faculty of Informatics, University of Karlsruhe.

[40] Wallace, D.L. (1983). A Method for Comparing Two Hierarchical Clusterings: Comment. Journal of the American Statistical Association, 78(383):569–576.

[41] Watts, M. (2009). Rules versus hierarchy: An application of fuzzy set theory to the assessment of spatial grouping techniques. In: Kolehmainen, M. et al. (ed.). Adaptive and naturals computing algorithms. Berlin, Heidelberg: Springer-Verlag: 517–526.

[42] Watts, M. (2013). Assessing Different Spatial Grouping Algorithms: An Application to the Design of Australia's New Statistical Geography. Spatial Economic Analysis, 8(1): 92–112.

FORECASTING ACCURACY AND CHANGE POINT DETECTION

Gregory Gurevich

Industrial Engineering and Management Department, SCE - Shamoon College of Engineering,
Bialik Sts. 56, P.O. Box 950, Beer Sheva 84100, Israel
E-mail: gregoryg@sce.ac.il

Yossi Hadad

Industrial Engineering and Management Department, SCE - Shamoon College of Engineering,
Bialik Sts. 56, P.O. Box 950, Beer Sheva 84100, Israel
E-mail: yossi@sce.ac.il

Baruch Keren

Industrial Engineering and Management Department, SCE - Shamoon College of Engineering,
Bialik Sts. 56, P.O. Box 950, Beer Sheva 84100, Israel
E-mail: baruchke@sce.ac.il

Abstract: The accuracy of time series forecasting often decreases because of the existence of change points in the data. This paper presents a novel method for time series forecasting that taking into account the possibility of a change point in past data. The proposed method can be applied to situations where the considered time series consists of independent or weakly dependent observations. Change point analysis prevents the omission of relevant data as well as the forecasting that may be based on irrelevant data. The study demonstrates that change point techniques may increase the accuracy of forecasts.

Keywords: change point, business forecasting, error indexes, homogeneous series.

1 INTRODUCTION

Business forecasting is the process of estimating future business conditions by analyzing past business data. A good forecasting can help to develop and to improve the business plans by increasing the knowledge of the marketplace. Many of the forecasting methods are based on a known historical data. However, the data depends on many factors that may be changed over time. Sometimes it is easy to identify the points where the data is changed. If the changes are obvious, decision-makers perceive the changes and modify their forecasts accordingly. However, there are situations where it is difficult to identify the existence of change points. In these cases, one can use statistical analysis that can accurately detect hidden change points.

There are also situations where decision-makers believe that a change point occurs due to external or internal events, but accurate statistical analysis demonstrates that there is no significant change point. Therefore, it is very important to use statistical methods for revealing change points in order to identify changes that enable us to adjust and improve the forecast by using only the relevant data.

2 FORECASTING METHODS

Forecasting methods can be classified into two groups: qualitative (subjective) techniques and objective methods. The subjective group includes forecasting methods that are based on human judgment. The objective or mathematical forecasting methods include forecasting methods based only on historical (past) data, causal models that allow prediction of the dependent variable through a number of independent variables that can be estimated and models that predict future phenomena according to the laws of nature.

There are many error indexes (loss functions), that help to evaluate the accuracy of the forecasting model (see, [8]). These indexes assist in selecting the model that is most appropriate under specific circumstances. Most of the indexes weight somehow the past forecast errors. Let us consider the following notations:

n - the known number of considered periods (i.e., $n > 1$ is a given sample size).

X_t - the observed value during period t, $t = 1, 2, ..., n$.

Y_t - the forecasted value for period t, $t = 1, 2, ..., n$, that is obtained by applying a forecast method.

e_t - the forecast error in period t. The forecast error is the difference between the forecasted value for that period and the observed value for that period, $e_t = Y_t - X_t$.

Three following error indexes will be used in this paper:

- Mean Absolute Error (MAE) is calculated as follows:

$$MAE = \frac{\sum_{t=1}^{n} |Y_t - X_t|}{n} = \frac{\sum_{t=1}^{n} |e_t|}{n} \ .$$

- Root Mean Squared Error (RMSE) is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (Y_t - X_t)^2}{n}} = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}} \ .$$

- Mean Absolute Percentage Error (MAPE) is calculated as follows:

$$MAPE = 100 \frac{\sum_{t=1}^{n} \left| \frac{Y_t - X_t}{X_t} \right|}{n} = 100 \frac{\sum_{t=1}^{n} \left| \frac{e_t}{X_t} \right|}{n} \ .$$

3 CHANGE POINT DETECTION AND ESTIMATION TECHNIQUES

The statistics and engineering literature has broadly discussed change point detection problems during last decades. Various engineering applications consider different forms of the classical AMOC (at most one change) change point problem, i.e., detection and estimation of a single change point in the distributions of a sequence of independent random variables (for more details, see [2], [6]). Problems with multiple change points can be defined similarly. Many of the natural applications give cause to consider, as well, a change point in regression models such as linear, logistic, or nonparametric regressions. This paper considers only the basic *iid* (independent and identically distributed) version of the AMOC change point problem.

Let $X_1, X_2, ..., X_n$ be a time series, where $X_1, X_2, ..., X_n$ are independent random variables. In the formal context of hypotheses testing, we state the basic change point detection as the following problem (1):

H_0, the null: $X_1, X_2, ..., X_n \sim F_0$ (1)

versus

H_1, the alternative: $X_i \sim F_0$, $X_j \sim F_1$, $i = 1, ..., \nu - 1$, $j = \nu, ..., n$,

where F_0 and F_1 are distribution functions that correspond to density functions f_0 and f_1, respectively. The unknown parameter ν, $1 \le \nu \le n$, is called a change point. According to

the statistical literature, problem (1) has been investigated in parametric and nonparametric forms, depending on assumptions made on the distribution functions F_0 and F_1. In the parametric case of problem (1), the distribution functions F_0 and F_1 are assumed to have known forms that can contain certain unknown parameters. In the nonparametric case of (1), the functions F_0 and F_1 are assumed to be completely unknown. When problem (1) is stated nonparametrically, the common components of change point detection policies have been proposed to be based on signs and/or ranks and/or U-statistics (e.g. [9], [1], [2], [3]). In particular, Gurevich [3] analyzed problem (1) when F_0, F_1 are unknown and for all x, $F_1(x) \leq F_0(x)$ (i.e., after a possible change the observations are stochastically larger than before the change). He suggested rejecting H_0 for large values of the statistics

$$MK = \sum_{k=2}^{n} \left(U_{k-1,n-k+1} - (k-1)(n-k+1)/2 \right),\tag{2}$$

$$MD = \sum_{k=2}^{n} \frac{U_{k-1,n-k+1} - (k-1)(n-k+1)/2}{\sqrt{(k-1)(n-k+1)(n+1)/12}},\tag{3}$$

where $U_{k-1,n-k+1} = \sum_{i=1}^{k-1}\sum_{j=k}^{n} I\left(X_i \leq X_j \right)$ is the Mann-Whitney statistic for two samples of size k-1 and n-k+1 ($I(\cdot)$ is the indicator function). In addition, the author presented the asymptotic ($n \to \infty$) distribution of the statistics MK and MD under H_0:

$$\lim_{n \to \infty} P_{H_0}\left(\frac{MK}{S1_n} > x \right) = 1 - \Phi(x),\tag{4}$$

$$\lim_{n \to \infty} P_{H_0}\left(\frac{MD}{S2_n} > x \right) = 1 - \Phi(x),$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution,

$$-\infty < x < \infty, \quad S1_n = \sqrt{\frac{n+1}{12}\left(\sum_{k=2}^{n}(k-1)(n-k+1) + 2\sum_{k=2}^{n}\sum_{r=1}^{n-k+1}(k-1)(n-k+1-r) \right)},$$

$$S2_n = \sqrt{(n-1) + 2\sum_{k=2}^{n}\sum_{r=1}^{n-k+1}\sqrt{\frac{(k-1)(n-k+1-r)}{(n-k+1)(k-1+r)}}}.$$

While current change point literature relies mainly on testing hypotheses (1), rather scant work has been done on the problem of estimating change point v (see e.g. [5], [6]). When H_0 is rejected, the issue of estimating the unknown parameter v can be stated. Considering a change point estimation problem in a nonparametric framework, Gurevich and Raz [4] studied the behavior of change point estimator based on statistic (3):

$$\hat{v}_D = \arg\max_{2 \leq k \leq n}\left\{ \left| U_{k-1,n-k+1} - (k-1)(n-k+1)/2 \right| / \left[(k-1)(n-k+1)(n+1)/12 \right]^{\frac{1}{2}} \right\}.\tag{5}$$

The authors confirmed the efficiency of this estimator even for small and average sample sizes ($n \approx 40$).

The recommendation here is to utilize the change point nonparametric tests based on statistics (2), (3) and the change point estimator (5).

The forecasting method used in this paper can be applied to situations where the considered time series consists of independent or weakly dependent observations. Moreover, the data before a possible change point seems to be a homogeneous series, and after the

possible change point, it also seems to be a homogeneous series. The data has 30-70 observations, making it likely that only one change point, if any, is expected. A larger number of observations can increase the possibility of more than one change point. If a change point exists, the forecast should be based only on the data after this point. It is clear that removing part of the observations when a change point does not exist decreases the accuracy of the forecast.

4 APPLICATION

Fig. 1 below demonstrates 50 observations that look like a homogeneous series. The data is based on the case study of Hadad et al. [6] and presented in Table 1.

Table 1: The observations.

t	1	2	3	4	5	6	7	8	9
X_t	5.33	5.39	5.21	4.54	6.00	5.40	5.38	5.64	5.41
t	10	11	12	13	14	15	16	17	18
X_t	4.45	4.97	5.00	5.62	4.98	4.94	5.30	5.76	4.78
t	19	20	21	22	23	24	25	26	27
X_t	5.36	5.18	4.52	5.06	4.14	4.93	5.35	5.55	6.15
t	28	29	30	31	32	33	34	35	36
X_t	5.21	4.89	5.07	5.97	6.06	5.26	5.86	5.44	6.04
t	37	38	39	40	41	42	43	44	45
X_t	6.68	7.06	5.93	6.17	6.54	6.27	6.50	5.26	6.21
t	46	47	48	49	50				
X_t	5.88	6.28	5.45	5.49	6.12				

It is reasonable to assume that if a change point exists, the observations before this point are from the same distribution, and the observations after this point are from another distribution.
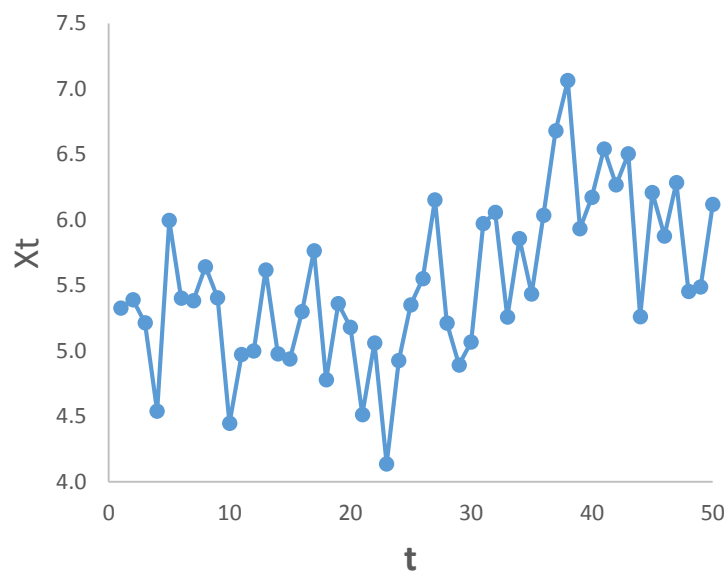
Figure 1: The series of observations.

Hadad et al. [6] proposed the following three-step procedure for forecasting a time series with a suspicion for a change point.

Step 1: Apply change point tests based on statistics (2), (3) to detect a change point.

Step 2: If a change point is not found – all the observations can be used for future forecasting. In such case, one can use a forecasting method that is appropriate for a stationary series. If a change point is found, go to step 3.

Step 3: Only the observations after the change point should be used for future forecasting. In such case, one can use a forecasting method that is appropriate for a stationary series based only on the observations that obtained after the change point.

Applying the above procedure to the data presented in Figure 1 yields the following results: $MK = 5686.50$, $MD = 124.47$, $S1_{n=50} = 1487.5$, $S2_{n=50} = 34.07$. Thus, by equation (4), the p-values of the tests based on statistics MK and MD are approximately equal to zero (i.e., less than 0.05): $p_{value} \approx 1 - \Phi(MK / S1_{50}) = 1 - \Phi(3.82) \approx 0$ and $p_{value} \approx 1 - \Phi(MD / S2_{50}) = 1 - \Phi(3.65) \approx 0$. The conclusion is that there is a change in the distribution of the observations and the change point estimator given by equation (5) is $\hat{v}_D = 31$. Therefore, the forecast should be calculated based only on the last 20 observations.

In order to demonstrate the difference between the forecasts, this is based on all the data $(t = 1,2,...,50)$ and the other is based only on the data after the change point $(t = 31,32,...,50)$, five common forecasting methods (see, [7]) were used:

(A) Simple average; (B) Linear trend; (C) Simple exponential smoothing; (D) Brown's linear exponential smoothing; (E) Holt's linear exponential smoothing.

Note that these methods used all known observations as input for their forecasts. Other methods that used only the latest observations (for example, moving average and double moving average), may generate an identical forecast to those obtained by change point analysis because the observations before the change point may be omitted by both forecasting methods. However, such methods not always lead to more accurate forecasts. In order to compare between the case where the forecast is based on all the data and the case where the forecast is based only on the data after the change point, three common error indexes (RMSE, MAE, MAPE) were calculated for both cases. The average error indexes were calculated only on the last 17 couples (forecast and observation) for both cases, as presented in Table 2.

Table 2: Values of the error indexes

Model	All the data			Proposed (partial data)		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
A	0.8135	0.6745	10.6822	0.5159	0.4064	6.7888
B	1.1687	1.0412	16.6786	0.6548	0.5424	8.8969
C	0.5280	0.4411	7.3800	0.5214	0.4170	7.0484
D	0.3427	0.2812	4.7946	0.3345	0.2684	4.5355
E	0.5791	0.4419	7.4557	0.5741	0.4364	7.5775
Min	0.3427	0.2812	4.7946	0.3345	0.2684	4.5355

Each of the statistics is based on one-ahead forecast errors, which are the differences between the data value at time (X_t) and the forecast made of that value (Y_t). Thus, the three statistics measure the magnitude of the errors. A better model will give smaller error values. The

values of the errors were computed only for the periods that both forecasts existed (in general, for periods 34-50).

Table 2 shows that for both forecasts model D (Brown's linear exponential smoothing model) is more appropriate. In spite of that, the proposed forecasts are based on fewer observations, with all five forecasting methods providing smaller error indexes. This fact confirms the conclusion that the data before the change point should not be used as input for the forecasting.

5 CONCLUSIONS

This paper presents a novel method for time series forecasting with the possibility of a change-point in the data. The proposed method uses change-point techniques to detect change points and to improve the forecasting process by taking into account the potential existence of a change point. The results of the example in this paper show that forecast accuracy is improved (a lower error index) by taking the change point into account. Therefore, the main conclusion is that one of the earliest stages in the forecasting process should be applying an appropriate statistical test for detecting change points.

References

[1] Gombay, E. 2001. U-statistics for Change under Alternatives. Journal of Multivariate Analysis, Vol. 78(1): 139-158.

[2] Gurevich, G. 2006. Nonparametric AMOC Change point Tests for Stochastically Ordered Alternatives. Communications in Statistics - Theory and Methods, Vol. 35(5): 887-903.

[3] Gurevich, G. 2009. Asymptotic distribution of Mann-Whitney type statistics for nonparametric change point problems. Computer Modelling and New Technologies, Vol. 13(1): 18-26.

[4] Gurevich, G., Raz, B. 2010. Monte Carlo analysis of change point estimators. Journal of Applied Quantitative Methods, Vol. 5(4): 659-669.

[5] Gurevich, G., Hadad, Y., Ofir, A. and Ohayon, B. 2011. Statistical analysis of temperature changes in Israel: an application of change point detection and estimation techniques. Global Nest Journal, Vol. 13(3): 215-228.

[6] Hadad, Y., Keren, B., Gurevich, G. 2017. Improving demand forecasting using change point analysis. International Journal of Business Forecasting and Marketing Intelligence, Vol. 3(2): 130-151.

[7] Nahmias, S. 2001. Production and operations analysis, McGraw-Hill/Irwin, New York.

[8] Wang, Y., Wu, C., Yang, L. 2016. Forecasting crude oil market volatility: A Markov switching multifractal volatility approach. International Journal of Forecasting, Vol. 32(1): 1-9.

[9] Wolfe, D.A., Schechtman, E. 1984. Nonparametric statistical procedures for the changepoint problem. Journal of Statistical Planning and Inference, Vol. 9(3): 389–396.

WHAT DRIVES CROATIAN REGIONAL EXPORT?

Saša Jakšić
Faculty of Economics and Business, University of Zagreb, Croatia, Department of Statistics
Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia
E-mail: sjaksic@efzg.hr

Nataša Erjavec
Faculty of Economics and Business, University of Zagreb, Croatia, Department of Statistics
Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia
E-mail: nerjavec@efzg.hr

Abstract: Export performance is a key element in establishing regional competitiveness. In line with recent empirical research which stress the importance of domestic factors, this paper investigates potential factors that could influence Croatian regional export dynamics for the period from 2004 to 2014. The results of the estimated panel models point to the strategic importance of strengthening manufacturing, especially in the coastal regions. On the other hand, weak domestic demand as well as high labour costs, limited wage flexibility and demographics present the main obstacles in improving export performance.

Keywords: regional export performance, Croatia, exports of goods, panel models, manufacturing.

1 INTRODUCTION

Balanced regional development is a key prerequisite for improvement of national well-being. The concept of balanced regional development is present as a key feature in most of European commissions' (EC) documents, recommendations and guidelines. Croatia has strong regional disparities on several dimensions: from individual dimensions such as unemployment [7] to overall indicators of regional economic development such as European regional competitiveness index (RCI) [1], [6] and regional competitiveness index for Croatia [13]. Regional export performance is a crucial element in establishing regional competitiveness. However, as the export variables constitute only a small segment of the RCI this paper adds a more thorough insight into regional export dynamics and its most important determinants.

The recent empirical work for euro area [12] and for Croatia [11] found that the domestic factors, and in particular domestic demand, play an important role in explaining export dynamics. It makes perfect sense to take external conditions as given as Croatia is a small open economy (SOE) and cannot influence them. Therefore, for a more precise pinpointing, this article focuses on potential domestic factors that could influence export dynamics. Furthermore, to avoid intermingling with tourism and export of services which have dynamics of their own, this paper focuses on exports of goods. In comparison to EC's indicators this paper analyses counties instead of NUTS 2 regions, as NUTS 2 regions combine the lowest and highest developed regions in one single unit (region Continental Croatia).

The remainder of the paper is organized in the following manner. Dynamics of Croatian regional export is depicted in section 2. Section 3 provides a description of the data sets as well as the model employed in the analysis of the regional export dynamics. Results of the empirical study are presented and discussed in section 4. Finally, section 5 concludes.

2 CROATIAN REGIONAL EXPORT DYNAMICS IN THE PERIOD 2004 – 2014

For a better picture of the regional export dynamics it is helpful to have an overview on the background, i.e. national level dynamics. Figure 1 depicts Croatian trade dynamics, hence both exports and imports. Due to the internationalization of production chains [5], exporting firms import components. This, in turn, leads to comovements between exports and imports.
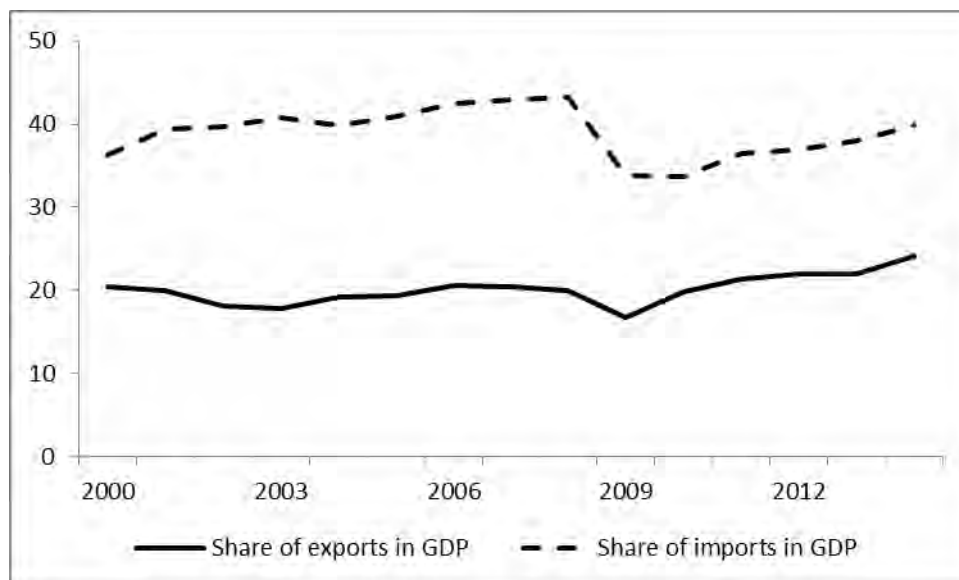
Figure 1: Share of exports and imports in Croatian GDP for the period from 2000 to 2014

Figure 1 confirms strong comovements between exports and imports which indicates that Croatian exports have a strong import component and that both rely on a similar set of factors. Croatian trade dynamics were quite stable in the analysed period with the share of exports and imports in Croatian GDP at around 20% and 40%, respectively. A large decline in domestic demand in 2009 influenced both imports and exports. However, the decline in imports was much more pronounced (around 10 percentage points) compared to exports (around 3 percentage points).

Table 1 presents the share of counties' exports of goods in the regional GDP at the beginning of the analysed period (2004) and in the end of the analysed period (2014). In 2014 Varazdin County had the highest share of exports in GDP (53.9%) while two counties (Lika-Senj and Dubrovnik-Neretva) had the share well below 10%. The lowest share was in Dubrovnik-Neretva County (only 2.3% in 2014). Lika-Senj County increased the share of its exports from 0.9 in 2004 to 6.7% of the regional GDP in 2014. However, this share is still the second lowest in 2014. Most counties increased their exports in 2014 compared to 2004. However, the share of exports in GDP decreased in Sisak-Moslavina (which had the highest share in 2004), Istria and Split-Dalmatia County. A striking figure is that all coastal counties have low share of exports in GDP (below 16%), with the exception of Istria County that is slightly better off at 22.9%. These basic descriptive statistics illustrate low importance of sectors that are considered to be of strategic importance, such as shipbuilding. Therefore, apart from export of services and tourism, exports of goods contribute very little to well-being of coastal counties.

Table 1: Share of counties' exports in regional GDP (2004, 2014 and index 2014 compared to 2004)

County	2004	2014	2014-2004 index
Varazdin	36.1	53.9	149.3
Medimurje	24.9	45.5	182.4
Krapina-Zagorje	27.9	44.8	160.4
Sisak-Moslavina	44.4	39.7	89.6
Koprivnica-Krizevci	16.3	27.9	171.2
Karlovac	22.0	27.3	123.7
Virovitica-Podravina	15.8	26.9	170.0
City of Zagreb	20.6	25.9	125.6
Pozega-Slavonia	15.1	22.9	152.1
Istria	33.8	22.9	67.7
Slavonski Brod-Posavina	8.7	19.7	226.6
Zagreb	7.9	18.7	235.4
Osijek-Baranja	19.1	18.6	97.2
Vukovar-Sirmium	6.0	18.0	300.4
Sibenik-Knin	15.2	15.9	104.7
Primorje-Gorski kotar	12.0	13.5	112.3
Bjelovar-Bilogora	9.4	13.5	142.8
Zadar	10.5	13.0	123.1
Split-Dalmatia	18.9	10.0	52.9
Lika-Senj	0.9	6.7	713.3
Dubrovnik-Neretva	1.7	2.3	137.9

3 DATA AND METHODOLOGY

The empirical analysis is performed for a cross-section of Croatian counties (N=21) for the period from 2004 to 2014 (T=11). The central variable of the study is a share of exports of goods in the regional GDP (variable *export*). Explanatory variables selection includes representatives of verious groups of variables that might influence exports. Namely; as a representative of macroeconomic factors regional gross value added (GVA, variable *gva*) in million HRK was considered. The variable represents the degree of the economic activity in the County and acts as a proxy for the domestic demand. Labour costs are represented by average monthly gross earnings in HRK, deflated by HICP, 2015=100 (variable *wages*). As descriptive statistics indicate strong comovements of exports and imports, a share of imports of goods in the regional GDP are also added as an explanatory variable (variable *import*). Although the share of manufacturing in EU value added and employment has been decreasing, it is still the most important sector for European international trade which takes up over 90% of overall exports of goods [9]. Therefore, a share of manufacturing in counties' GVA (variable *gvaman*) is another explanatory variable used to describe Croatian regional export dynamics. Furthermore, to account for the influence of the economic crisis illustrated in Figure 1, the model additionally includes dummy variable D2009 which equals 1 in a period 2009–2014 and zero for other years.

Additionally, the model also accounted demographic factors related to workforce such as migration, population growth, the share of population with university degree and regional unemployment rate. However, all of the considered variables turned out to be statistically insignificant and hence were not included in the final specification of the model. The lack of statistical significance of the demographic variables indicates that in the future demographic factor could represent a substantial obstacle in improving export performance or even maintaining its current level.

Data on exports, imports, regional and national GDP, average monthly gross earnings, education, regional GVA are obtained from the Croatian Bureau of Statistics (CBS). The source for HICP data is Eurostat.

In the empirical analysis, estimates of static and dynamic panel models are compared. The static panel model is specified as follows:

$$y_{i,t} = \alpha + \beta X_{i,t} + v_i + \varepsilon_{i,t}, \quad i=1,2,...N, \quad t=1,2,...,T, \tag{1}$$

where i denotes a county and t time. Dependent variable $y_{i,t}$ is defined as a share of exports of goods in the regional GDP (variable *export*), and X is a set of regressors. Namely; regional gross value added (*gva*), average monthly gross earnings (*wages*), a share of imports of goods in the regional GDP (*import*), a share of manufacturing in counties' GVA (*gvaman*), and crisis dummy variable (*D2009*).

Although the analysis is based on a cross-section of all Croatian counties and inference is made about the group, which suggests the appropriateness of a panel model with fixed-effects, the formal tests were performed to verify the model selection.

For selecting the appropriate static panel model, two tests were applied: the *F*-test for poolability and the Breusch-Pagan Lagrange multiplier (LM) test for random effects. The *F*-test for poolability (*F*-test for Fixed Effects) is performed to verify if individual (group-specific) effects exist. If the null hypothesis is rejected, the conclusion is that there is a significant fixed effect, i.e. that the fixed effect model is preferred to the pooled ordinary least squares (OLS) model.

The Breusch-Pagan Lagrange multiplier (LM) test examines if any random effect exists. The null hypothesis is that the error variance components are zero. If the null hypothesis is not rejected, the pooled OLS model is preferred. Otherwise, the random effect model outperforms pooled OLS.

To decide between fixed or random effects model, a Hausman test was performed with the null hypothesis that the preferred model is the random effects model vs. the alternative of the fixed effects model.

Additionally, to account for possible dynamics, lagged dependent variable was added in the model. Estimation of a dynamic panel model was also a robustness check for the results obtained by estimation of static panel models. The dynamic model specification is given as follows:

$$y_{i,t} = \alpha + \gamma y_{i,t-1} + \beta X_{i,t} + v_i + \varepsilon_{i,t}, \quad i=1,2,...N, \quad t=1,2,...,T, \tag{2}$$

where counties are denoted by subscript i, while t stands for years. Model (2) is estimated using Arellano-Bover [2] and Blundell-Bond [4] two-step estimator with robust standard errors. To test for the appropriateness of a dynamic model, two specification tests were performed: Arellano-Bond test for zero autocorrelation in first-differenced errors and Sargan test of overidentifying restrictions.

4 EMPIRICAL RESULTS

The estimates with robust standard errors for each model (pooled ordinary least squares OLS, FE and RE static panel models and dynamic panel model) are presented in Table 2. When tested for the significance of individual effects, both *F*-test for fixed effects and Breusch-Pagan Lagrangian multiplier test for random effects indicated the existence of significant differences across counties. The *F*-test for poolability (*F*=84.10, *p*-value=0.0000) supports the significance of fixed effect, i.e. that the fixed effect model is preferred to the pooled ordinary least squares (OLS) model. Breusch and Pagan Lagrangian multiplier test for random effects (Chi-square(1)=762.55, *p*-value=0.0000) indicates that the random effect model outperforms pooled OLS.

Table 2: Estimation results (dependent variable: a share of exports of goods in the regional GDP, *export*).

Variable	OLS	FE	RE	Dynamic panel
Lagged dependent variable				0.356**
				(0.122)
gva	0.0121	.0894	-0.106	-0.0616
	(0.0821)	(0.106)	(0.0700)	(0.183)
wages	-0.00403***	-.00134	-0.000986	-0.0000529
	(0.00120)	(0.00147)	(0.00135)	(0.00184)
import	0.337***	0.494***	0.4756***	0.375*
	(0.0991)	(0.121)	(0.109)	(0.160)
gvaman	0.551***	0.415*	0.428**	0.586***
	(0.0727)	(0.172)	(0.159)	(0.154)
D2009	2.615**	2.945***	2.763***	1.801*
	(0.964)	(0.726)	(0.712)	(0.837)
Model diagnostics				
N	231	231	231	210
F or Wald test	87.79***	10.04***	47.17***	273.53***
R²	0.6191			
SSE or $\hat{\sigma}_\varepsilon$	6.9623	8.5737	7.1044	
$\hat{\sigma}_v$		2.8567	2.8567	

Note: A constant is also included in the model specification but is not reported. Robust standard errors in parenthesis;

*, **, *** statistical significance at 5%, 1% and 0.1% respectively.

Even though both model specifications (FE and RE) give almost identical estimates, formal test was also performed in order to decide which model is appropriate. A robust version of Hausman's specification test, with the null hypothesis that a difference in estimates is not systematic, gives chi-squared test statistic of 10.99 (*p*-value=0.0516) supporting the selection of FE-model.

Regarding specification of dynamic model, Arellano-Bond test for zero autocorrelation in first-differenced errors rejects the null of no autocorrelation at order one (*p*-value=0.0301) while at order two the null is not rejected (*p*-value=0.3762) indicating that there is no serial correlation in the (original) error as desired. The results of Sargan test, suggest that the null hypothesis that overidentifying restrictions are valid cannot be rejected (Chi-square(17)= 17.51853, *p*-value=0.4198). Therefore, both specification tests indicate the validity of the estimated dynamic model. Finally, as the analysed period is relatively short [3] the stationarity of the variables is not explored by means of unit root tests.

The results of the estimated models are similar regardless the estimation method. This resemblance means that the obtained results and conclusions are robust to model specification. The statistical insignificance of the variable that proxies regional economic activity indicates weak domestic demand. Wages are also statistically insignificant which is not surprising considering high labour cost and limited wage flexibility [8], [10]. As for the sequence of statistically significant variables, one has the expected sign (share of manufacturing) while positive effect of crisis dummy variable is a bit puzzling but could be explained by the fact that the relative decrease in export was not as large as the decline in the Croatian economic activity. Positive sign for the imports variable confirms the internationalization of production chains that induces strong comovements between exports and imports.

5 CONCLUSION

The results of the estimated panel models indicate strong comovements between exports and imports thus confirming the internalization of product chains. Increase of share of manufacturing has a positive effect on exports. Consequentially, regions with higher share of manufacturing in regional GVA also have higher exports. It is also important to note that the positive correlation between share of manufacturing and exports increased after EU accession, from 0.6 in 2004 to 0.8 in 2014. Thus, it is of strategic importance to strengthen manufacturing, especially in the coastal regions. On the other hand, weak domestic demand as well as high labour costs and limited wage flexibility highlight most important problems. Furthermore, in the future demographic factor could represent a substantial obstacle in improving export performance.

References

[1] Annoni, P., Dijkstra, L., Gargano, N. 2017. The EU Regional Competitiveness Index 2016. Directorate-General for Regional and Urban Policy. WP 02/2017, Brussels.

[2] Arellano, M., Bover, O. 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics, 68: 29–51.

[3] Blackburne, E. F., Frank, M. W. 2007. Estimation of nonstationary heterogeneous panels. The Stata Journal, 7: 197–208.

[4] Blundell, R., Bond, S. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87: 115–143.

[5] Bussière, M., Chudik, A., Sestieri, G. 2012. Modelling global trade flows: results from a GVAR model. Federal Reserve Bank of Dallas. Globalization and Monetary Policy Institute. Working Paper 119, Dallas.

[6] Dijkstra, L., Annoni, P., Kozovska, K. 2011. A New Regional Competitiveness Index: Theory, Methods and Findings. Directorate-General for Regional Policy. WP 02/2011, Brussels.

[7] Erjavec, N., Jakšić, S. 2016. Differences in Croatian Regional Labour Market Adjustment Mechanisms. Proceedings of the ISCCRO – International Statistical Conference in Croatia, Zagreb, Croatia, 05-06.05.2016., 118-122.

[8] Erjavec, N., Jakšić, S. 2015. Regional unemployment in Croatia: evidence from dynamic panel model. Proceedings of the 13th International Symposium on Operational Research SOR'15, Bled, Slovenia, 23.9.-25.9.2015., 485-489.

[9] European Commission. 2012. A stronger European industry for growth and economic recovery. Industrial Policy Communication Update. SWD(2012) 297 final, Commission Staff Working Document, Brussels.

[10] Jakšić, S. 2017. Explaining regional unemployment in Croatia: GVAR approach. Revija za socijalnu politiku, 24 (2): 189-217.

[11] Jakšić, S., Žmuk, B. 2014. Modelling Croatian Export Dynamics Using Global Macroeconometric Model. Zagreb International Review of Economics and Business, 17, Special Conference Issue: 31-48.

[12] Rua, A., Soares Esteves, P., Staehr, K., Bobeica, E. 2015. Exports and domestic demand pressure: a dynamic panel data model for the euro area countries. European Central Bank. Working Paper Series 1777, Frankfurt am Main.

[13] Singer, S., Gable, J. 2014. Regional Competitiveness Index Croatia 2013. National Competitiveness Council and United Nations Development Programme (UNDP), Zagreb.

THE RELATIONSHIP BETWEEN SUSTAINABLE PROFIT AND SUSTAINABLE BUSINESS IN COMPANIES IN CROATIA

Vedran Kojić, Tihana Škrinjarić and Nidžara Osmanagić Bedenik
University of Zagreb, Department of Mathematics and Department of Managerial Economics
Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia
vkojic@efzg.hr, tskrinjaric@efzg.hr, nosmanagic@efzg.hr

Abstract: In this paper, we present theoretical and empirical analysis of the relationship between sustainable profit and sustainable business in company's business by using the phase diagram model. The research is based on the research studies from Thailand. This paper has two main objectives. The first objective is to give some theoretical remarks and improvements to the analysis of the model, which are missing in the initial research. The second objective is to apply this methodology on the sample of Croatian companies in order to help domestic companies to improve their business by achieving the sustainable profit and sustainable business. To the best of our knowledge, this is the first research in Croatia of this kind.

Keywords: mathematical modelling, phase diagram, econometric methods, sustainable profit, sustainable business, Croatian Business Council for Sustainable Development

1 INTRODUCTION

Generally, it is known that profit maximization is not the unique and the most important goal of the company's business. On the contrary, another very important goal is the certain level of sustainable business. Thus, it is very important to describe the relationship between profit maximization and investment into sustainability and how these two goals affect one another. Beside other research studies, in the recent literature we can find papers which deal with mathematical modelling of influence of corporate social responsibility on business, and vice versa. This includes certain mathematical methods such as bi-level programming or multicriteria optimization (see for instance, [4] and [5]). New foreign research studies show how important is monitoring the relationship between sustainable profit and sustainable business. This relationship can be mathematically described and econometrically measured by using the phase diagram, as it is shown in the case of several companies in Thailand (see [1], [8], [9] and [10]). Although there are research studies about non-financial reporting and sustainable business in Croatia (see for example [6] and [7]), applications of quantitative models are still missing. Therefore, the idea and purpose of this paper is to apply the phase diagram approach to Croatian companies and to see how domestic companies can improve their sustainable business without deterioration of sustainable profit and vice versa. To the best of our knowledge, this paper is the first application of the phase diagram approach to the Croatian companies. The structure is as follows. After introduction, the second section deals with methodology, data and empirical analysis, while the third section presents results and discussion. The final, fourth section gives conclusion and guidelines for further research.

2 METHODOLOGY, DATA AND EMPIRICAL ANALYSIS

We followed the methodology given in the papers [9] and [10]. It is important to emphasize that sustainability is newer approach in business, so mathematical research studies in this field are new as well.

2.1. Construction of the phase diagram of the sustainable profit and sustainable business

In the paper [10], Suriya et al. started from the profit function $\pi = PQ - (FX + \delta FX + CQ + RD + S)$, where π is the profit, P is the price and Q is quantity sold to the market, FX is the fixed cost, δFX is maintenance cost when δ is the depreciation rate, CQ is the variable cost where C is the unit cost, RD is the research and development expenditure, and S is expenditure for the ecological (environmental) component of sustainable business. Let $\Phi = P - C$ be a unit profit. Further, define the effect of S on the quantity sold to the market as $Q = (\ln S)^{\alpha}$, where α is a positive real number.[1] By taking some simple manipulations, from [9] we know that the time derivative of the profit, i.e. the profit change over time is given by

$$\dot{\pi} = (\ln S)^{\alpha} \dot{\Phi} + \Phi \alpha (\ln S)^{\alpha-1} \ln \dot{S} - (1+\delta) FX - RD - e^{\ln S} \ln \dot{S} . \qquad (1)$$

For simplicity, it is assumed that $\ln \dot{S}$, $\dot{\Phi}$, FX, RD are positive constants. The condition of the sustainable profit in the lon run is given by the no change of the profit over time, i.e., mathematically it means that the first time derivative of the profit function equals zero: $\dot{\pi} = 0$. Thus, by taking the left side of (1) to be zero, we get the unit profit as a function of ln$S$:

$$\Phi = \frac{(1+\delta) FX + RD + e^{\ln S} \ln \dot{S} - (\ln S)^{\alpha} \dot{\Phi}}{\alpha (\ln S)^{\alpha-1} \ln \dot{S}} . \qquad (2)$$

In the paper [10] it is shown that $\lim\limits_{\ln S \to 0^+} \Phi = +\infty$. However, finding the limit $\lim\limits_{\ln S \to +\infty} \Phi$ is missing. Thus, let us find it. By using the well-known result $\lim\limits_{x \to +\infty} \dfrac{x^p}{e^x} = 0$, $p \in \mathbb{R}$ (see for instance [14]), it is easy to see that

$$\lim\limits_{\ln S \to +\infty} \Phi = \lim\limits_{\ln S \to +\infty} \frac{(1+\delta) FX + RD + e^{\ln S} \ln \dot{S} - (\ln S)^{\alpha} \dot{\Phi}}{\alpha (\ln S)^{\alpha-1} \ln \dot{S}} = +\infty . \qquad (3)$$

Now, in the paper [10], Suriya et al. took $\alpha=1$ to conclude that the function $\Phi=\Phi(\ln S)$ is U-shape with global minimum, such that its whole graph is contained in the quadrant I of a Cartesian coordinate system. However, the assumption $\alpha=1$ is not necessary. In fact, it can be proved that it is sufficinet to assume that $\alpha\geq 1$ to get the same conclusion (but this analysis is beyond the scope of this paper).

On the other hand, sustainable business in the long run is given by $\dot{S} = 0$. Using the condition when the first time derivative of S equals zero, the Suriya et al. (see [10]) concluded that ln$S$ as a function of Φ does not depend of Φ:

$$\ln S = \left(\frac{\dot{\pi} + (1+\delta) \dot{FX} + \dot{RD}}{\dot{\Phi}} \right)^{\frac{1}{\alpha}} . \qquad (4)$$

Now, by calculating the partial derivatives $\dfrac{\partial \dot{\pi}}{\partial \Phi}$ and $\dfrac{\partial \dot{S}}{\partial \ln S}$, the Suriya et al. (see [10]) constructed and described the phase diagram of the sustainable profit and sustainable business given by the Figure 1 a). The point E (see Figure 1 a), b) and c)) is the intersection

---

[1] The assumption $\alpha>0$ is very important and it is missing in the paper [10]. In fact, the assumption $\alpha\geq 1$ is needed, as it will be seen in the further analysis.

of the functions for which $\dot{\pi}=0$ and $\dot{S}=0$ hold. It represents the steady state, i.e. the equilibrium point (it is called the point of double or twin stability) where the both goals are achieved: sustainable profit and sustainable business (in the long run). The phase diagram contains four areas. Suriya et al. (see [10]) named them as warm glow area, frozen area, decayed area and charitable area (see Figure 1 b) and c)). For instance, in the warm glow area stream lines force the company to go only to sustainable profit, but not to sustainable business (see Figure 1 b)). In that case, to achieve the equilibrium E, the company should reduce the the unit profit and expenditure for sustainable business. In fact, as it is said in [10], the company located at positions A, B, C, or D (see Figure 1 c)) cannot achieve both goals automatically without policy modification. The company should adjust its profit policy and sustainable business policy according to its location in the phase diagram.
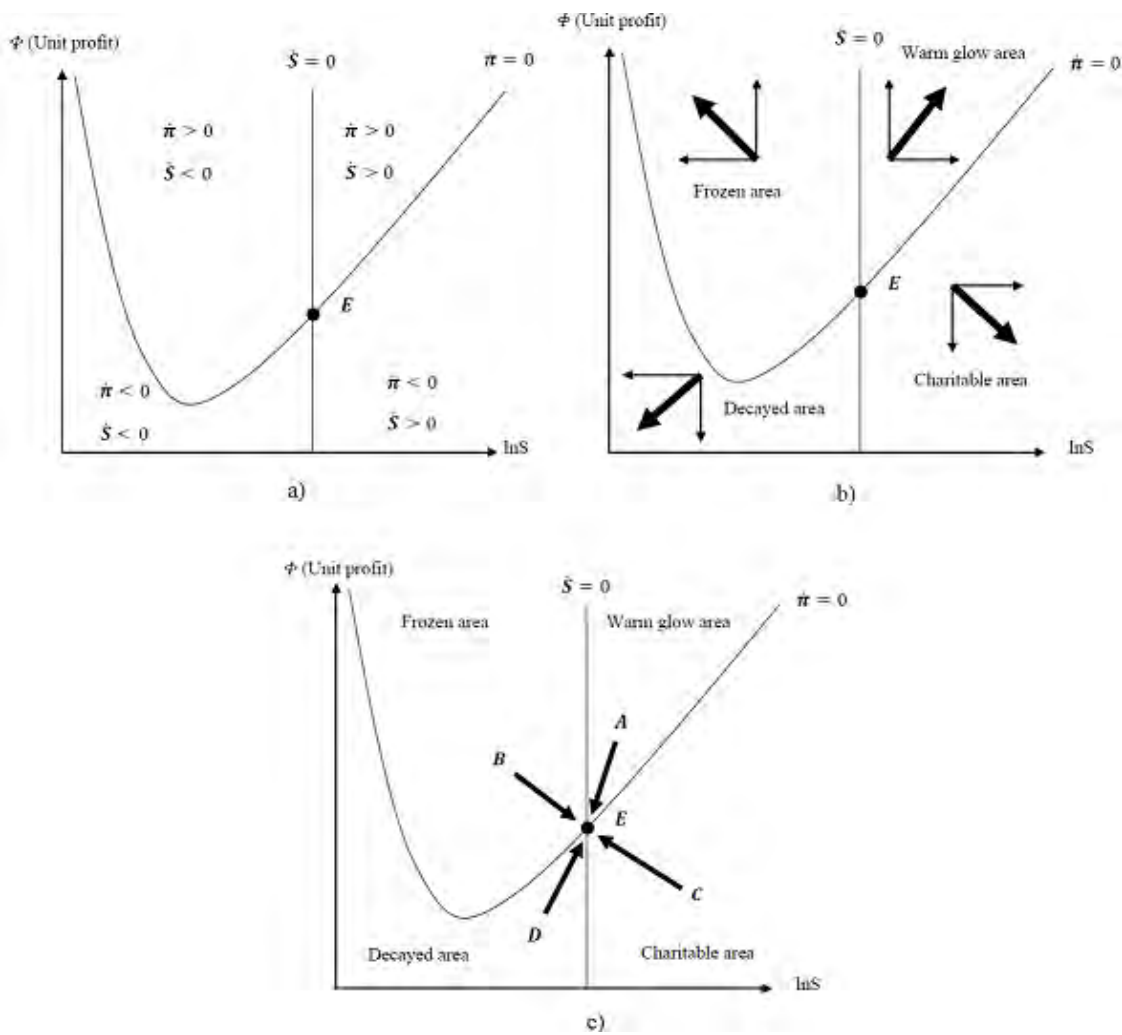
Source: Taken from [10].

Figure 1: a) Phase diagram of the sustainable profit and sustainable business. b) Phase diagram with stream lines and area classification. c) Policy adjustment to move the company toward the sustainable profit and sustainable business.

2.2. Data collection

In this research, we have studied the sample of the 46 companies – members of the Croatian Business Council for Sustainable Development, which is member of World Business Council

for Sustainable Development. Only 14 companies had reachable online data both for profit[2] and expenditure for ecological (environmental) component of sustainable business. In general, sustainability encompasses three components: economic, social and ecological component, where, in our case, data for social component is given only in non-monetary way[3].

2.3. Econometric estimation

The usual methodology to estimate the parameters of term $\Phi(\ln S)^{\alpha}$ is panel data estimation. The basic model is pooled model, defined as:

$$y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{itk} + \varepsilon_{it}, \ i \in \{1, 2, ..., N\}, t \in \{1, 2, ..., T\},$$ (5)

where N is the number of companies, T number of time periods, y_{it} the dependent variable for i-th company in period t, parameter β_0 is constant for each company in each period, x_{ijk} independent variable k, β_k parameters and ε_{it} the error terms. Pooled model is assumed to be used when the sample is random. Two other models which are used more frequently in research are the fixed effects and random effects models. The fixed effects model can assume that the constant term in (5) changes for each company, in each period or both. If we assume that the constant changes both for each company and in each period, model (5) can be written as:

$$y_{it} = \beta_i + \lambda_t + \sum_{k=1}^{K} \beta_k x_{itk} + \varepsilon_{it}.$$ (6)

The random effects model assumes that the error term in (5) changes for each company, in each period or both. Thus, (5) can be rewritten as:

$$y_{it} = \beta_0 + \sum_{k=1}^{k} \beta_k x_{itk} + \gamma_t + e_i + \varepsilon_{it}.$$ (7)

Usual assumptions in these models and methods of estimation can be seen in [11] and [12] or [13]. In order to use the most adequate model, an F test is used to compare the fixed effects model to the pooled regression and Hausman test (see [3]) is used to compare the fixed effects to random effects model. Details can be seen in [2].

3 RESULTS AND DISCUSSION

Firstly, pooled regression, model with fixed effects and with random effects have been estimated. The results of the Hausman test indicated that the model with cross section random effects does not yield consistent estimates (test value is equal to 14.97 with p-value 0.000). Period random and period fixed effects both result with consistent estimates (test value is equal to 2.94 with p-value 0.086). When comparing fixed effects model to the pooled regression, F test indicates that fixed effects should be used for cross section (test value is equal to 34.99 with p-value 0.000), but not for period effects (test value is equal to 0.74 with p-value 0.574). Thus, model with cross section fixed effects was estimated in order to obtain values of unit profit Φ and parameter α:

$$\ln Q = 10.54 + 2.41 \ln(\ln S).$$ (8)

---

[2] Profit before taxation.
[3] The data can be requested from authors.

In the next step, individual unit profit for each company i in period t ($\hat{\Phi}_{it}$) have been obtained from equation (8). Now we can observe phase diagrams for individual companies. In order to do so, we need to estimate Φ as the function of $\ln S$ when $\dot{\pi} = 0$ and the value of $\ln S$ when $\dot{S} = 0$. Function $\Phi(\ln S)$ for each company was estimated as:

$$\Phi_t = \gamma_0 \frac{1}{\hat{\alpha}(\ln S_t)^{\hat{\alpha}-1}} + \frac{e^{\ln S_t}}{\hat{\alpha}(\ln S_t)^{\hat{\alpha}-1}} + \gamma_1 \frac{\ln S_t}{\hat{\alpha}} + \varepsilon_t \qquad (9)$$

where Φ_t has been used from (8), $\Phi_t \equiv \hat{\Phi}_{it}$, as well as $\hat{\alpha} = 2.41$. $\ln S_t$ was obtained from original data on S for each company. Estimated values of γ_0 and γ_1 are used to retrieve the value of $(1+\delta)FX + \dot{RD}$ in equation (2). To calculate the value of $\ln S$ when $\dot{S} = 0$ we use the equation (4), for which we now have all the needed elements. $\dot{\pi}$ and $\dot{\Phi}$ were calculated from the data as averages for the whole period. Only 2 out of 14 companies resulted with positive value of $\ln S$. Thus, we provide more detailed results for those 2 companies.
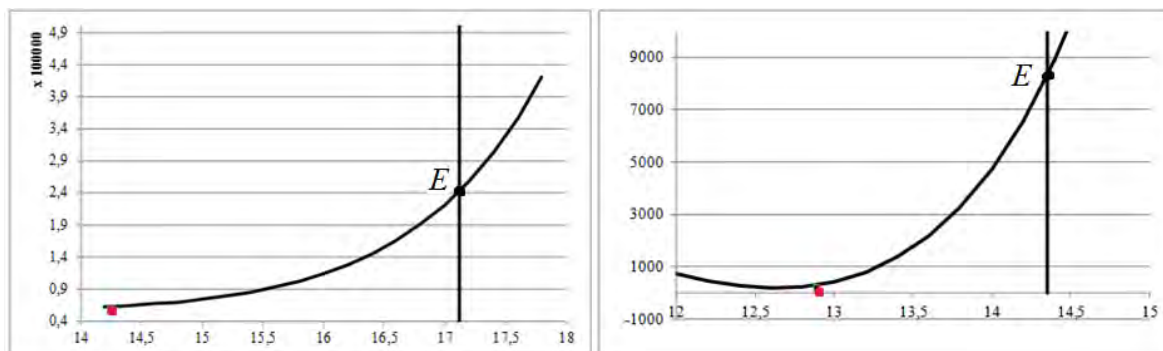
Source: author

Figure 2: Left is the location of the Company 1 and right is the location of the Company 2.

The equilibrium point E of the Company 1 has the coordinates (17.117, 242055.1). In 2015, this company had actual coordinates (14.215, 57184.01), which means it was in the decayed area. The other company, Company 2, has the following coordinates: E(14.349, 8259.74) and 2015(12.888, 212.382). Thus, both companies are in the same (decayed) area. To achieve the point E, i.e. sustainable profit and sustainable business in the long run, both firms should increase the profit and expenditure for sustainable business. One can say that the usefulness of the proposed phase diagram model is low, since only 2 of 14 Croatian companies fitted into this analysis, and we are aware of that remark. However, Suriya et al. (see [10]) are also aware of disadvantages of this model, but they gave the same analysis for 26 companies in Thailand and only 4 companies fitted (see [8]). Including more data in the future and estimating other functional forms of relationship between unit profit and $\ln S$ could result with better outcomes. Although this model has to be improved, we believe that it is a good starting point in the analysis of the relationship between sustainable profit and sustainable business.

4 CONCLUSIONS AND FURTHER RESEARCH

In this paper, we analysed the relationship between sustainable profit and sustainable business in the long run on the sample of 14 companies in Croatia by using the phase diagram model proposed by the scholars from Thailand. We added some remarks in the mathematical analysis of the model, which are missing in the original papers from Thailand, such as the U-shape of the function $\Phi=\Phi(\ln S)$ and sufficiency of the assumption that $\alpha \geq 1$. From the empirical analysis side, we have shown that only 2 of 14 Croatian companies fitted

into this analysis, where both of the companies are located into decayed area. To achieve the equilibrium point, i.e. sustainable profit and sustainable business in the long run, we proposed the appropriate policy modification. To the best of our knowledge, this is the first attempt of quantitative modelling and describing of the relationship between sustainable profit and sustainable business in companies in Croatia, but we hope not the last one. For instance, for further research, we are planning to increase the sample of the Croatian companies, compare the results according to the company's size (large, medium and small organization) and industries. In the end, recall that the effect of the expenditure for the sustainable business on the quantity sold to the market is given as the power of the ln$S$. It will be interesting to see what improvements we can get if we assume that the mentioned effect is given by different function form.

References

[1] Chanthao, P., Suriya, K. (2014). An analysis of the profit sustainability and social responsibility of companies listed in the food and beverage sector in the Stock Exchange of Thailand. The Empirical and Quantitative Economics Letters, 3 (1): 55-66.

[2] Greene, W. H. (2003). Econometric Analysis, 5th edition. Prentice Hall.

[3] Hausman, J. (1978). Specification Tests in Econometrics. Econometrica, 46 (6): 1251-1271.

[4] Hseuh, C-F. (2015). A bilevel programming model for corporate social responsibility collaboration in sustainable supply chain management. Transportation Research Part E, 73: 84-95.

[5] Huo, Y. (2012). Supply Chain Network Equilibrium Model Based on Corporate Social Responsibility with Multicriteria under the Revenue-Sharing Contract. Advanced Materials Research, 452-453: 282-288.

[6] Osmanagić Bedenik, N., Strugar, I., Labaš, D., Kojić, V. (2016). Nonfinancial Reporting – The Challenge of Sustainable Business (Empirical research in companies in Croatia). Zagreb: M.E.P.

[7] Osmanagić Bedenik, N., Prebežac, D., Strugar, I., Kojić, V. (2017). Non-financial Reporting in Hotel Enterprises in Croatia. Zagreb: Naklada Veble.

[8] Srikaew, P.,Suriya, K. (2014). Profit sustainability and social responsibility of companies listed in the media and publishing sector in the Stock Exchange of Thailand. The Empirical Econometrics and Quantitative Economics Letters, 3 (1): 67-76.

[9] Suriya, K., Sudtasan, T. (2014). How to estimate the model of sustainable profit and corporate social responsibility. Business and Economic Horizon, 9 (4): 1-7.

[10] Sudtasan, T., Suriya, K. (2013). Sustainability of profit and corporate social responsibility:Mathematical modelling with phase diagram. The Empirical Econometrics and Quantitative Economics Letters, 2 (4): 1-12.

[11] Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge: MIT Press.

[12] Wooldridge, J. M. (2008). Introductory Econometrics: A Modern Approach. Thomson South-Western, Mason, OH.

[13] Verbeek, M. (2005). A Guide To Modern Econometrics, Second edition. London: John Wiley & Sons.

[14] Zorich, V. A. (2004). Mathematical Analysis I. Berlin: Springer Verlag.

TECHNICAL NOTE: THE SHAPE OF THE MACAULAY'S DURATION AS THE FUNCTION OF COUPON BOND MATURITY DERIVED WITHOUT DERIVATIVES

Vedran Kojić and Zrinka Lukač
University of Zagreb, Department of Mathematics
Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia
vkojic@efzg.hr, zlukac@efzg.hr

Abstract: In literature, the common approach is to consider Macaulay's duration of coupon bonds as a differentiable function. However, in reality bond maturity is a discrete variable, meaning that Macaulay's duration, as a function of maturity, is in fact a sequence of real numbers. It is not a differentiable function. Therefore, the analysis of properties of Macaulay's duration by using the differentiable calculus is not justified. There are some papers known in the literature which analyse properties of Macaulay's duration without the use of calculus, however the results presented there are not complete. In this paper we fill the gap by pointing out the shortcomings of the existing results regarding the non-calculus approach and completing the analysis of Macaulay's duration considered as a sequence of real numbers.

Keywords: coupon bond, Macaulay's duration, bond maturity, sequence of real numbers, without derivatives

1 INTRODUCTION

In this paper we consider the properties of Macaulay's duration of coupon bonds by using the non-calculus approach. The common approach to this problem is to use differential calculus of real functions with several real variables (see for instance [2], [4], [5], [6], [9], [10]), although there are also papers which approach the problem of the bond's price properties by using a student-friendly approach via non-calculus ([7] and [8]). The main obstacle to using calculus in analysis of the Macaulay's coupon bond duration is the fact that duration with respect to coupon-rate (cet. par.), yield to maturity (cet. par.) or maturity (cet. par.) is not a continuous function, not to mention differentiable. In practice, coupon rate, yield to maturity and maturity are discrete variables, so the duration has to be seen as the sequence of real numbers, rather then a continuous function. Therefore, in papers [1] and [3] we have presented the analysis of the Macaulay's coupon bond duration properties by using a non calculus approach. However, in [3] the proof of the shape of the Macauly's duration as the function of coupon bond maturity is not complete, so in this paper we present the complete proof.

The structure of this paper is as follows. After the introduction, in the second section we present the notation. The third section presents previous results, while the fourth section contains improved analysis and complete proof of the main problem of this note. The fifth section is conclusion.

2 NOTATION

We use the same notation as in [1] and [3]:

N	face value, bond's par value
i	contractual interest rate, bond's coupon-rate
I	annual coupon payment, a year interest payment
n	bond maturity, number of payments, n years

k	annual yield to maturity of the bond
$r=1+k$	annual period discount factor at rate k
P	(market) bond's price
$P(k, i, n)$	market price bond as a function of k, i and n
$P(k)$	market price bond as a function of k (cet. par.)
$P(i)$	market price bond as a function of i (cet. par.)
$P(n)$	market price bond as a function of n (cet. par.)
D	Macaulay's bond duration
$D(k, i, n)$	Macaulay's bond duration as a function of k, i and n
$D(k)$	Macaulay's bond duration as a function of k (cet. par.)
$D(i)$	Macaulay's bond duration as a function of i (cet. par.)
$D(n)$	Macaulay's bond duration as a function of n (cet. par.)

3 PREVIOUS RESULTS

Here we give an overview of the results given in [3].

Lemma 1. For all real numbers a and b such that $a>0$ and $b>0$, there exists a positive integer n such that $na > b$.
Proof. In literature, this lemma is called **The Principle of Archimedes**. For proof, see for example [11].

Theorem 2. Properties of D with respect to change of its variables are as follows:
  (a) Value of duaration is always positive.
  (b) Value of duration is always less or equal to bond's maturity, i.e. $D(n) \leq n$.
  (c) Function $D=D(i)$ is monotonically decreasing ceteris paribus, i.e. for two coupon-rates $i_1 \leq i_2$ inequality $D(i_1) \geq D(i_2)$ holds.

  (d) Function $D=D(k)$ is monotonically decresing ceteris paribus, i.e. for two yields to maturity $k_1 \leq k_2$ ineqaulity $D(k_1) \geq D(k_2)$ holds.

  (e) The limit of function $D=D(n)$ when n approaches infinity is $L_D = \lim_{n \to +\infty} D(n) = 1 + \dfrac{1}{k}$. In

case $i \geq k$, function $D(n)$ is increasing with maximal asymptotic value L_D. However, in case of $0 < i < k$, the shape of the duration as the function of maturity $D(n)$ is as follows: duration first increases, reaches its maximum, and after that decreases and asymptotically approaches the limit value L_D.

In [3], we have presented the proof for statements (a), (b), (c) and (d) of Theorem 2. As for the statement (e), [3] contains the proof for the case of $i \geq k$ and the proof for that case is complete. However, the proof of (e) in the case of $0 < i < k$ is not complete. Therefore, in this paper we bring the complete proof for this case.

Let us first quote the proof of (e) in the case of $0 < i < k$ from [3].

Proof of Theorem 2, statement (e), case $0 < i < k$:
In [3], we have derived the following formula:

$$P(n)[D(n+1) - D(n)] = \frac{N}{(1+k)^{n+1}}[-(i-k)D(n+1) + (ni+i+1-nk)]. \qquad (1)$$

Since $P(n)>0$, from (2) it follows that the sign of the term $D(n+1)-D(n)$ is positive, i.e.

negative, if and only if the right hand side in (2) is positive, i.e. negative.

There are two cases. In the case of $i \geq k$ the proof is complete and it can be found in [3]. The proof of the second case is given as follows.

In the second case, if $i < k$, i.e. $k - i > 0$, let us show that then there exists a positive integer n_0 such that for all positive integers $n > n_0$ the second factor in the square brackest on the right hand side of (1) is less than 0. Thus, using (1), it is trivial to see that inequality

$$P(n)\big[D(n+1) - D(n)\big] \leq 0, \tag{2}$$

is equivalent to

$$(k-i) \cdot \big(D(n+1) - n\big) + i + 1 \leq 0, \tag{3}$$

i.e.,

$$(k-i) \cdot D(n+1) + i + 1 \leq (k-i) \cdot n. \tag{4}$$

Since $D(n+1)$ approaches limit L_D when n approaches infinity, it means that for given $\varepsilon > 0$ there exists positive integer n_0 such that for all positive integers such that $n > n_0$ inequalities $-\varepsilon + L_D < D(n+1) < L_D + \varepsilon$ hold. To prove inequality (4), let us prove equivalent but stronger inequality

$$(k-i) \cdot (L_D + \varepsilon) + i + 1 < (k-i)n. \tag{5}$$

Given that

$$k - i > 0, \tag{6}$$

let $a = k - i$ and $b = (k-i)(L_D + \varepsilon) + i + 1$. Notice that a, $b > 0$. By applying Lemma 1., we know that there exists positive integer n' such that (5) holds for all $n > n'$. Thus, for all integers $n > \max\{n_0, n'\}$ the inequality (2) holds, which means that after the nth term (that is after the element of the seqence on the position $\max\{n_0, n'\}$, to be more precise) the function $D(n)$ becomes a decreasing function, and as n approaches infinity, $D(n)$ approaches limit L_D. Notice that in this case ($i < k$), $D(n)$ reaches its maximum which is greater than L_D, and after that decreasingly and asymptotically approaches the limit value L_D. It is easy to show that

$$D(1) = 1 < D(2), \tag{7}$$

and $D(1) = 1 < 1 + \dfrac{1}{k} = L_D$. Therefore, $D(n)$ first increases, reaches its maximum which is necessarily greater than L_D and after that begins to decrease and asymptotically approaches its limit L_D. Figure 1. shows the shape of the function $D = D(n)$, where n is a continuous variable.
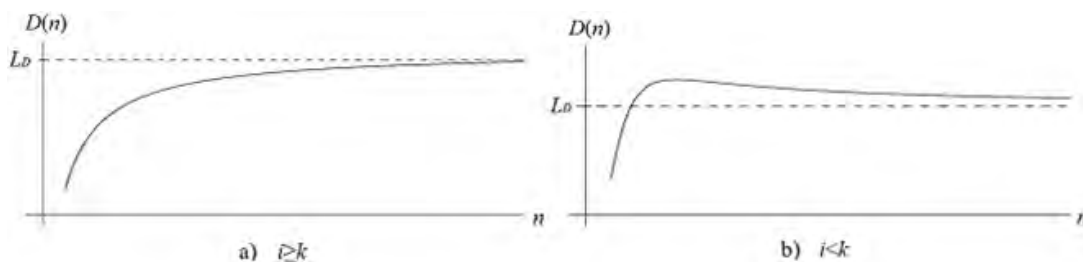
Source: Taken from [3].

Figure 1: The shape of the Macaulay's duration as the function of maturity

This is the end of the proof of the statement (e) taken from [3]. Figure 1, taken from [3], illustrates the shape of the duration as the function of maturity.

4 IMPROVED ANALYSIS

The proof presented in the previous section is not complete. In fact, it proves only the case shown in Figure 2 a). However, there is an open question remaining why the cases shown in Figures 2 b) and c) are not possible.
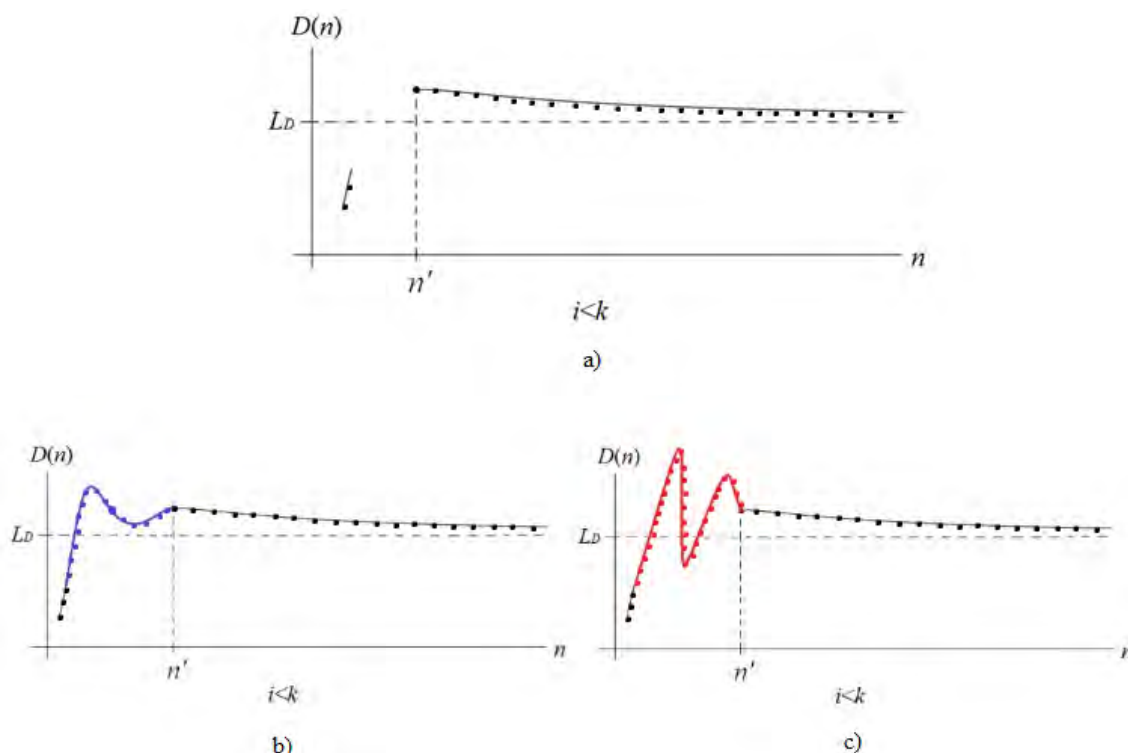
Source: authors

Figure 2: Hypothetical possibilities for the shape of Macaulay's duration as the function of maturity

From [3], by using The principle of Archimedes, we know that there exists a number $n*$ such that sequence $D(n*)$, $D(n*+1)$, $D(n*+2)$, $D(n*+3)$, … is decreasing. Let n_{min} be the smallest positive integer number such that $D(n)$ after the $n_{min}{}^{th}$ element is a decreasing sequence. Then, it is obvious that inequlities

$$D(n_{min}-1) \leq D(n_{min}) \geq D(n_{min}+1) \tag{8}$$

hold (otherwise we would have the contradiction with the definition of the number n_{min}). Thus, $(n_{min}, D(n_{min}))$ is a local maximum of the function $D(n)$.

It remains to prove that $(n_{min}, D(n_{min}))$ is the uniqe (global) maximum. Consider the following equation:

$$D(n+1)-D(n)=0. \tag{9}$$

Let $l(n)=k(k-i)n-k^2-2k+i$ and $d(n)=\dfrac{(i-k)^2}{(1+k)^n}$. It is easy to see that equation (9) is

equivalent to the equation

$$l(n)=d(n). \tag{10}$$

335

Also, obviously $l(0) = -k^2 - 2k + i = -k^2 - k - (k - i) < 0$ and $d(0) = (i - k)^2 > 0$. Since $l(n)$ is strictly increasing line and $d(n)$ is strictly decreasing exponential function, it is obvious that equations (10) and (9) have at most one positive integer-valued solution, but exactly one real-valued solution (see Figure 3).
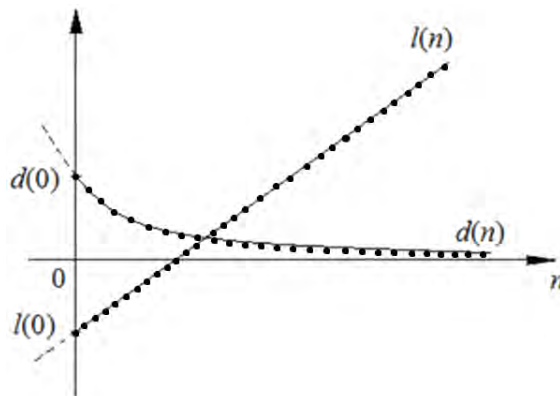
Source: authors

Figure 3: Intersection of the functions $l(n)$ and $d(n)$

Thus, this gives the complete proof of the statement (e) of Theorem 4.

5 CONCLUSION

The common approach in the financial modeling literature is to consider the Macaulay's duration of coupon bonds as a differentiable function. In accordance to this assumption, the properties of Macaulay's duration as a function of coupon-rate i (ceteris paribus), yield to maturity k (ceteris paribus) and maturity n (ceteris paribus) are also being analysed by using differential calculus. However, in reality Macaulay's duration is a sequence of real numbers because variables i, k and n are discrete, meaning that the use of differential calculus is not justified. There are some recent papers (see [1] and [3]) which derive the properties of Macaulay's duration without the use of calculus. However, the analysis of Macaulay's duration as a function of n presented in these papers is not complete. Therefore, in this paper we fill the gap and complete the above mentioned analysis without the use of calculus. It is important to highlight that the properties of Macaulay's duration of coupon bonds derived without the use of differentiable calculus confirm the results already obtained by the use of calculus.

References
[1] Gardijan, M., Kojić, V., Šego, B. (2012). Bond Duration: Propositions and Applications. In: Aljinović, Z., Marasović, B., ed., in Mathematical models in analysis of the Croatian financial market (5-26). Split: University of Split - Faculty of Economics,.

[2] Fabozzi, F. J. (2006). Bond Markets, Analysis and Strategies. New Jersey: Prentice Hall Inc.

[3] Kojić, V., Lukač, Z. (2015). A non-calculus Analysis of Macaulay's Coupon Bond Duration Properties. Proceedings of the 3rd International Symposium on Economics and Social Sciences, Tokyo, Japan, 156 – 168.

[4] La Grandville, O. (2000). Bond pricing and portfolio analysis. London: MIT Press.

[5] Hawawini, G. A. (1982). On the Mathematics of Macaulay's Duration. In Hawawini, G. (ed), Bond Duration and Immunization: Early Developments and Recent Contributions, Garland. Retrieved from https://flora.insead.edu/fichiersti_wp/inseadwp1982/82-03.pdf [26/4/2015]

[6] Hawawini, G. (1984). On the relationship between Macaulay's Bond Duration and the Term to Maturity. Economics Letters, 16: 331-337.

[7] Lawrence, E. R., Shankar, S. (2007). A Simple and Student-friendly Approach to the mathematics of bond proces. Quarterly Journal of Business & Economics, (46) 4: 91-99.

[8] Malkiel, B. (1962). Expectations, Bond Prices and the Term Structure of Interest Rates. Quarterly Journal of Economics, 76: 197-218.

[9] Pianca, P. 2005. Maximum Duration of Below Par Bonds: A Closed –form Formula. http://dx.doi.org/10.2139/ssrn.738445 [13/6/2017]

[10] Smith, D. J. (1998). A Note on the Derivation of Closed-Form Formulas for Duration and Convexity Statistics On and Between Coupon Dates. Journal of Financial Engineering, 7 (2).

[11] Zorich, V. A. (2004). Mathematical Analysis I. Berlin: Springer–Verlag.

ON DEFINITION OF SAMPLE QUARTILES

Darja Rupnik Poklukar

University of Ljubljana, Faculty of mechanical engineering, Aškerčeva 6, Ljubljana
e-mail: `darja.rupnik@fs.uni-lj.si`

Janez Žerovnik

University of Ljubljana, Faculty of mechanical engineering, Aškerčeva 6, Ljubljana
e-mail: `janez.zerovnik@fs.uni-lj.si`

Abstract: There are a large number of various methods for finding sample quartiles for discrete data. The results of these methods are different on small datasets, and may be a source of confusion when teaching elementary of introductionary statistics courses. Convergence properties of various methods on large samples are briefly diuscussed.

Keywords: Quartiles, Percentiles, Sample quartiles.

1   INTRODUCTION

It is well known among statisticians that there are a number of different definitions in the literature of the first and third quartile values of a finite data set. Quartiles are often introduced in national curricula at secondary and even at elementary school level early together with box-and-whisker plots. However, as statistics deals with variation and uncertainty, a sound introduction of statistical notions even at elementary school level is not necessarily straightforward. Already a brief look at some textbooks for secondary and elementary school shows that this seemingly trivial exercise, the computation of sample quartiles may be far more than confusing for the pupils because in the textbooks [1, 10] there are given different methods yielding different results ! The problem does not only arise in Slovenia. The differences in definition of quartiles for quantitative data are well known elsewhere. For school or introductory students in english speaking world, the situation is very similar. Many definitions are offered (see [6, 7, 2], and the references there). Furthermore, different definitions are used by some statistical computing packages [4].

We have discussed the problem from teachers' perspective elsewhere [12, 13], and argued that Langford's method is the best choice for considering quartiles at elementary level. In [7] Langford writes "the situation is, I believe, far worse than most realize". He examined various methods which are actually used in elementary statistics textbooks and the methods employed by various commonly-used calculator and computer packages, and using a precise definition of percentile, identified which of the methods satisfy this definition. More elementary methods (in more detail) are provided in [12, 13].

When discussing the issue with (anonymous) reviewers and editors it seems that in general, the statisticians do not find the issue to be very serious. Namely, it is believed that in statistics, due to consideration of large populations, large samples, and many repetitions the difference in the definitions of quartiles will vanish: "It is of the utmost importance to emphasize that data quartiles and medians are providing estimates of theoretical or population values, and that differences we see in small samples in ways of calculating data values, do not tend to matter in large samples." At first, we agreed with the statement [13], even illustrated the vanishing effect with an example in [13]. After a second thought, we however think that the inconsistencies may not be so easy to overcome, or at least a deeper argument is needed. In particular, the questions related may be much more tricky when considering discrete distributions.

In contrast to elementary school level, where only the quartiles of finite dataset are regarded aiming to provide a box-and-whisker plots to present the data in a nicer way, in statistics, a set of observed values (maybe with repetitions) is usually a sample drawn from some distribution, and a natural question is how to estimate the quartiles of the original distribution from the given

sample. Sometimes there is additional information available about the type of distribution that may allow some assumptions giving arguments for various methods for estimation of quartiles yielding a variety of methods for computing the quartiles in the literature and in software [4, 7, 2]. While there may be good reasons to use different algorithms (and different definitions !) for quartiles in various contexts, it is very important to know that various methods will either converge (the differences will vanish) or to understand that they may provide different results.

In the next section, definition of sample percentiles is recalled and simple examples of sample quartiles and medians given using it. In Section 3, the calculation method proposed by Langford is recalled and compared with two frequently used methods that fail to provide the results in accordance with previously proposed definition that is recommended to be used at elementary level [7, 12, 13]. While for continuous distributions, the differences among the various methods tend to vanish as the sample sizes are growing, the issue is much more complicated when considering discrete distributions. In the last section, we provide an example and briefly discuss the natural questions that arise.

2   DEFINITIONS AND EXAMPLES

For a continuous random variable X with cumulative distribution function $F(x) = P(X \leq x)$, the p-th quantile q_p is defined by $F(q_p) = p$. The p-th quantile is often called the $100p$-th percentile. The quartiles are special cases of percentiles: **first quartile** $Q_1 = q_{0.25}$, **second quartile** (or **median**) $Q_2 = M = q_{0.5}$, **third quartile** $Q_3 = q_{0.75}$.

For discrete quantitative random variables, the definition of quantiles and quartiles is much more complicated. For data, the most common idea is to arrange the data in ascending order and divide them into fixed number of roughly equal parts. ("For finding quartiles, divide the ordered sample into four quarters having the same number of observations in each quarter", [6]) However, there are some differences in the details and also different definitions are used by various statistical software packages [4]. We do not discuss all the variants, and restrict ourselves to three most popular, in particular among these are the two methods usually used in introductory textbooks [1, 10].

Langford (Definition 2 of [7]) uses the following definition of data percentiles.

Definition of data percentile. A i-th percentile value is a number which puts at least i percent of the data values at that number or below and at least $(100 - i)$ percent of the data values at that number or above. If more than one such number exists, there will be an entire interval of such and we choose the i-th percentile value to be the midpoint of that interval. We propose to call this value the **canonical value** of the quartile.

The 25-th, 50-th, and 75-th percentiles of the data are the first quartile, the median, and the third quartile of the data, and these can be used as estimates of the quartiles and median of the (continuous) distribution. In some cases however, there may be reasons for some other (slightly different) choice of data percentiles [4, 6].

Examples: By definition, the quartiles in some simple examples are given below (of size $n = 9, 10, 11, 12, 13$, some with repetitions). All the data are ordered with all repetitions listed as separate observations.

1. $n = 9 :$ 12, 15, $\boxed{22}$, 23, $\boxed{25}$, 44, $\boxed{46}$, 51, 59

   $Q_1 = 22$, $Q_2 = M = 25$, $Q_3 = 46$.

2. $n = 10 :$ 12, 15, $\boxed{22}$, 23, $\boxed{25, 26}$, 44, $\boxed{46}$, 51, 59

   $Q_1 = 22$, $Q_2 = M = 25.5$, $Q_3 = 46$.

3. $n = 11 : 12, 15, \boxed{22}, 22, 23, \boxed{23}, 25, 44, \boxed{46}, 51, 59$

   $Q_1 = 22, Q_2 = M = 23, Q_3 = 46.$

4. $n = 12 : 12, 15, \boxed{\textbf{22, 22}}, 22, \boxed{\textbf{23, 23}}, 25, \boxed{\textbf{44, 46}}, 51, 59$

   $Q_1 = 22, Q_2 = M = 23, Q_3 = 45.$

5. $n = 13 : 12, 15, 15, \boxed{\textbf{22}}, 22, 22, \boxed{23}, 23, 25, \boxed{44}, 46, 51, 59$

   $Q_1 = 22, Q_2 = M = 23, Q_3 = 44.$

3 DISCRETE DATA

The two methods found in the previously mentioned textbooks [1, 10] are among the most popular methods listed in [7]. Both methods first compute the data median, and then compute a median of the two *halves*. In the first method (M1) the median is included in both halves if the number of data points in the entire set is odd and excluded if the number of data points is even (Tukey, [11]). In the second method (M2) the median of the entire set is excluded from both halves (a method described by Moore and McCabe, see [7]).

Langford's method is given as an alternative to a more complicated method, (Method 4 in [7], Definition 2 of [4] or Method 4 in [6]), which uses extrapolation between data values in the cases $n = 4k + 1$ and $n = 4k + 3$. Langford suggests to "divide the data set into two halves, a bottom half and a top half. If n is odd, include or exclude the median in the halves so that each half has an *odd* number of elements. The first and the third quartiles are then the medians of the bottom and top halves respectively." In the n even case the median is taken to be the average of the middle two values.

Langford's method can be summarized by the following four rules (write $n = 4q + r$, and assume the dataset is ordered $v_1 \leq v_2 \leq v_3 \cdots \leq v_n$):

if $r = 3$, then $Q_1 = v_{q+1}$, $Q_2 = v_{2q+2}$ and $Q_3 = v_{3q+3}$.
if $r = 2$, then $Q_1 = v_{q+1}$, $Q_2 = \frac{v_{2q+1}+v_{2q+2}}{2}$ and $Q_3 = v_{3q+2}$.
if $r = 1$, then $Q_1 = v_{q+1}$, $Q_2 = v_{2q+1}$ and $Q_3 = v_{3q+1}$.
if $r = 0$, then $Q_1 = \frac{v_q+v_{q+1}}{2}$, $Q_2 = \frac{v_{2q}+v_{2q+1}}{2}$ and $Q_3 = \frac{v_{3q}+v_{3q+1}}{2}$.

The proof is straightforward (see, c.f. [12]), and the following figure may be helpful:
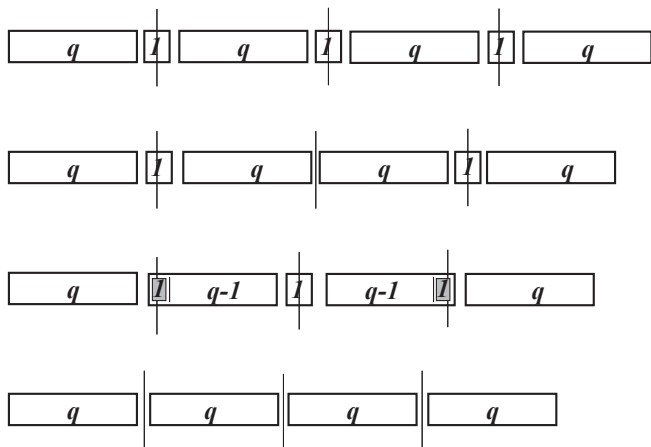
In the next table, first and third quartiles are computed first by definition and then by the three methods M1, M2, and M3. The values not in accordance with our definition above are in bold.

	Q_1				Q_3			
	Def	M1	M2	M3	Def	M1	M2	M3
Example 1	22	22	**18.5**	22	46	46	**48.5**	46
Example 2	22	22	22	22	46	46	46	46
Example 3	22	22	22	22	46	**45**	46	46
Example 4	22	22	22	22	45	45	45	45
Example 5	22	22	**18.5**	22	44	44	**45**	44

We can make several observations from the table. It is easy to see that M1 gives some different results (according to our definition) for any $n = 4k + 3$, and M2 for any $n = 4k + 1$. M3 provides results, consistent to our definition for all cases, see [7].

Note that from a statistical viewpoint all three methods are valid and can be justified. For students, it is easy to explain what theoretical or population quartiles are for continuous (measurement) variables, as the only concept needed is that of chance. Data quartiles are estimates of the theoretical and population values (involve errors of estimation), and slightly different definitions of these are slightly different ways of estimating the theoretical or population values.

While we stress the importance of an unambiguous definition of quartiles at elementary level, it is on the other hand well known that on larger datasets the differences among the various methods for computation of quartiles tend to be very small. Consider the following two examples with datasets of size 1001 and 999.

Examples:

6. Randomly generated dataset of 1001 ($n = 4k + 1$) numbers, normally distributed (with mean 120 and standard deviation 30, rounded to one decimal place) gives the following quartiles, comparing methods M1, M2 and M3[1] :

   M1: $Q_1 = 98.7$, $Q_2 = M = 119.8$, $Q_3 = 137.9$.

   M2: $Q_1 = 98.7$, $Q_2 = M = 119.8$, $Q_3 = \mathbf{138}.05$.

   M3: $Q_1 = 98.7$, $Q_2 = M = 119.8$, $Q_3 = 137.9$.

7. With the same parameters, randomly generated dataset of 999 ($n = 4k + 3$) numbers, normally distributed, gives the following quartiles:

   M1: $Q_1 = 98.3$, $Q_2 = M = 119.1$, $Q_3 = \mathbf{138}.45$.

   M2: $Q_1 = 98.3$, $Q_2 = M = 119.1$, $Q_3 = 138.2$.

   M3: $Q_1 = 98.3$, $Q_2 = M = 119.1$, $Q_3 = 138.2$.

There are some expected (negligible) differences, based on previous considerations on smaller datasets; M1 gives some different results for $n = 4k + 3$, and M2 for $n = 4k + 1$. Here, the differences are already very small, and it is expected that for larger datasets the errors tend to be even smaller.

However, observe that in the last example the samples are of a normally distributed random variable. While the example may well serve to reduce the confusion among students caused by introduction of different definitions, we have to be aware that the more interesting phenomena may occur when considering other distributions. In particular, special care should be taken when considering discrete distributions.

---

[1]All statistical computations were done with R ver. 3.1.0 (2014), The R Foundation for Statistical Computing.

4   SAMPLE QUARTILES OF DISCRETE DISTRIBUTIONS

In case of absolutely continuous distribution it is well known that the asymptotic distribution of sample quantiles in the classical definition is normal (see [8], Theorem 1 and further references therein). If the underlying distribution is discrete, the situation is much more delicate ([5]). Sample quantiles may not even be consistent in general for the population quantiles in this case. Now consider an example of discrete distribution.

Example: simulation throwing (fair) dodecahedron 1000-times gives some interesting results (in only four attempts):

1. Frequencies (throwing 1, 2, ..., 12, respectively) are $69, 84, 86, 69, 92, 88, 71, 77, 86, 104,$ $87, 87$. $Q_1 = 4, M = 7, Q_3 = 10$.

2. Frequencies are $72, 76, 89, 81, 83, 84, 75, 88, 76, 91, 89, 96$.
   $Q_1 = 4, M = 7, Q_3 = 10$.

3. Frequencies are $76, 76, 85, 92, 102, 77, 98, 87, 75, 83, 76, 73$.
   $Q_1 = 4, M = 6, Q_3 = 9$.

4. Frequencies are $74, 82, 95, 82, 81, 90, 95, 83, 84, 74, 72, 88$.
   $Q_1 = 3, M = 6, Q_3 = 9$.

This issue occurs due to the fact that the cumulative distribution function (cdf) in this case is a step function. This leads to inconsistency if the level of the quartile lies in the image of the cdf and, consequently, the Central limit theorem do not hold anymore. In other words, the first quartiles of samples are most likely either 3, 4 or 3.5, regardless of the sample size. We may however anticipate that the expected value of the sample quartiles will converge, hopefully to the quartile values of the population.

Now suppose a dodecahedron is thrown independently n times and we observe a sequence $X_1, \ldots, X_n$ of scores. (Assume that the faces of dodecahedron are labelled with numbers $1, 2, \ldots, 12$.) Clearly, we can assume that $P(X_i = k) = \frac{1}{12}$ for $i = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, 12$. The random variables X_i are thus distributed uniformly over the set of twelve values. The population quartiles are thus $Q_1 = 3.5$, $Q_2 = M = 6.5$ and $Q_3 = 9.5$.

Consider the sample sizes $n = 4q + 3$. Let the sample quartiles as defined and computed by the Langford's method be $Q_1^L = v_{q+1}$, $Q_2^L = M = v_{2q+2}$, and $Q_1^L = v_{3q+3}$.

On the other hand, denote the sample quartiles, as computed by the method M1 (put the median in both halves and set quartiles to be the median of the halves), with $Q_1^T = \frac{v_{q+1} + v_{q+2}}{2}$, $Q_2^T = M = v_{2q+2}$, and $Q_3^T = \frac{v_{3q+2} + v_{3q+3}}{2}$.

We conclude discussion with several observations, and questions or conjectures that are of interest for future work. Let the ordered dataset $v_1 \le v_2 \le \ldots \le v_n$ represents a sample from some population with underlying discrete distribution, sample quartiles Q_i^L and Q_i^T defined as above and $E(\cdot)$ the expected value of random variable. Then

- It can be shown that at fixed sample size $n = 4q + 3$, the expected values of quartiles due to different methods are different $E(Q_1^T) < E(Q_1^L)$.

- A question arises, whether the difference $E(Q_1^L) - E(Q_1^T)$ vanishes as $n \to \infty$.

- We conjecture that there is a constant $\Delta > 0$ such that

$$\lim_{n\to\infty} E(Q_1^L) - \lim_{n\to\infty} E(Q_1^T) > \Delta.$$

References

[1] Bon Klajnšček, M., Dvoržak, B., Felda, D. (2009). MATEMATIKA 1, učbenik za gimnazije, DZS, Ljubljana.

[2] Dr. Twe (2000). Reply to Tom about quartiles. Online at
http://mathforum.org/library/drmath/view/60969.html
(accessed 24.12.2016).

[3] Franklin, C. and Kader, G. (2010). Models of Teacher Preparation Designed Around the GAISE Framework. Proceedings of the Eighth International Conference on Teaching Statistics.
http://www.stat.auckland.ac.nz/~iase/publications/icots8/
ICOTS8_3E3_FRANKLIN.pdf (accessed 15.5.2017).

[4] Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician* **50**(4), 361 – 365.

[5] Jentsch, C. and Leucht, A. (2014). Bootstrapping sample quantiles of discrete data. https://ub-madoc.bib.uni-mannheim.de/36588/1/Jentsch und Leucht 14-15.pdf (accessed 23.6.2017).

[6] Joarder, A. H. and Firozzaman, M. (2001). Quartiles for Discrete Data. *Teaching Statistics* **23**(3), 86 – 89.

[7] Langford, E. (2006). Quartiles in Elementary Statistics. *Journal of Statistics Education* **14**(3), 16 pp.
https://ww2.amstat.org/publications/jse/v14n3/langford.html

[8] Ma, Y., Genton, M. G. and Parzen, E. (2011), Asymptotic properties of sample quantiles of discrete distributions. *Ann Inst Stat Math* **63**, 227 – 243. DOI 10.1007/s10463-008-0215-z

[9] Magajna, Z. in Žakelj, A. (1999). Ali sodi obdelava podatkov k pouku matematike? *Obzornik mat. fiz.* **46**, 113 – 119.

[10] Pavlič, G., Kavka, D., Rugelj, M., Šparovec, J. (2011). LINEA NOVA, matematika za gimnazije, Modrijan, Ljubljana.

[11] Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley.

[12] Žerovnik, J. (2017). Računanje kvartilov v elementarni statistiki, *Obzornik mat. fiz.* **64**, 20 – 31.

[13] Žerovnik, J. and Rupnik Poklukar, D. (2017). Elementary methods for computation of quartiles, *Teaching Statistics* **39** (3), 88 – 91.

CLUSTER ANALYSIS OF THE POST-TRANSITION COUNTRIES OF EUROPEAN UNION ACCORDING TO THE INCOME INEQUALITY AND SOCIAL SPENDING

Nika Šimurina
Department of Finance, Faculty of Economic and Business, University of Zagreb
Trg J.F. Kennedyja 6, 10000 Zagreb, Croatia
E-mail: nsimurina@efzg.hr

Nataša Kurnoga
Department of Statistics, Faculty of Economic and Business, University of Zagreb
Trg J.F. Kennedyja 6, 10000 Zagreb, Croatia
E-mail: nkurnoga@efzg.hr

Blaženka Knežević
Department of Trade, Faculty of Economic and Business, University of Zagreb
Trg J.F. Kennedyja 6, 10000 Zagreb, Croatia
E-mail: bknezevic@efzg.hr

Abstract: The aim of this paper is to classify the post-transition countries of EU according to differences in their income inequality and social spending. As an instrument of fiscal policy, social spending should play a significant role in reducing income inequality. We also analyze important questions regarding income inequality, such as the unemployment and internet usage data. We perform cluster analysis to group the eleven Central and Eastern European (CEE) post-transition countries of EU. Comparison of the results of cluster analyses carried out, first for three and then for four variables, reveals a similar cluster structure.

Keywords: cluster analysis, income inequality, social spending, internet use, poverty, post-transition countries

1 INTRODUCTION

In this paper we discuss the effects of income inequality and social spending in the post-transition countries of EU. Questions regarding income inequality, such as unemployment and internet use, will also be discussed.

Increasing overall employment is the first-best way to reduce inequality, as unemployment and underemployment undermine welfare of individuals. To reduce inequality, workers must be competitive in a quasi-global labor market and must continually invest in their knowledge and skills. "Up-skilling" of the workforce is one of the most powerful instruments at the governments' disposal which can be used to counter rising inequality and increase employment rates.

While inequality in wages and salaries are the most important contributor to income inequality, transfer payments, as well as income taxes and social security contributions play a role in countering it. Low wages contribute to greater demands on the social safety net to maintain living standards. Similarly, fiscal authorities use the redistributive effects of a progressive taxation and social security contribution to distribute income more equally. Limits of the effectiveness of such policies depend on the respond of the individuals and firms to a change in relative prices and other elasticities, such as labor supply.

The remainder of the paper is organized as follows. Section 2 summarizes the extant literature on the determinants of inequality in post-transition countries. Section 3 discusses the data used and summarizes the empirical strategy. Section 4 discusses the results. Section 5 gives some concluding recommendations.

2 LITERATURE REVIEW

According to Stiglitz [13] the new financial crisis has made inequalities worse in innumerable ways, beyond higher unemployment, lost homes, and stagnating wages. The crisis had a wide range of impacts throughout the European Union. In response, the EU member states have adopted a wide range of policies.

Income inequality is relatively easy to define. It is frequently used as an indicator of relative poverty or prosperity. Inequality is often invoked as an argument for income redistribution policies. The difficulty, however, lies in understanding the origins of income inequality. Some previous studies have emphasized family structure, technology, globalized markets, immigration, property rights or trade. Other studies have focused on regulatory reforms and institutional changes and showed that the effects achieved with these measures are contradictory. Although such regulations and reforms increase employment possibilities, they can also contribute to wider wage disparities.

One of the leading authorities on the topic of income inequality, Milanovic [8] argues that the observed increase in inequality in transition countries is driven mainly by higher inequality in wage distribution. Keane and Prasad [5] also find that the reallocation of workers from the public sector (with a compressed wage distribution) to the private sector (with much higher wage inequality), accounts for the bulk of increased income inequality during transition. They also highlight the role that increased social transfers have on limiting increases in inequality.

According to OECD [10] studies the redistribution achieved by public cash transfers is on average twice as large as that achieved through taxes. The recent financial and fiscal crises have resulted in a contraction of public spending (including transfers) and in increased taxes, particularly on top income brackets, among most post-transition countries. Whereas taxes and transfers have significant redistributive impact and consequently affect income inequality, the redistributive impact of taxes depends on their size, mix, and their progressivity. Among most popular forms of social spending in CEE economies are pensions, family benefits, disability benefits and unemployment benefits.

Karaman Aksentijević et al. [4] identify education as the most influential tool of economic and social policy that reduces poverty and economic inequality and Obadić et al. [9] argues that only highly skilled workers benefit from a more dynamic economy.

Kurnoga Zivadinovic at el. [6] classified the selected European countries regarding the structural economic indicators, such as: GDP per capita, total employment rate, comparative price levels, employment rate of older workers, long term unemployment and productivity of national economies. Cluster analysis was used to classify selected countries and Croatia was grouped with the countries that have similar political and historical background (Central and Eastern European countries).

Numerous studies emphasize that development indicators of internet and telecommunications can help us to understand modern aspects of economic development and economic inequalities. Saunders et al. [12] explain the role of telecommunications in economic development and give a comprehensive analysis of internet contribution to macroeconomic development of various countries. Graham [1] argues that digital divide as a consequence an absence of modern ICTs creates not only technical divide between countries, markets and economies, but also social and democratic divide making less developed economies lagging behind economies which invest in ICT and, thus, actively participate and benefit from "internet revolution" and "internet society". Supporting findings and descriptions of existing correlation between internet, economic inequalities and economic development are given by Lucas and Sylla [7], Pradhan et al. [11] and Jimenez et al. [2].

3 DATA SET AND METHODOLOGY

This paper seeks to classify the post-transition countries of EU according to their income inequality and social spending policy. This instrument of fiscal policy should play a significant role in reducing income inequality. Furthermore, the paper also addresses important questions regarding income inequality, such as unemployment and internet use.

The observed EU post-transition countries are: Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia. The analyzed variables and their abbreviations in brackets are as follows: Gini coefficient of equalized disposable income – EU-SILC survey (*gini*), long-term unemployment rate – percentage of active population aged 15-74 (*unempl*), total expenditure on social protection by type – percentage of total expenditure (*socprot*) and internet use – percentage of individuals (*internet*). The data source was Eurostat. The last available data for selected countries were for 2014.

To group the eleven Central and Eastern European (CEE) economies we used cluster analysis. Cluster analysis groups similar objects into homogenous groups/clusters. That is, the objects in one cluster are similar to each other and dissimilar to objects in other clusters. There are various methods of hierarchical and non-hierarchical cluster analyses. The most common hierarchical cluster analysis is Ward's method. The most common non-hierarchical method is the k-means method. For grouping similar objects in clusters, we must use similarity measures. That is, distance measures because the objects with smaller distances between them are more similar to each other than are those with larger distances.

First, we performed hierarchical and non-hierarchical cluster analyses on only three variables (poverty, unemployment and expenditures on social protection) with impact on income inequality, to identify homogenous groups among the post-transition countries. We use the Gini coefficient as the measure of income inequality, which is defined as the relationship of cumulative shares of the population arranged according to the level of equalized disposable income, to the cumulative share of the equalized total disposable income received by them. We want to group the eleven Central and Eastern European (CEE) economies only according to those variables, because they should have strong synergy and influence on income inequality. Hierarchical cluster analysis and dendrogram reported the cluster solution. Non-hierarchical cluster analysis then confirmed the results of chosen cluster solution.

Since, unemployment affects mainly low-skilled workers, who are then more reliant on state assistance, we introduced fourth variable of internet use, to explore the effect of "up-skilling" of the workforce on the cluster structure of chosen countries. Unemployment or underemployment limits people's ability to access decent income and cuts people off from social networks. To escape unemployment in a globalized world, people have to be competitive on labor market. For that reason, they must continually invest in their knowledge and skills.

We then performed hierarchical and non-hierarchical cluster analyses on the aforementioned variables as well as the additional variable of internet use. Finally we compared the results of cluster analyses conducted separately for three and four variables.

4 DISCUSSION OF RESULTS

Before cluster analysis is conducted, the presence of multicollinearity should be revealed. The numerical reliability of results could be questionable in the presence of multicollinearity. One indicator that multicollinearity may be present is the variance inflation factor (VIF). VIF quantifies how much the variance of the estimated regression coefficient is inflated due to

collinearity with the other regressors. VIF values smaller than 5 indicate the absence of multicollinearity.

We examined the multicollinearity for the analyzed variables. To check the multicollinearity we converted variable into a dependent variable and regressed it against the remaining independent variables. All VIF values were smaller than 5. There was no high multicollinearity.

The cluster analysis is quite sensitive to measurement differences among the variables because of the distance measures. Therefore, the clustering variables were standardized and standardized values were then used in cluster analyses.

After the assumptions were examined and found to be satisfied, we conducted hierarchical and non-hierarchical cluster analyses on three variables (poverty, unemployment and expenditures on social protection) with impact on income inequality. Figure 1 gives the hierarchical and non-hierarchical cluster analysis results, respectively dendrogram and plot of means for each cluster. First, the hierarchical cluster analysis was used. Ward's method with Squared Euclidean distances was performed and according to the dendrogram given in Figure 1 the four-cluster solution was chosen: the first cluster (Poland, Hungary, Slovenia and Czech Republic), the second cluster (Estonia, Latvia and Romania), the third cluster (Croatia and Slovakia) and the fourth cluster (Bulgaria and Lithuania). The non-hierarchical cluster analysis then confirmed this chosen solution. We used k-means method for the four-cluster solution and obtained the same cluster solution.

According to the obtained plot of means shown in Figure 1, in the first cluster the variable of expenditures on social protection has a small range of values regarding the fact that all CEE economies are using a "Continental European model" by Joumard, Pisu and Bloch [3] characterized by large cash transfers with the lion's share of old-age pension (redistributing income mostly over the lifecycle instead of across individuals). The unemployment variable shows that two countries (Croatia and Slovakia) are outliers with a high unemployment rate, while in other countries the unemployment rate is not too high (below 7%). According to the Gini coefficient, differences in income inequality among CEE countries are not big, consistent with their similar historical background as transition economies.
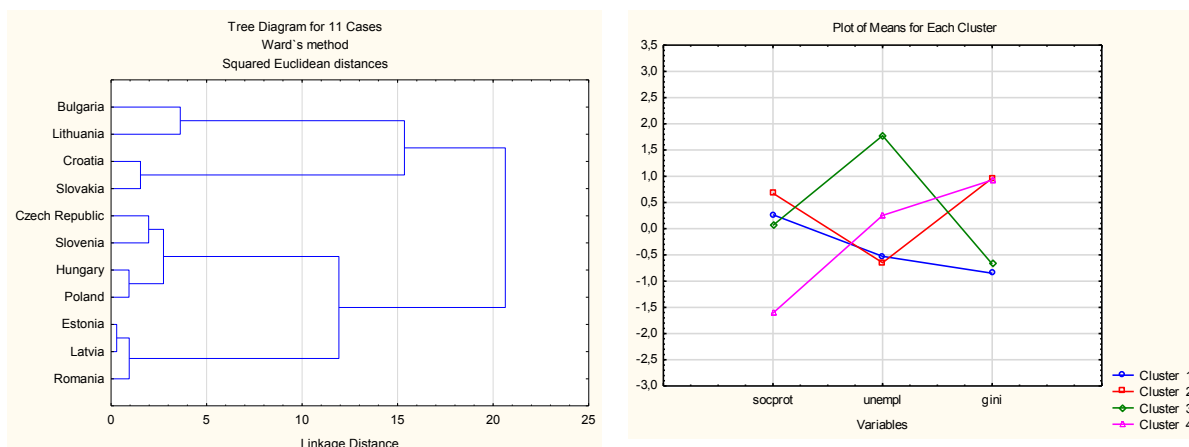
Figure 1: Dendrogram for three variables and plot of means for each cluster

After we conducted cluster analysis on three variables, we introduced an additional variable on internet use to the analysis.

Again, we conducted hierarchical and non-hierarchical cluster analyses. The results for both analyses are shown in Figure 2, respectively dendrogram given by the Ward's method with Squared Euclidean distances and plot of means for each cluster given by the k-means

method. According to the dendrogram shown in Figure 2 the four-cluster solution has the following cluster composition: the first cluster (Estonia and Latvia), the second cluster (Hungary, Slovenia and Czech Republic), the third cluster (Croatia and Slovakia) and the fourth cluster (Poland, Romania, Bulgaria and Lithuania).

The difference from the cluster solution on three variables is in the clustering of Poland and Romania. In this cluster solution, Poland is, due to the low percentage of internet use (69%), in the cluster with Romania (59%), Bulgaria (59%) and Lithuania (73%) compared to the previous solution where it was with Hungary (77%), Slovenia (74%) and Czech Republic (81%). Furthermore, also considering the low percentage of internet use, Romania (59%) is now with Bulgaria (59%), Poland (69%) and Lithuania (73%) but previously was in the cluster with Estonia (86%) and Latvia (77%). It can be seen also, that Lithuania has greater percentage of internet use (73%), than Poland (69%), Romania (59%) and Bulgaria (59%). In a five-cluster solution, Lithuania is alone in the fifth cluster.

After the hierarchical cluster analysis was carried out, we then used the non-hierarchical cluster analysis to confirm the four-cluster solution. The k-means method was carried out for the four-cluster solution and resulted with the following cluster composition: the first cluster (Bulgaria, Romania and Lithuania), the second cluster (Hungary, Slovenia, Poland and Czech Republic), the third cluster (Estonia and Latvia) and the fourth cluster (Croatia and Slovakia). The only difference compared to the results of hierarchical cluster analysis is in the clustering of Poland. In the solution of hierarchical cluster analysis, Poland is in the cluster with Romania, Bulgaria and Lithuania, and now is in the cluster with Hungary, Slovenia and Czech Republic.
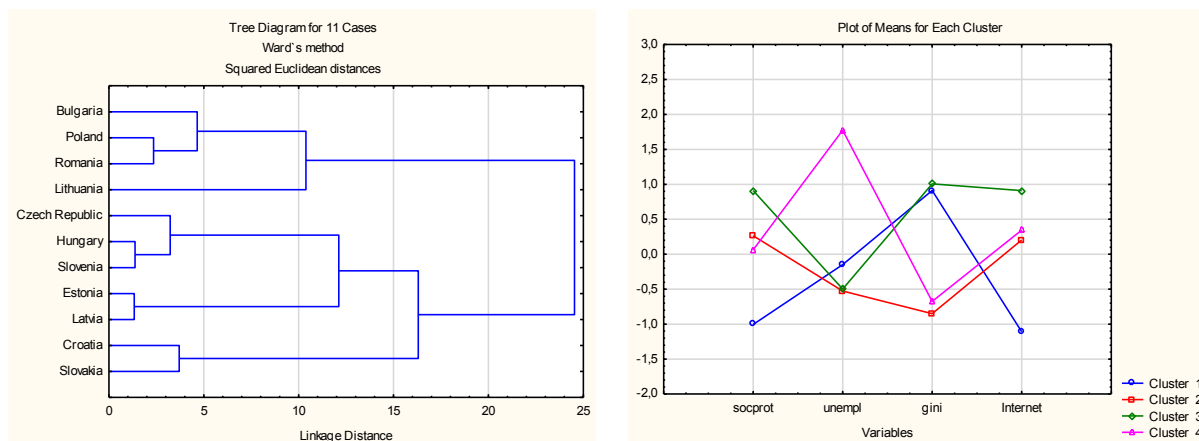
Figure 2: Dendrogram for four variables and plot of means for each cluster

By comparing the results of non-hierarchical cluster analyses, first for three and then for four variables, we reveal a similar cluster structure. The only difference is in the clustering of Romania. Romania is in the cluster analysis on three variables clustered with Estonia and Latvia, and in the cluster analysis on four variables with Bulgaria and Lithuania. This is due to the introduction of the variable on internet use. Namely, Romania has a lower percentage of internet usage as Bulgaria and Lithuania.

5 CONCLUSION

The aim of this paper is to classify the post-transition countries of EU according to differences in their income inequality and social spending. Furthermore, important questions regarding income inequality, such as unemployment and internet use are analyzed.

The focus of the research was how to classify all the eleven CEE economies. Cluster analysis supports the notion that all post-transition countries experience similar problems regarding income inequality but each country has its own experience with different fiscal instruments (expenditures) and theirs uses different options to achieve distributive objectives in an efficient manner.

It also must be stressed that due to technological progress (internet use) better-educated individuals have better possibilities for finding and keeping jobs and they are also paid better. More unequal wages have contributed to the fact that more people need the help of the social safety net in order to maintain their living standards.

Acknowledgement

References

[1] Graham, M. (2008). Warped Geographies of Development: The Internet and Theories of Economic Development. Geography Compass, 2 (3): 771-789.

[2] Jiménez, M., Matus, J.A., Martínez, M.A. (2014). Economic growth as a function of human capital, internet and work. Applied Economics, 46 (26): 3202-3210.

[3] Joumard, I., Pisu, M., Bloch, D. (2012). Tackling income inequality: The role of taxes and transfers, OECD Journal: Economic Studies, http://dx.doi.org/10.1787/eco_studies-2012-5k95xd6l65lt [Accessed 13/04/2017].

[4] Karaman Aksentijević, N., Denona Bogović, N., Ježić, Z. (2006). Education, poverty and income inequality in the Republic of Croatia, Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu/Proceedings of Rijeka Faculty of Economics: Journal of Economics and Busniness, 24 (1): 19-37.

[5] Keane, M., P., Prasad, E., S. (2002). Inequality, Transfers, And Growth: New Evidence From The Economic Transition In Poland. The Review of Economics and Statistics, 84(2): 324-341.

[6] Kurnoga Zivadinovic, N., Dumicic, K., Ceh Casni, A. (2009). Cluster and Factor Analysis of Structural Economic Indicators for Selected European Countries, Wseas Transactions on Business and Economics, 6 (7), 331-341.

[7] Lucas, H., Sylla, R. (2010). The Global Impact of the Internet: Widening the Economic Gap Between Wealthy and Poor Nations?, Prometheus - Critical Studies in Innovation, 21(1): 1-22.

[8] Milanovic, B. (1999). Explainig the Increase in Inequality during transition. Economics of Transition, 7(2): 299-341.

[9] Obadić, A., Šimurina, N., Sonora, R. (2014). The effects of tax policy and labour market institutions, Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu/Proceedings of Rijeka Faculty of Economics: Journal of Economics and Busniness, 32(1): 121-140.

[10] OECD, (2008). Growing Unequal? Income Distribution and Poverty in OECD Countries. Paris: OECD Publishing.

[11] Pradhan, R.P., Arvin, M.B., Norman, N.R., Bennett, S.E. (2015). Financial depth, internet penetration rates and economic growth: country-panel evidence. Applied Economics, 48(4): 331-343.

[12] Saunders, R.J., Warford, J.J., Wellenius (1994), Telecomunications and Economic Development. Baltimore and London: Johns Hopkins University Press for the World Bank.

[13] Stiglitz, J. (2012). The Price of Inequality – How Today′s Divided Society Endangers our Future. New York, London: W. W. Norton & Company.

NONLINEAR CONNECTIONS IN STRUCTURAL EQUATION MODELING: THE CASE OF SERVICE SETOR COMPANIES IN SLOVENIA

Živa Veingerl Čič, Simona Šarotar Žižek, Vesna Čančer
University of Maribor, Faculty of Economics and Business
Razlagova 14, 2000 Maribor, Slovenia
E-mails: zivana.veingerl1@um.si; simona.sarotar-zizek@um.si; vesna.cancer@um.si

Abstract: The paper argues the necessity of considering nonlinear connections in structural equations modelling. By using WarpPLS 5.0, it presents and discusses the non-linear connections between the individual constructs of the conceptual model of individual performance management system of the employees in the survey among service sector companies in Slovenia. The results show that the use of comprehensive development methods have significant positive impact on employee satisfaction at work, individual performance and psychical well-being.

Keywords: WARP PLS, structural equation modelling, nonlinear connections, service sector

1 INTRODUCTION

The vast majority of the models developed by structural equation modeling (SEM) are based on the linear connections between constructs. In case that there are nonlinear connections between constructs, the failure of the tools based on the linearity to comply with nonlinearity can result in unacceptable low quality of the obtained models regarding model consistency, convergent, discriminant and nomological validity, and statistically insignificant connections between the constructs in the obtained conceptual models. The aim of this paper is to present and discuss the non-linear connections between the individual constructs of the conceptual model using the tools Warp PLS 5.0, together with the quality of the model based on nonlinear connections in the case of the individual performance management system in the survey among employees in service companies in Slovenia. In this study, we focused on the individual employee performance management through the use of comprehensive development methods and the impact on psychic well-being, employee satisfaction at the workplace and individual performance. Individual constructs of the conceptual model are comprehensive development methods such as coaching, mentoring, intergenerational cooperation and sponsorship, employee satisfaction at the workplace, psychic well-being, and individual employee performance. Conceptual model was verified using SEM. For this purpose, we conducted a survey among service sector companies in Slovenia in 2016: 418 respondents from 334 companies completed a questionnaire about Individual Performance. The data were collected from 1th March to 31 of May 2016. The collected data were processed using the programs: IBM SPSS 23.0 and WarpPLS 5.0.

2 NONLINEARITY IN STRUCTURAL EQUATION MODELING

Researchers in quantitative research in the social sciences use models as mental, physical or formal structures [13]. Considering the fact that our research falls within the scope of social sciences and business sciences, we applied SEM, because it is a relatively simple tool for verifying the connections [1]. SEM is in practice most often performed with software tools LISREL, SmartPLS, WarpPLS and AMOS. Within the framework of our research we used software tool WarpPLS as it identifies the non-linear relationship between the studied latent variables. When the relationships between the studied latent variables are non-linear., as stressed by Šebjan [14], the coefficients obtained with WarpPLS are often higher than

coefficients presented by other software tools. Algorithm WarpPLS uses PLS-regression which reduces multicollinearity between the latent variables [14].

Taking nonlinearity into consideration sometimes leads to results that are markedly different from the corresponding linear results, particularly when the underlying relationships take the form of U-curves [8]. We decided to use the software Warp PLS 5.0 in our study, because unlike other programs such as LISREL or Smart PLS, Warp PLS 5.0 identifies non-linear connection between the latent variables [9]. The linear algorithm does not perform any warping of relationships. The Warp2 algorithm tries to identify U-curve relationships among linked latent variables, and, if those relationships exist, the algorithm transforms (or "warps") the scores of the predictor latent variables so as to better reflect the U-curve relationships in the estimated path coefficients in the model [9]. This option allows users to customize their analyses based on theory and past empirical research. If theory or results from past empirical research suggest that a specific link between two latent variables is linear, then the corresponding path can be set to be analyzed using the linear algorithm. Conversely, if theory or results from past empirical research suggest that a specific link between two latent variables should have the shape of a U curve (or J curve), the corresponding path can be set to be analyzed using the Warp2 algorithm or the Warp2 Basic algorithm [9].

The fact is that we cannot deny that the most connections between variables are non-linear (although it is often assumed that there are linear connection). To verify this assumption, we decided to use the software Warp PLS 5.0 to check whether the link between the latent variables in our model is linear or non-linear. The Warp PLS 5.0 was applied in order to exploit its ability to establish a non-linear relationship between the variables (if any). At the same time, the advantage of Warp PLS is also that the values of the individual coefficients are calculated automatically and not manually [9]. Kock [10] lists the following advantages of Warp PLS compared to other tools:
- automatically evaluates the value of the statistical significance (p) of the coefficients of connections and thus determining the appropriateness of statistically significant correlations,
- automatically builds the structure of indicators,
- allows the user to display warped relationships charts for each connection between the variables in the regression curve (the term "warped" is used for relationships that are clearly nonlinear [9]),
- provides the calculation and display of coefficient for verifying multicolinearity.

WarpPLS works well with larger samples (well over 500) [2]. WarpPLS can also be used in cases with large number of indicators [4]. Nonlinear analyses employing the software WarpPLS also allow for the identification of linear segments emerging from a nonlinear analysis, but without the need to generate subsamples.

3 RESEARCH ANALYSIS RESULTS AND DISCUSSION

First we tried to build a model for the study of individual performance management of employees based on the linear connections. For this purpose, we used the Smart PLS tool where the quality assessment of key indicators of conceptual model were lower (not statistically significant, $p > 0.05$) than those obtained with WARP PLS. Because the indicators show that there are no significant connections between the constructs, we used Warp PLS to examine the nonlinear connections between the constructs. In the context of the SEM, the model consistency was verified with the index Goodness-of-fit (GoF), keeping in mind the criteria of 0.1, 0.25 and 0.36.

Linear relationships between the pairs of latent variables, that is, those relationships best described by a line, are relatively easy to interpret [9]. They suggest that an increase in one

variable either leads to an increase (if the slope of the line is positive) or decrease (if the slope is negative) in the other variable. Figure 1 presents that the relationship between the basic constructs in the conceptual model are nonlinear.

Nonlinear relations between the latent variables could provide a much more nuanced view of the data and at the same time they are also much more difficult to explain [7]. The data in Figure 1 can be explained so that we can divide each graph into two parts according to the average value of the standard deviation for each latent variable separately. Kock [9] states that each section of the curve, represented in each picture, is treated separately.
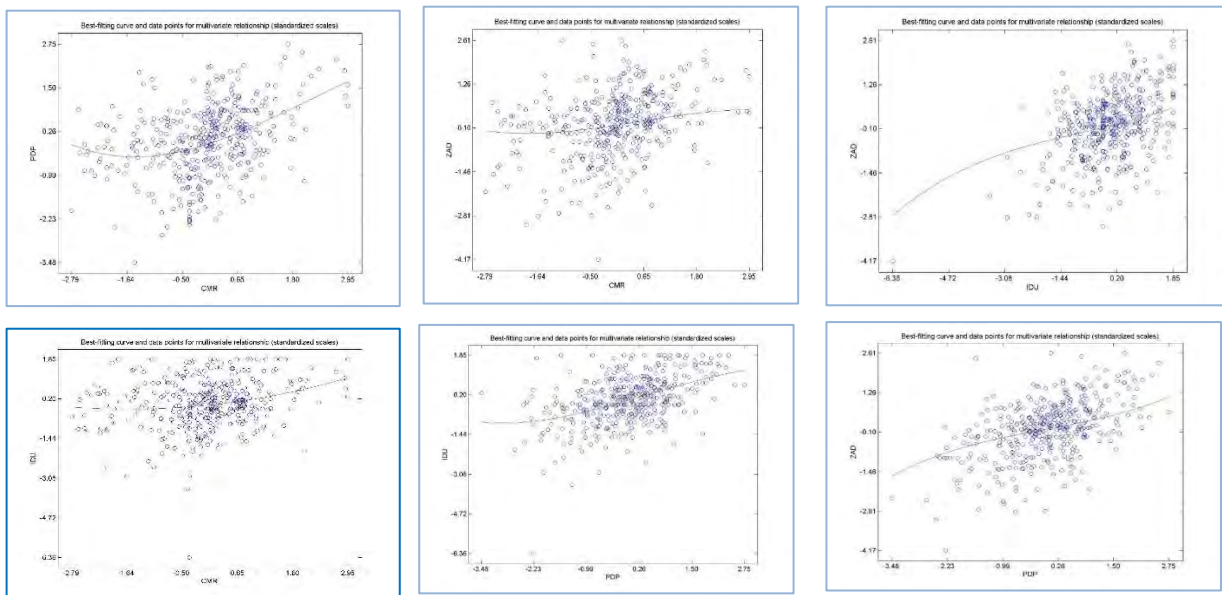
Figure 1: Forms of connections between the constructs of basic conceptual model

The research therefore applied WarpPLS in the following steps: (1) determining the structural model, (2) determining the measurement models, (3) data collection and review, (4) evaluating the PLS model, (5) evaluating the results of the PLS SEM, (6) evaluating the results of the PLS SEM model structure, (7) advanced PLS analysis, and (8) the interpretation of the results and conclusions. Below is the summary of the results obtained in this survey. In this article we will focus further to present the results of steps six to eight.

The key indicators for assessing the quality of the conceptual model by WarpPLS are presented hereinafter [9]. For a final research model is important to determine the degree of data fit to research model. The model is validated when they are confirmed, for all three models: measuring, general structural and conceptual model. For the realization of this objective within the framework of PLS modeling we usually verify communalities index, the Redundancy Index and Gof [11], but in the context of software support WarpPLS we could check only index Gof (0,390). In the PLS model, the composite reliability coefficient (CR), where CR > 0.60, and average variance extracted index (AVE), where AVE > 0.50, are also crucial. The measure of a critical evaluation of the structural model presents adjusted R-squared coefficient (R^2), which should be reviewed for each latent variable; in this case it was higher than 0.15. It reflects the percentage of explained variance of latent variables in the structural model [12].
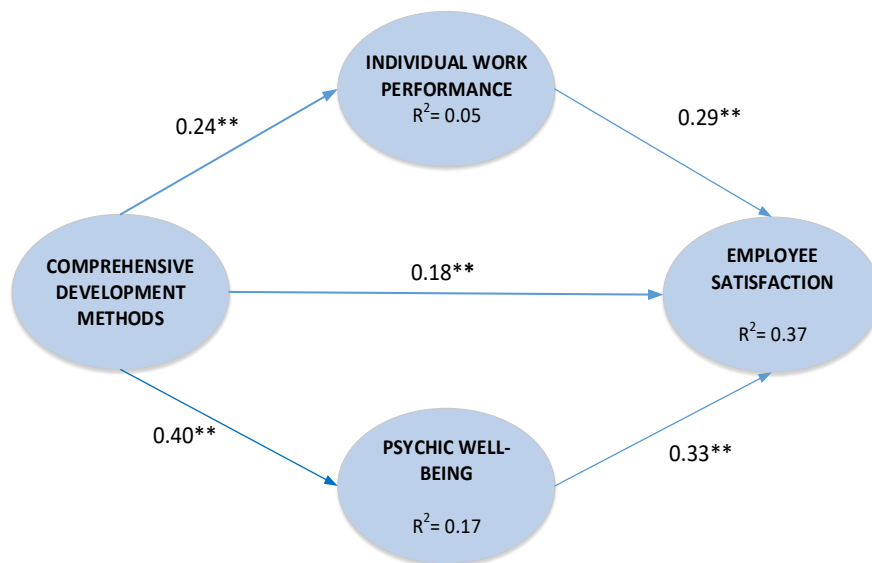
Note: ** p<0.01
Source: Authors

Figure 2: The conceptual model with the results of the links between constructs

Values of Cronbach α are higher than 0.6 except for the construct individual employee performance where Cronbach α is slightly below 0.6: it is 0.584. According to the values of Cronbach α, we met the criterion of the reliability of the measuring instrument. When checking the reliability of the composite we considered that all CR values of all constructs were higher than 0.6. We met this criterion, since all CR values in individual constructs were higher than 0.7. When verifying the convergent validity measured with AVE we found that all the values exceeded the prescribed value 0.5; that is in accordance with the Fornell and Larcker rule [3]. The criterion CR > AVE was also met for all constructs. With the help of the variance inflation factor (VIF) we examined multicollinearity. When the VIF is below 3.3, it can be argued that there is no multicollinearity, other more conservative condition for determining multicolinearity is VIF > 5.0, or VIF > 10.0 [5]. Because the VIF of all constructs is less than 1.6, we met the criterion of low multicollinearity since the value is less than 3.3 or 5.0. Discriminant validity was verified with the highest total variance (MSV) and the average total variance (ASV). When checking the MSV and ASV we took into account the criterion: MSV <AVE and ASV <AVE [6]. These results indicate that we can accept discriminant validity for measuring conceptual model. Nomological validity was evaluated by using the correlation between the constructs with which we verified the consistency of theoretical constructs; we took into account the criteria that the correlation coefficients are statistically significant [14]. The correlation coefficients are for all constructs statistically significant and positive, therefore we accept the nomological validity. By modeling the structural equation, we found that all the connections between the constructs are positive and statistically significant and have rate risk $p < 0.01$. Table 1 shows the values of the indicators by which we evaluated the quality of the advanced conceptual model, and the consistency of the data with conceptual model.

Table 1: Key quality assessment indicators of conceptual model with WarpPLS

Coefficient	Criterion of coefficients	Calculated values of indicators for model
Average path coefficient (APC)	p<0.05	0.279. p<0.001
Average R-squared (ARS)	p<0.05	0.190. p<0.001
Average adjusted R-squared (AARS)	p<0.05	0.187. p<0.001
Average block variance inflation factor (AVIF)	AVIF<=5.0 or AVIF<=3.3	1.279
Average block variance inflation factor (AFVIF)	AFVIF<=5.0 or AVIF<=3.3	1.434
Tenenhaus Gof (Gof)	GoF≥ 0.1 (low) GoF≥ 0.25 (medium) GoF≥ 0.36 (high)	0.339
Sympson's paradox ratio (SPR)	SPR≥0.7 or SPR=1.0	1.000
R-squared contribution ratio (RSCR)	RSCR≥0.9 or RSCR=1.0	1.000
Statistical suppression ration (SSR)	SSR≥0.7	1.000
Nonlinear causality direction ration (NLBCDR)	NLBCDR≥0.7	1.000

The coefficients of the APC, ARS, and SRA are statistically significant at $p < 0.05$, the coefficient AVIF and AFVIF are below the prescribed value 3.3 or 0.5 and are suitable. GoF shows us the power of the underlying conceptual model which is medium powerful (0.339), which is above 0.25, which shows that the model is moderate appropriate. The values of other coefficients: SPR, the SSR and NLBCDR are higher than the prescribed value 0.7; their value is exactly 1.000. The value of the Sympson's paradox ratio coefficient (SPR) is equal to 1, which means that there are no examples of Sympson's paradox in the model.

Links between individual constructs are non-linear. The results show that the comprehensive development methods have positive impact on the psychical well-being ($\gamma_{CMR \rightarrow PDP} = 0.40$, $p < 0.01$). The value of Cohen's coefficient ($f^2 = 0.156$) shows that the effect of predictive latent variables is of medium strength. The comprehensive development methods have positive, statistically significant and weak impact on individual employee performance ($\gamma_{CMR \rightarrow IDU} = 0.24$, $p < 0.01$), as well as the strength of the effect f^2 is greater than 0, 02, ranging from 0,047. The results of SEM and structural coefficients of links of basic conceptual model are presented in Table 2.

Table 2: Structural coefficients of links of basic conceptual model

The link between constructs	Link direction	Shape of link	Direct effect (γ)	f²	Standard error
CMR→PDP	+	nonlinear	0.40**	0.156	0.046
CMR→IDU	+	nonlinear	0.24**	0.047	0.049
CMR→ZAD	+	nonlinear	0.18**	0.064	0.048
IDU→ZAD	+	nonlinear	0.29**	0.135	0.046
PDP→ZAD	+	nonlinear	0.33**	0.175	0.047

Note: **p<0,01; f²-size effect, ZAD – employee satisfaction at work, CMR – comprehensive development methods, IDU – individual employee performance, PDP – psychical well being

4 CONCLUSION

During the history researchers who have conducted multivariate quantitative analyses have been restricted to the use of linear algorithms in their estimation of coefficients of association. This situation has changed recently with the availability of software tools for nonlinear analysis, such as the software WarpPLS [2, 7]. So researchers in other fields where

multivariate statistical methods are employed, can now account for the existence of underlying nonlinearity when estimating coefficients of association among linked variables. The research results show that the quality of the conceptual model obtained with Warp PLS that supports considering nonlinear connections between constructs. As the considered constructs were psychological constructs we can explain the cause of non-linearity in quantitative terms as a result of the current opinion and well-being of individual respondents. Persons can differently perceive presented constructs and therefore the answers are influenced by the environment, the personality of respondents, and their experience and other relevant factors.

References

[1] Bollen, K. A. 1989. Structural Equations with Latent Variables. New York: John Wiley and Sons.

[2] Chen, S.P. 2014. A Verification of Factors Influencing Taiwan College Students' Well-Being Model. Life Science Journal, 11(9): 1013–1023.

[3] Fornell, C., Larcker, D. F. 1981. Evaluating structural equation models with unobservable variables and measurement error, Journal of Marketing Research, 18: 39–50.

[4] Gefen, D., Rigdon, E.E., Straub, D. 2011. An Update and Extension to SEM: Guidelines for Administrative and Social Science Research, MIS Quarterly, 35(29): iii–xiv.

[5] Hair, J. F. Jr., Black, W.C., Babin, B. J., Anderson, R. E. 2005. Multivariate data analysis. Upper Saddle River: Pearson Prentice Hall.

[6] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. 2010. Multivariate Data Analysis. Seventh Edition. New Jersey, Upper Saddle River: Prentice Hall.

[7] Kock, N. 2011. A mathematical analysis of the evolution of human mate choice traits: Implications for evolutionary psychologists, Journal of Evolutionary Psychology, 9(3): 219–247.

[8] Kock, N., Gaskins, L. 2016. Simpson's paradox, moderation, and the emergence of quadratic relationships in path models: An information systems illustration, International Journal of Applied Nonlinear Science, 2 (3): 200–234.

[9] Kock, N. 2015. Warp PLS 5.0 User Manual. Laredo, TX: Script Warp Systems.

[10] Kock, N. 2016. Advantages of nonlinear over segmentation analyses in path models, International Journal of e-Collaboration, 12(4): 1–6.

[11] Milfelner, B., Mumel, D., Snoj, B. 2006. Meta analiza dveh pristopov k raziskovanju kompleksnih marketinških problemov (A meta-analysis of two research approaches to complex marketing problems). Naše gospodarstvo (Our Economy), 52 (5/6): 37–51.

[12] Sharma, P. N., Kim, K. H. 2012. Model selection in information system research using partial least squares based structural equation modeling. In Proceedings of the 33rd International Conference on Information Systems (pp. 1-13), Orlando, FL.

[13] Šarotar Žižek, S. 2012. Vpliv psihičnega dobrega počutja na temelju zadostne in potrebne osebne celovitosti zaposlenega na uspešnost organizacije (The impact of psychical well-being based on the sufficient and necessary personal integrity of an employee on organizational performance). Doctoral dissertation, Maribor: Faculty of Economics and Business, University of Maribor.

[14] Šebjan, U. 2015. Kritični dejavniki uporabe informacijskih rešitev analitičnega upravljanja odnosov s strankami (Critical factors of analytical customer relationship management information solutions use). Doctoral dissertation, Maribor: Faculty of Economics and Business, University of Maribor.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 2:

Environment and Human Resources

CONTRIBUTION OF PRIVATE SECTOR TO THE EFFECTIVENESS OF HEALTH CARE PROVISION

Josip Arnerić
University of Zagreb, Faculty of Economics and Business
Trg J. F. Kennedy 6, 10000 Zagreb, Croatia
E-mail: josip.arneric@efzg.hr

Lana Kordić
University of Split, Faculty of Economics
Cvite Fiskovića 5, 21000 Split, Croatia
E-mail: lana.kordic@efst.hr

Abstract: The purpose of this paper is to evaluate the effectiveness of inpatient healthcare facilities in Croatia, in order to find out if private facilities are more effective than public. Using PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluation), we analyzed 99 selected departments representing various specialised, private and public inpatient health facilities in Croatia. For that purpose the 20 indicators are grouped into five criteria to measure effectiveness of inpatient healthcare facilities. Hence, the purpose of the paper is to define tools allowing decision makers to manage their facilities more effectively. Furthermore, this would further improve responsiveness to publicdemands.

Keywords: inpatient healthcare, performance measurement, effectiveness, public and private ownership, PROMETHEE

1 INTRODUCTION

The transition of health care system in developing countries required the government to acquire a new function, such as provision of laws and regulations that support a market economy.In that context, there is a need for good organization concerning the distribution of work, product factor consumption, the encouragement of production force for efficient and effective work, etc. The second goal of modern healthsystemsis related to the microanalysis of the healthcarefacilities' performance. [2]The private health care providers can contribute to the availability of adequate infrastructures to support sustainable growth and social development in the time when infrastructure investment will be challenged by a range of fundamental long-term trends, i.e. ageing population, increasing constraints on public finances, rising quality standards, technological progress and the growing importance of maintaining the existing health care infrastructure [7].Thus, the hypothesis of this paper isthat departments in private inpatient health care facilities are on the average more effective than the departments in public inpatient health care facilities.

Hence, the objective of this paper is todefine the system of indicators for measuring effectiveness developed for inpatient health care facilities. Their purposeis to eliminate most of the imbalance in theinformation available to managers and owners when planning and comparing between healthcarefacilities, i.e. to make decision-making easier for the management and to promote effectiveness within facilities.

The contribution of this study stems from several areas: definingthe system of indicators for measuring effectiveness of inpatient healthcare facilities, measuring the difference in relative effectiveness between public and private healthcare facilities using the defined set of indicators and critical analysis of different scenarios for ranking institutions using PROMETHEE.

The rest of the paper is organized as follows. The second section presents indicators for measuring effectiveness of inpatient healthcare facilities, while the third section provides a

detailed description of sample and model specification. Results are presented in the fourth section. Conclusions are given in the last, fifth section.

2 INDICATORS FOR MEASURING EFFECTIVENESS OF INPATIENT HEALTHCARE FACILITIES

In order to provide a big picture of healthcare facility's current state, a system of indicators has to be defined. Effectivenessevaluates the outcomes of health care,i.e. encourages us to ask if the necessary inputs and outputs are being used in order to produce the best possible outcomes [8].Effectiveness is difficult to assess asthe outcome of health system is under various influencesthat go beyond the health system. Furthermore, the boundary between the output and outcome is not always clear so the efficiency and effectiveness are not easy to isolate. Also, policy makers often want the results of their political program to be presented in terms of outputs. [6]

Hence, in this paper we used corporate social responsibility as a measure of effectiveness in the public sector, for inpatienthealthcarefacilities respectively.The set of indicators can be summarised in five groups of indicators. Each of them describes a part of corporate social responsibility. Thus, philanthropic responsibility was analyzed through outcome indicator, ethical responsibility was analyzed through organizational and quality indicators, legal responsibilitywas analyzed through employee and quality indicators, while economic responsibilitywas analyzed through technical, employee and organizational indicators. [1]

Defined set of indicators are present as follow:

Technical indicators:
- T1: number of patients/number of beds
- T2: number of surgeries/(number of operating rooms*365)
- T3: number of patients/number of beds in the intensive care units

Employee capabilities indicators:
- P1: number of patients/number of doctors
- P2: number of patients/number of nurses
- P3: number of patients/number of administrative staff
- P4: number of patient/number of technical staff
- P5: number of surgeries/number of doctors
- P6:number of surgeries/number of nurses

Organizational indicators:
- O1: number of beds/number of doctors
- O2: number of beds/number of nurses
- O3: number of doctors/number of operating rooms
- O4:number of beds in the intensive care units/number of doctors
- O5: number of nurses/number of operating rooms
- O6:number of beds in the intensive care units/number of nurses

Quality:
- Q1: number of cured patients and patients with improved condition/number of patients
- Q2: number of bed days of cured patients and patients with improved condition/number of cured patients and patients with improved condition

Outcome:
- O1: number of bed days/number of patients
- O2:[(number of bed days/number of beds)/365]*100
- O3: [365-(number of bed days/number of beds)]/number of patients per bed

3 EMPIRICAL FINDINGS ABOUT THE DIFFERENCE IN EFFECTIVENESS BETWEEN THE PUBLIC AND PRIVATE INPATIENT HEALTHCARE FACILITIES

3.1 Data

The sample used in this analysis includes inpatient healthcare departments but not hospitals. Healthcare departments are considered in relation to their homogeneity [4].These departments, chosen in the sample, cover different type of speciality which are included in both, public and private facilities, i.e. cardiosurgery, gynaecology and obstetrics, orthopaedics,surgery and urology. From the 99 selected departments for acute treatments, three are of cardio surgery departments (two public and one private), 29 of them arepubic gynaecology and obstetrics departments, 18 of them are orthopaedics (15 public and 3 private), 31 are the surgery departments (one of them is private) and 18 of them are urology departments (one of them is private). However, all mentioned departments come from facilities which include university hospital centres, teaching hospitals and clinics, clinics, general hospitals and special hospitals.

For each department in the sample five variables were analyzed, technical, employee capabilities, organizational, quality and outcome indicator,respectively.Each variable presents one ranking criteria. Namely, all departments will be ranked according tofive criteria simultaneously. All criteria are standardized (normalized) in order to achieve requirements for maximization. This means that each criterion wasranked on the scale between 0 and 1, higher value being better. All data were collected from the CIPH (Croatian Institute of Public Health) for 2011.

3.2 Methodology

As the main purpose of this paper is ranking 99 hospitals' departments, according to five criteria simultaneously, it is appropriate to use ranking methodology PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluation).This is a multicriteria decision making method.

Multicriteria approach with the PROMETHEE has been previously used in the area of health care. D'Avignon and Mareschal (1989) applied PROMETHEE I and II to determine specialization degrees for hospital services offered in the region of Quebec in Canada, while Urli and Beaudry (1995) applied PROMETHEE for allocation annual funds for development programs in the health and welfare [3]. In 2013, Khamda and Pilevari used PROMETHEE for ranking hospital's health care departments according to the quality criteria [5].

After complete ranking by PROMETHEE II, departments were compared to analyzed differences between departments of public and private health care facilities. Namely, each chosen criteria is the measure of effectiveness. Thus, all departments are ranked in relation to their effectiveness. Department which is the first on the ranking list is most effective one. Also, different scenarios will be considered according to various preferences between criteria but with constant weights, i.e. each criterion has the same weight (Table 1).

Table 1: Different scenarios for complete ranking

	Scenario 1	*Scenario2*	*Scenario 3*
Preference function	usual	U-shape	U-shape
Indifference threshold	/	10%	10%; for quality criterion 5%
Weight	Same weight	Same weight	Same weight

In the first step, the preferences were defined for each criterion. In first scenario, usual preference function was chosen while in the second and third scenario U-shape preference function was chosen. The usual preference function was chosen because in that case all differences were taken into account in the process of ranking, i.e. there is strict preference ofdepartments with the higher values of criteria.In the second and third scenario we want to examine ranking in the case when we are indifferent to small differences.Hence, in the second scenario, wedefined indifference threshold at 10%, i.e. when the difference is smaller than the 10% it is considered as negligible by the decision maker. Finally, in third scenario, we defined indifference threshold at 5% for quality criterion, while for other criteria it was 10%.

As forthe cases in which using PROMETHEE I is not possible for each department to conclude is it effective in comparison to another one, additionally, we solved multicriteria decision making problem by employing PROMETHEE II.

3.3 Results

Table 2 presents ranking results according to the three scenarios using Promethee II, assuming equally weighted criteria but different preference functions.

Table 2: Complete ranking of public and private inpatient healthcare departments

CODE	R_S1	R_S2	R_S3	CODE	R_S1	R_S2	R_S3
PUBLIC_1	80	79	79	PUBLIC_47	64	67	67
PUBLIC_2	99	99	99	PUBLIC_48	91	92	92
PRIVATE_1	1	1	1	PUBLIC_49	87	81	82
PUBLIC_3	85	83	83	PUBLIC_50	48	50	53
PUBLIC_4	73	74	72	PUBLIC_51	56	56	56
PUBLIC_5	76	75	75	PUBLIC_52	79	84	84
PUBLIC_6	62	60	58	PUBLIC_53	49	47	48
PUBLIC_7	93	91	91	PUBLIC_54	15	15	16
PUBLIC_8	18	25	23	PUBLIC_55	62	57	57
PUBLIC_9	71	72	73	PUBLIC_56	96	96	96
PUBLIC_10	47	53	50	PUBLIC_57	61	58	60
PUBLIC_11	9	9	9	PUBLIC_58	97	97	97
PUBLIC_12	57	62	61	PUBLIC_59	40	37	38
PUBLIC_13	60	61	62	PUBLIC_60	92	93	93
PUBLIC_14	44	49	49	PUBLIC_61	41	39	39
PUBLIC_15	75	77	79	PUBLIC_62	11	10	10
PUBLIC_16	89	90	90	PUBLIC_63	25	22	26
PRIVATE_2	6	6	6	PUBLIC_64	83	85	85
PRIVATE_3	22	23	24	PUBLIC_65	66	63	64
PUBLIC_17	88	88	88	PUBLIC_66	74	82	81
PPRIVATE_4	70	77	77	PUBLIC_67	81	75	78
PUBLIC_18	51	48	47	PUBLIC_68	77	69	69
PUBLIC_19	43	45	43	PUBLIC_69	42	44	42
PUBLIC_20	65	68	65	PUBLIC_70	67	65	65
PUBLIC_21	78	71	71	PUBLIC_71	55	54	55

PUBLIC_22	36	35	35	PUBLIC_72	52	55	52
PUBLIC_23	72	73	74	PUBLIC_73	33	28	27
PUBLIC_24	68	70	69	PUBLIC_74	69	66	67
PUBLIC_25	7	8	8	PUBLIC_75	86	87	87
PUBLIC_26	30	24	22	PUBLIC_76	35	37	39
PUBLIC_27	20	19	19	PRIVATE_5	21	19	19
PUBLIC_28	26	18	18	PUBLIC_77	5	5	5
PUBLIC_29	16	14	14	PUBLIC_78	38	43	45
PUBLIC_30	98	98	98	PUBLIC_79	95	95	95
PUBLIC_31	8	7	7	PUBLIC_80	33	32	33
PUBLIC_32	29	34	32	PUBLIC_81	39	40	37
PUBLIC_33	2	2	2	PUBLIC_82	28	29	29
PUBLIC_34	37	36	36	PUBLIC_83	30	31	33
PUBLIC_35	59	59	59	PUBLIC_84	84	86	86
PUBLIC_36	54	52	54	PUBLIC_85	19	21	21
PUBLIC_37	50	46	46	PUBLIC_86	4	4	4
PUBLIC_38	94	94	94	PUBLIC_87	10	13	13
PUBLIC_39	45	42	44	PUBLIC_88	13	11	11
PUBLIC_40	12	12	12	PUBLIC_89	14	17	17
PUBLIC_41	27	27	27	PUBLIC_90	46	41	41
PUBLIC_42	82	79	76	PUBLIC_91	22	26	25
PUBLIC_43	32	32	30	PUBLIC_92	52	50	51
PUBLIC_44	90	89	89	PUBLIC_93	3	3	3
PUBLIC_45	24	29	31	PRIVATE_6	57	64	62
PUBLIC_46	17	15	15				

From results in Table 2 it can be noticed that departments in private inpatient healthcare facilities are higher ranked in comparison to the public departments which means that departments in private inpatient health care facilities are more effective on the average than the public healthcare departments. This hypothesis can be tested using statistical test for the difference in the ranking means, i.e. one-tailed test t test is performed for each ranking scenario (Table 3).

Table 3: Difference in ranking between public and private departments

	Scenario 1		Scenario 2		Scenario 3	
	PUBLIC	PRIVATE	PUBLIC	PRIVATE	PUBLIC	PRIVATE
Mean	51.25806	29.5	51.08602	31.66667	51.09677	31.5
df	6		6		6	
t-statistics	1,84987		1,47582		1,51005	
P(T<=t) one-tail	0,05690		0,09522		0,09089	

According to all scenarios,p-value is lower thanthe significance level of 0.10and therefore null hypotheses can be rejected. This means that departments in private inpatient healthcarefacilities are higher ranked in comparison topublic ones. The higher ranking means more effective providers. The same conclusion is obtained for each scenario. Furthermore,

the ranks between scenarios do not differ significantly according to the Friedman test (p-value=0.914).

4 CONCLUSION

In this paper we have evaluated effectiveness of public and private departments in inpatient health care facilities using data for 99 departments in the Croatian health care system. For that purpose, we defined a set of indicator groupsaccording to five criteria. Each of the aforementioned criteria describes a part of the corporate social responsibility and thus represents the measure of effectiveness. For the purpose of ranking department according to five criteria simultaneously, ranking methodology PROMETHEEII has been applied.

The hypothesis that departments in private inpatient health care facilities are on the average more effective than the departments in public inpatient health care facilities has been confirmed by using the statistical test for the difference in the ranking means. Furthermore, ranks between scenarios do not differ significantly according to the Friedman test, which means that different preferencesare not significant.

Due to lack of knowledge about the impact of different agreements between the public and private sector on effectiveness of health care provision and due to fact that a lot of researchon public private partnership remains inconclusive, all the above mentionedhighlights the contribution of this paper. This study can be improved and further extended to include other parts of the health care system on one hand, and more indicators for analysing a part of the health care system on the another hand. In this moment, the latter is disabled due to data limitation.

References

[1] Andoljšek, Ž. 2005. Družbena odgovornost – paradigma uspešnosti v državni upravi. Uprava, Letnik III, 1: 101-118.

[2] Andoljšek, Ž. 2013. Public Institutions and Performance in EU. 5th International Conference «Economic Integration, competition and cooperation», Croatia, Opatija.

[3] Behzadian et al. 2010. Decision Support PROMETHEE: A comprehensive literature review on methodologies and applications. European Journal of Operational Research, 200:198-215.

[4] Hollingsworth et al. 1999. Efficiency measurement of health care: a review of non-parametric methods and applications. Health care management Science, 2:161-172.

[5] Khamda, Z., Pilevari, N. 2013. Presenting a Model to rank Hospital Wards on the Base of Service Quality Case Study: An Iranian Hospital. Journal of Basic and Applied Scientific Research, 3 (1): 807-817.

[6] Mandl, U. et al. 2008. The effectiveness and efficiency of public spending. European Economy, Economic papers, No. 301.

[7] OECD 2007. Infrastructure to 2030: Mapping policy for electricity, Water and transport. Volume 2, Paris: OECD.

[8] Ozcan, Y. A. 2016. Health Care Benchmarking and performance Evaluation: An Assessment Using Data Envelopment Analysis (DEA), New York: Springer.

PANEL CAUSALITY BETWEEN STATE AID FOR TRAINING AND EMPLOYMENT AND EMPLOYMENT RATE IN EUROPEAN UNION

Kožul Blaževski Renata
University of Split / University Department of Professional Studies
Kopilica 5, 21000Split, Croatia
E-mail: <rkozulb@oss.unist.hr>

Vidović Jelena
University of Split / University Department of Professional Studies
Kopilica 5, 21000Split, Croatia
E-mail: <jvidovic@oss.unist.hr>

Abstract: The un/employment rate is one of the priority political issues meaning that states often interfere in economy in order to boost employment. In this paper we explore connection between state aid expenditures for training and employment and employment rate using panel causality test. Panel data set consists of eleven EU member states in period from 2002 to 2015. Our results suggest that the changes in state aid for employment and training will not create changes in employment rate.

Keywords: state aid, employment, EU member states, training, panel causality.

1 INTRODUCTION

In this paper we explore panel cointegration and causality evidence between state aid for horizontal objectives, primarily state aid for unemployment and training on employment rate using balanced panel data set on eleven EU member states.

The article is organized as follows: in part two we define state aid for unemployment and training and give an overview of the relevant literature on the effects of aid measures on employment. Data and methodology are presented in the third part of paper. In the fourth part of the paper the results of empirical results are presented. Main conclusions are drawn in the fifth part of paper.

2 STATE AID

According to the category, state aid can be horizontal and sectorial, regional and support at the level of local and regional governments. Horizontal aid does not distort competition in relation to the sectorial aid provided to certain sectors or individual entrepreneurs. Types of horizontal measures are aid for research and development and innovation, environmental protection and energy saving, small and medium enterprises, training, employment, culture and the similar. From July 2014 Member States do not longer have to notify in advance to the Commission less distortive aid measures. Most State aid fostering economic growth, jobs and other common interest objectives could therefore be directly implemented by Member States without prior notification or approval by the Commission. At the same time, measures that might seriously harm competition or fragment the Single Market are subject to more careful attention.

There are only few papers which deal with influence of state aid measures for employment and training on the actual unemployment rate. Vidović and Kožul Blaževski (2015) explored connection between state aid expenditures for horizontal and sectorial purposes and GDP using Granger causality test on panel data set of ten EU member states in period from 2000 to 2011. Results suggest that the changes in GDP will create changes in horizontal aid but changes in horizontal aid will not create changes in GDP. [7] To the same conclusions came

Vidović (2015) using Granger causality test on data relating 27 EU countries from 1992 to 2011. [12]

Boon and Van Ours observed effect of all subsidies on the employment on the company level observing 50 Belgian companies. About two years after the subsidies have been granted, beneficiary firms experience, on average, a higher increase in their employment levels, measured in terms of full time employees, than comparable firms that did not receive subsidies. This effect does not seem to last as it disappears two to three years later, when there is again no difference in the evolution of employment of the firms that received subsidies in 2006 and the ones that did not benefit. [1]

Buts and Jegers observed aggregate effect of subsidies on employment level focusing on 50 companies receiving the highest amounts of subsidies in 2006 and comparing them to similar companies that did not receive subsidies on their levels of employment. It appears that firms receiving high subsidies experience a significantly higher increase in full time employment than firms not receiving subsidies. This effect becomes visible about two years after the subsidies have been granted. The effect of the subsidies is thus rather short-lived. The beneficiary firms are not able to transform the extra resources into a lasting effect in terms of employment growth. [3]
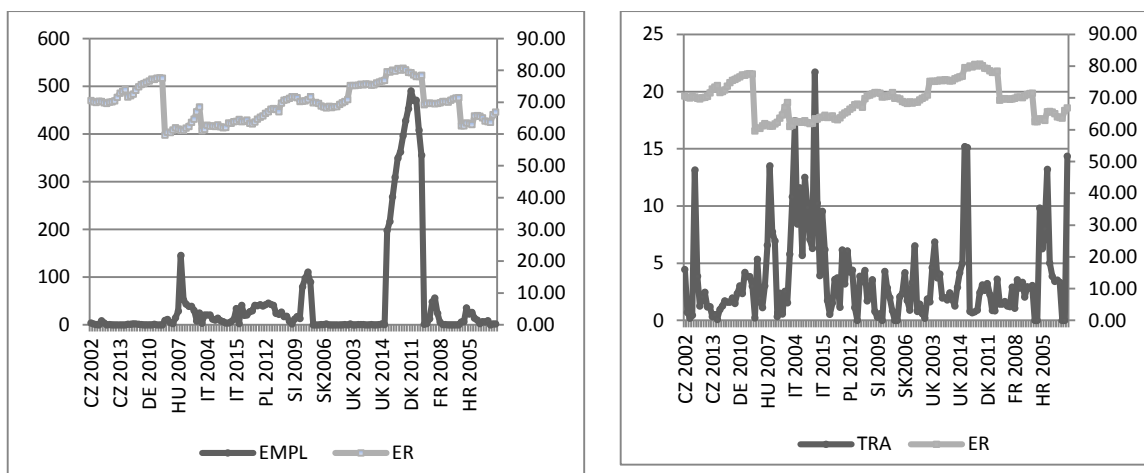
Figure 1: State aid for employment and training per active person, employment rate in period from 2002- 2015 in eleven EU member states. (Prepared by author according to data from Eurostat and EU Comission)

Amount of given state aid between countries is very different, highest state aid for employment in observed period is in Denmark (307 EUR per active person - average for observed period), average state aid for employment per person in remaining ten countries is 14,31 EUR per active person. Highest state aid for training per active person was observed in Italy (21,71 EUR) in 2012 and in Denmark 15,2 EUR in 2002. Average state aid for training in the whole sample in observed period is 3,63 EUR per active person. Highest employment rate was 80,70%, it was registered in Denmark in 2008., and the lowest is 59,70%, it was registered in Hungary in 2002.

3 DATA AND METHODOLOGY

In this study, the causal relationship between employment rate and state aid for employment as well as between employment rate and state aid for training in panel date context is examined.

The Granger causality method, originating from the seminal work of Granger [5], is employed. According to that method, the causal relationship between two variables can be

determined by examining the way they move with respect to each other over time. A variable *x* is said to Granger-cause another variable *y* if future values of y can be predicted better by using past values of *x* and *y*, than by using the past values of *y* only.

The presence of cross-section dependence in the data generating process has implications for the validity of panel unit root tests and Granger causality test results. So, the empirical analysis of this study begins with the application of the cross-sectional dependence tests proposed by Breusch and Pagan [2], Pesaran [10] and Pesaran et al. [11].

At the second stage of analysis the presence of a unit root in the variables is tested by using Hadri-Kurozumi [6] approach. The Hadri-Kurozumi unit root panel test accounts for cross sectional dependence that may be present in the panel. The null hypothesis under the Hadri-Kurozumi unit root panel test is that the series are stationary.

After establishing the order of integration in the series, the existence of a long-term equilibrium relationship between variables of interest is investigated by applying the heterogeneous panel cointegration test developed by Pedroni [8], [9].

At the final stage of the analysis to determine causality relationship between variables Dumitrescu-Hurlin [4] panel causality test is used. One of the advantages of the test is that it considers cross-sectional dependence. Dumitrescu-Hurlin [4] test proposes homogeneous non causality (*HNC*) hypothesis by taking into account both the heterogeneity of the regression model and that of the causal relation.

Table 1: Countries included in analysis

Czech Republic	Germany	Hungary	Italy	Poland
Slovakia	Slovenia	United Kingdom	Denmark	France
Croatia				

The panel data set used in this study consists of annual observations of 11 EU countries. The period under consideration runs from 2002 to 2015. The data, a balanced panel, is chosen based on the availability of data and consists of 154 observations. In Table 1, 11 EU countries observed in this paper are listed. Variables used in the study are: state aid for employment per active person (EMPL), state aid for training per active person (TRA) and employment rate (ER). State aid for employment per active person (EMPL) is defined as ratio between state aid for employment in EUR and number of active persons. State aid for employment per active person (TRA) is defined as ratio between state aid for training in EUR and number of active persons. Employment rate (ER) is defined as employed persons aged 15-64 as a percentage of the population of the same age group. Annual observations on state aid for employment and state aid for training are obtained from web sites of EU Commission relating Competition and State aid and annual observations on employment rate and active population are obtained from Eurostat.

4 EMPIRICAL RESULTS

Empirical analysis starts by testing the presence of cross-section dependence in data. Table 2 presents results from the cross-section dependence test proposed by Breusch and Pagan [2], Pesaran [10] and Pesaran et al. [11].The results suggest that the null hypothesis of no cross-section dependence should be rejected for variables ER and EMPL, with the exception of the results from the CD test for EMPL. For variable TRA, test statistics indicate that there is no cross-section dependence.

Table 2: Results of cross-section dependence tests

Test	EMPL		TRA		ER	
	Statistic	p-value	Statistic	p-value	Statistic	p-value
LM	108.2375***	0.0000	53.56652	0.5295	354.7849***	0.0000
CD$_{lm}$	4.027189***	0.0001	-1.185485	0.2358	27.53456***	0.0000
Bias adjusted CD test	3.604112***	0.0003	-1.608562	0.1077	27.11148***	0.0000
CD	-0.72766***	0.4688	-1.103726	0.2697	11.191***	0.0000

Notes: *** denote the rejection of the null hypothesis at 1% level.

To account for the presence of a unit root in the variables Hadri-Kurozumi [6] panel unit root test is employed and results at levels and first differences are presented in Table 3.

Table 3: Hadri-Kurozumi panel unit root test results

Variable	Level			First difference		
	Test	Statistic	p-value	Test	Statistic	p-value
EMPL	Z_A^{SPC}	6.38638***	0.0000	Z_A^{SPC}	3.03777***	0.0012
	Z_A^{LA}	4.04486***	0.0000	Z_A^{LA}	0.36106	0.3590
TRA	Z_A^{SPC}	0.46804	0.3199			
	Z_A^{LA}	0.86362	0.1939			
ER	Z_A^{SPC}	13.1495***	0.0000	Z_A^{SPC}	1.46236*	0.0718
	Z_A^{LA}	68.0460***	0.0000	Z_A^{LA}	1.53057*	0.0629

Notes: *** and * denote the rejection of the null hypothesis at 1% and 10% levels, respectively.

The results suggest that the null hypothesis of stationarity can be accepted for variable TRA. For variable ER and EMPL results suggest the existence of unit root at level. The null hypothesis of stationarity can be accepted for the first difference of variable ER at the 5% level of significance. In the case of variable EMPL, the results for the first difference of variable from the Z_A^{SPC} test suggest rejection of the null hypothesis of stationarity while that from the Z_A^{LA} test suggest acceptance of the null hypothesis. The Z_A^{LA} test is preferred over the Z_A^{SPC} test if there is evidence of cross sectional dependence in the panel, Hadri -Kurozumi [6]. It can be concluded that the first difference of variable EMPL is stationary.

Since variable TRA is stationary and variable ER is integrated of the order one, it can be concluded that the variables TRA and ER are not cointegrated. The order of integration for EMPL and ER is one, so the existence of a long-term equilibrium relationship among them is explored.

Pedroni [8] suggested a number of panel cointegration tests. Pedroni [9] derives seven panel cointegration test statistics, four are based on within-dimension, and three are based on between-dimension. Table 4 reports the results of the cointegration tests.

	no deterministic intercept and trend		no deterministic trend		deterministic intercept and trend	
	statistic	p-value	statistic	p-value	statistic	p-value
Panel v-statistic	0.965290	0.1672	-0.972613	0.8346	1.786222**	0.0370
Panel rho-statistic	-1.081775	0.1397	0.525034	0.7002	0.609233	0.7288
Panel PP-statistic	-2.472485***	0.0067	-1.106663	0.1342	0.721861	0.7648
Panel ADF-statistic	-2.382472***	0.0086	-0.910058	0.1814	-3.163382***	0.0008
Group rho-statistic	-0.746424	0.2277	-0.292932	0.3848	0.977336	0.8358
Group PP-statistic	-3.740078***	0.0001	-3.731531***	0.0001	-2.941475***	0.0016
Group ADF-statistic	-4.203102***	0.0000	-3.330292***	0.0004	-4.174430***	0.0000

Notes: ***, ** denote the rejection of the null hypothesis at 1% and 5% levels, respectively. Lag length for based on Schwarz information criterion (SIC)

Based on the p-values, corresponding to the seven test statistics, the null hypothesis of no cointegration cannot be accepted at the significance level of 5% in ten (out of 21) cases. Thus, it can be concluded that variables EMPL and ER are not cointegrated, although the evidences that support such conclusion are rather weak.

Table 5: Dumitrescu-Hurlin panel causality test results

Null Hypothesis	Lags (k)	W statistic	Zbar statistic	p-value	Conclusion
ER does not homogeneously cause EMPL	k=1	0.76690	-0.74815	0.4544	ER ---EMPL
	k=2	2.46491	-0.34154	0.7327	ER ---EMPL
EMPL does not homogeneously cause ER	k=1	1.26854	-0.02476	0.9802	EMPL ---ER
	k=2	2.15999	-0.53616	0.5918	EMPL ---ER
ER does not homogeneously cause TRA	k=1	1.39518	0.15785	0.8746	ER --- TRA
	k=2	2.26121	-0.47156	0.6372	ER --- TRA
TRA does not homogeneously cause ER	k=1	1.27453	-0.01613	0.9871	TRA --- ER
	k=2	3.91883	0.58648	0.5576	TRA --- ER

Notes: X --- Y denotes that the null of no causal effect from X to Y cannot be rejected.

Granger causality test will help us conclude whether past values of one variable affect another variable in the current period. These test results also indicate the directions of causal relationships between variables. Dumitrescu-Hurlin [4] panel causality test is used to test for causality between ER and EMPL as well as between ER and TRA. In order to assess sensitivity of results to the choice of the common lag order, both statistics are computed for one and two lags. More than two lags would lead to the degrees of freedom problem for model. The loss of degrees of freedom due to over-parameterization can lead to inefficient

estimates. Also, variable TRA at the level and first difference of variables EMPL and ER are used in the model.

The results of Dumitrescu-Hurlin [4] panel causality test are presented in Table 5. For lags one and two, the null hypothesis of non-causality from ER to EMPL, from EMPL to ER, from ER to TRA and from TRA to ER cannot be rejected. So, it can be concluded that there is no causal effects between EMPL and ER as well as between TRA and ER.

5 CONCLUSIONS

When observing state aid for employment and training one should be aware that state aid expenditures are political decision and data relating state aid do not have continuous behavior. Absolute pioneer of state aid expenditures for employment is Denmark while some countries in certain periods had zero expenditures: Chez Republic, Slovakia and Croatia. Italy and Denmark had in certain periods highest state aid expenditures for training, while some countries had zero expenditures: Croatia, Slovakia and Slovenia. Results of Dumitrescu-Hurlin panel causality test indicate that state aid expenditures for training and employment had no influence on employment rate. Due to shortness of the time period under consideration it was not possible to apply Dumitrescu-Hurlin panel causality test for the common lag order greater than two. This needs further attention since results indicate on possible existence of such causality. Similar conclusions had Boon and Van Ours and Buts and Jegers. Existence of a long-term equilibrium relationship between state aid for employment and employment rate should be also further examined since the evidences of no cointegration are rather weak.

References
[1] Boone, J., van Ours, J. C. 2004. Effective Active Labor Market Policies. CEPR Discussion Paper No. 4707. Available at SSRN: https://ssrn.com/abstract=641561 [Accessed 11/06/2017]

[2] Breusch, T. and Pagan, A.R. 1980. The Lagrange multiplier test and its applications to model specification tests in econometrics. Review of Economic Studies, 47(1): 239-253

[3] Buts, C., & Jegers, M. 2013. Does State Aid Create Jobs: The Short and Mid-Term Employment Effects of Subsidies. Eur. St. Aid LQ, 651

[4] Dumitrescu, E. and Hurlin, K. 2012. Testing for Granger non-causality in heterogeneous panels, Economic Modelling, 29(4): 1450–1460

[5] Granger, C. W. J. 1969. Investigating causal relation by econometric and cross-sectional method. Econometrica, Vol (37): 424–438

[6] Hadri, K. and Kurozumi, E. 2012. A simple panel stationarity test in the presence of serial correlation and a common factor, Economics Letters, 115 (1): 31–34

[7] Kožul Blaževski, R., Vidović, J. 2015. Causality between State Aid and GDP in European Countries, Proceedings of the 13th International Symposium on Operational Research SOR'15 in Slovenia, Bled, September 23 - 25, 2015., p. 407-412

[8] Pedroni, P. 1999. Critical values for cointegrating tests in heterogeneous panels with multiple regressors. Oxford bulletin of economics and statistics, Vol (61): 653–670.

[9] Pedroni, P. 2004. Panel cointegration; asymptotic and finite application to the PPP hypothesis, Econometric theory, Vol (20): 597-624

[10] Pesaran, H.M. 2004. General diagnostic tests for cross-section dependence in panels. University of Cambidge, Faculty of economics, Cambridge working papers in economics No. 0435

[11] Pesaran, H.M., Ullah, A. and Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. Econometrics Journal, 11(1): 105-127

[12] Vidović, J. 2015. Analysis of state aid in the European Union with emphasis on economic growth. Proceedings of the University of Dubrovnik, 2(2), 69-84

PREDICTING THE ACCEPTABILITY OF MUSIC WITH ENTROPY OF HARMONY

Lorena Mihelač
School center Novo mesto, IT department, Slovenia
E-mail: lorena.mihelac@sc-nm.si

Janez Povh
University of Ljubljana, Faculty of Mechanical engineering, Slovenia
Institute of mathematics, physics and mechanics, Ljubljana, Slovenia
E-mail: janez.povh@fs.uni-lj.si

Abstract: Although the information theory of entropy was developed to determine the optimal method of encoding the message for radio transmission, since 1948 it has been taken over as an analytical tool in various fields, including linguistics, literary criticism and musical theory. In the field of music, the measurement of complexity (entropy) has opened a path for understanding why certain musical compositions are more acceptable to listeners than some other musical compositions. The article presents the results of the research from 2016, which has shown the influence of the entropy of harmony on the acceptability of music, and raised the awareness that the entropy of harmony could be used as a predictive model of auditory acceptability of music.

Keywords: acceptability of music, harmony, symmetry, entropy of the harmony

1 INTRODUCTION

Music, like language, uses the rules and ways in which symbols are made up of smaller sound structures into larger sound structures [9], and, as with language, it may be expected that symbols or even smaller / larger audio structures are repeating.

A higher frequency of a symbol or sound structure means that there is a certain degree of predictability [8], which enables the listener to identify, for example, what will follow a single tone, a shorter melody, or a chord in harmony as well.

Predictability in music is associated with complexity [6]. Continuity in the musical composition leads to low complexity, a change to a higher complexity. If musical material is repeated all the time (from the point of view of melody, harmony, rhythm ...), then it requires significantly less attention (mental space) from the listener as the material that is constantly changing [8].

If we proceed from the idea that a musical composition is an information [6], then from the point of view of predictability in sound content and its complexity and from the point of view of the information theory, that measures this complexity, we can speak of a lower or higher entropy in musical composition.

Higher level of predictability in musical composition therefore means lower complexity or a lower degree of indeterminacy (lower entropy) and thus a more acceptable composition, while lower predictability results in a higher complexity (higher entropy) and a less acceptable composition [3], [8].

In music, the measurement of complexity (entropy) has opened the way to understanding why certain musical compositions are more acceptable to listeners than some other musical compositions [6].

Initial research in this area was mostly focused on the measurement of the entropy of the musical style [4], [7], [2], but less on the measurement of the entropy of individual elements in music (melody, rhythm, harmony), although a multiplicative measurement of the entropy of musical elements [1], [11] has shown that it is necessary to take into account various elements of music (melody, rhythm, harmony) and their influence on entropy of musical composition.

Therefore, this paper is presenting a research which was carried out in 2016 with the main purpose to find out, what is the impact of entropy in harmony on the acceptability of a musical composition. A short explanation about harmony will be provided in chapter 2 and about the measuring of entropy of harmony in chapter 2.1 and chapter 3 as well. The main results of the research will be presented in chapter 4, in chapter 5 some questions about indicators of entropy in harmony as also the indicators for the acceptance of a musical piece will be opened for further discussion and research work.

2 THE ENTROPY OF HARMONY

Some research works about entropy in musical style has shown that the entropy of the musical style is rising due to the musical building elements which complexity is also growing through the historical development of music [4], [10].

Other research works about music entropy are explaining that the entropy is the same or almost identical in musical styles from different musical periods, regardless of the complexity of musical elements [2]. The question arises, what to measure in a musical composition if we want to determine the impact of entropy on the acceptability of musical composition in listener.

If the height of entropy in musical styles from different musical periods is the same or almost the same, nevertheless that the complexity of the same music building elements that define a musical style (melody, rhythm, harmony) is the same, then obviously one should not analyze entropy of the musical style (with all the entire musical elements). It is necessary to find the answer on the question which musical elements are actually affecting the value of entropy in a musical composition, and, consequently, also the predictability and acceptability of music content.

As harmony plays a fundamental role in all musical styles, from classical, pop, rock, Latino or jazz music and is one of the foundations of Western tonal music, the authors have decided to explore the impact of harmony on the entropy of musical composition as also how the entropy of harmony affects the acceptance of music.

2.1 The measurement of unigram and bigram in music harmony

Harmonic structure describes a musical work as a time series of discrete, overlapping musical elements (chords) composed of a comparatively small-defined alphabet [9]. Similar to a text document, where it is possible to present letters, syllables, words or phonemes with a plurality of *n-grams* (the sequence of *n* elements), it is also possible in the music to present a music smaller/larger content with a plurality of *n-grams* samples [13], [14].

In measuring the entropy of harmony, we are interested in *n-gram*, which consists of the harmonic current (sequence) of the chords and the choice $n = 1$ (unigram), that is, the individual chords in the harmonic flow or the choice of $n = 2$ (bigram), where we look at the connection of two adjacent chords as it is shown in figure 1:
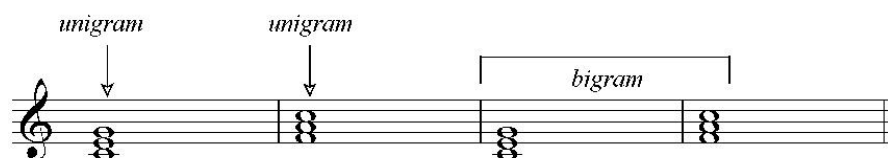
Figure 1: Unigram and bigram in harmonic flow

With the distribution of unigrams and bigrams, the idea is to find out which chord functions in a harmonic flow are more important (frequent) and how these functions affect the acceptance of music.

3 MEASURING THE ENTROPY OF HARMONY

An empirical study on predicting the acceptance of music based on the entropy of harmony was carried out in October 2016. For the purpose of the research, musical examples were selected based on two criteria: a) given their popularity or frequency of listening and listing on the list of the most listened music on Billboard; and b) given the complexity of the harmonic flow, which is higher in relation to the higher number of transitions between the chords in the harmonic flow, or higher in relation to the number of functions (main and lateral functions).

With these criteria, a set of 160 musical pieces was obtained, of which 80 of them are popular musical pieces (highly ranked on popularity and low ranked in harmonic complexity) and 80 unpopular musical pieces (low ranked on popularity and high ranked in harmonic complexity).

The sound quality of each musical example was analyzed. The entire set of musical pieces went through tonal processing (in order to cut each musical piece to the appropriate length and to balance the appropriate volume) with the Audacity program, and through the analysis of the harmonic flow using the Sonic Visualizer program. The entropy (unigram and bigram) of each individual music piece was also measured using the R Studio program. Shannon's formula was used to measure entropy:

$$H\ (X) = -\sum_{x \epsilon X} p(x) log_2 p(x)$$

(1)

In this formula, the set $X = \{x_1, x_2, ..., x_n\}$ is the set of chords in a harmonic flow, or the set of functions of chords in the harmonic flow, and *p(x)* the probability that an event occurs (chord function ...).

We evaluated these 160 music pieces by 21 evaluators. Firstly they have answered five general questions in order to determine gender, age, music education, music engagement (amateur/professional) and one special (sixth) question in order to determine whether an individual can define the acceptability of a music example only on the basis of the initial part (introduction) of the musical piece.

Later, each evaluator evaluated each of these 160 music pieces by four main criteria with the scale from 1 to 5: the difficulty of listening to the musical piece, the pleasantness of the musical piece, the recognition of the musical piece and the repeatability of the piece (the readiness of the evaluator to listen to the whole music piece, not only to a part of it).

4 RESULTS

The results of evaluating the 160 musical pieces, obtained by R Studio with ANOVA analysis have shown:
- the style of music influences the acceptance (pleasure) of a musical piece,
- the entropy of the harmonic flow correlates with the acceptance (pleasantness) of a musical piece,
- the average of entropy is different in each music style, which means that there is a connection between musical style and entropy.

The research has also shown that the acceptance of a musical piece and the height of the entropy in the harmonic flow is influenced by symmetry, which is indicated by the fact that a certain chord in the harmonic flow is followed by another chord with high probability. For the purpose of an accurate measurement, a new symmetry scale was used where the degree of symmetry (full symmetry, partial symmetry, no symmetry) and type of symmetry (linear symmetry, successive symmetry, and rotational symmetry) were given.

Hi-square tests and ANOVA analysis have shown that entropy is low or moderately low in musical pieces with full symmetry or partial symmetry. This means that there is such a harmonic syntax (flow) in which harmonic patterns (chords) are more often repeated and the degree of variation/change in the harmonic flow is low or relatively low.

From the point of view of the acceptability of music in listener, this means that in a harmonic flow tonal relations are obviously important because they create a sense of relaxation and predictability [5], [12].

Hi-square test and ANOVA analysis have confirmed that the evaluators are experiencing and evaluating musical pieces different according to the different entropy values. This means that the entropy of a musical piece could be used to predict the acceptance (pleasure) of a musical piece. This assumption was tested with linear discriminant analysis (LDA) where musical pieces were divided into $r = 50:50$ (in each group approximately the same number of musical pieces).

To pleasant musical examples was given the value "1" and to less pleasant one the value "0". In this relationship, 54 unpleasant and 49 pleasant musical examples were classified correctly, yielding classification accuracy = 64%. The classification precision for less pleasant musical pieces was 66% and for more pleasant musical pieces 63%, as can be seen from the following results in table 1:

```
> ldapred <- predict(data1.lda,data1) # predict 1
> ldaclass <- ldapred$class           # predict 2
> table(ldaclass,data1$prijetnost)    # predict 3 (matrix)

ldaclass   0   1
       0  54  28
       1  29  49
> ct <- table(ldaclass,data1$prijetnost) #accuracy
> diag(prop.table(ct,1)) # percent correct for each category
         0          1
0.6585366  0.6282051
> sum(diag(prop.table(ct))) # total percent correct
[1] 0.64375
```

Table 1: The accuracy and precision of the prediction model for music acceptability

5 CONCLUSIONS

Although the research covered a very large area of entropy in harmony, there remain several open questions: (i) which indicators should be taken into account when measuring the entropy in the harmonic flow; (ii) which features should have a musical composition to be accepted by listener.

We foresee that the most likely are the three indicators of entropy:
- number of functions in the harmonic flow. A detailed analysis of the harmonic flows of all 160 musical examples has shown: a) that entropy is getting higher with the number of different chords; b) that the entropy in two harmonic flows with the same number of different chords can be different because in one of these two harmonic flows one of the chords is repeated;
- symmetry in the harmonic flow. Musical pieces with a higher degree of symmetry have a lower degree of complexity and hence a lower entropy value;
- acceptance of the musical piece. Unpredictable chords in a harmony flow, unusual tonal relationships, or excessive variation in harmonic flow affect the acceptance of the musical piece.

As to question two (ii), we foresee that a musical composition should have these features for a higher acceptability in listener:

- the musical composition should contain a certain measure of predictability, which allows the individual to identify musical patterns, which of course depends on his acquired knowledge, his belonging to a culture and his exposure to certain music styles;
- the musical composition should contain a certain measure of predictability, because it creates greater comfort and, consequently, greater pleasure;
- the musical composition should contain a certain measure of variation, or such musical content that is not heard (too) often, otherwise the musical composition is perceived as less pleasant due to the fact that an individual has already heard similar musical content;
- the musical composition should include such chord relations in a harmonic flow that are interesting, "fresh" and do not cause excessive listening tension in listener;
- the musical composition should contain clearly transmitted sound content in the sense that at least one of the musical elements (melody, rhythm, harmony ...) is clearly understandable and definable. If the musical structure is not clearly transmitted or clearly defined by the listener, the acceptance of a musical piece is lower due to the fact that the pleasure in experiencing music content is lower;
- the musical composition should contain a certain degree of symmetry, since symmetry affects the lowering of the complexity in the harmonic flow or influences the lowering of the value of entropy, and has an impact on higher hearing acceptance of music in listener.

References
[1]   Conklin, D. and Witten, I. H. 1995. Multiple viewpoint systems for music prediction. Journal of New Music Research, Vol. 24, 51-73.
[2]   Febbres, G. and Jaffe, K. 2015. Music viewed by its Entropy content: A novel window for comparative analysis. Retrieved from https://arxiv.org/ftp/arxiv/papers/1701/1701.04064.pdf [Accessed 10/October/2016].
[3]   Hansen, N. C. and. Pearce, M. Shannon entropy predicts perceptual uncertainty in the generation   of melody pitch expectations. Retrieved from http://pure.au.dk/portal/files/51824943/Hansen_Pearce_2012_Shannon_entropy.pdf [Accessed 2/January/2017]
[4]   Hiller, L. and Bean, C. 1966. Information Theory Analyses of Four Sonata Expositions. Journal of Music Theory, Vol 10 (1): 96-137.
[5]   Krumhansl, C. L. 2002. Music: A link between cognition and emotion. Current Directions in Psychological Science, Vol (11): 45-50.
[6]   Madsen, S. T. and Widmer, G. 2006. Music complexity measures predicting the listening experience. Proceedings of the 9t International Conference on Music Perception & Cognition
[7]   Margulis, E. H.and Beatty, A. 2008. Musical style, psychoaesthetics, and prospects for entropy as   an analytic tool. Computer Music Journal, 32 (4), 64 – 78.
[8]   Maršík, L. 2013. Music Harmony Analysis: Towards a Harmonic Complexity of Musical Pieces. Master thesis. Bratislava: Department of Computer Science, Comenius University in Bratislava.Music Theory, Vol. 10 (1), 96–137.
[9]   Mihelač, L. 2017. Predicting acceptability of music with entropy of harmony. Master thesis. Novo mesto: Faculty of information studies.
[10]   Patel, A. D. 2007. Music, language, and the brain. Oxford: Oxford University Press.
[11]   Pearce, M. 2007. Early Applications of Information Theory to Music. Retrieved from http://webprojects.eecs.qmul.ac.uk/marcusp/notes/music-information-theory.pdf [Accessed 10/June/2016].
[12]   Scott, J. S. 2005. A multi-dimensional entropy model of jazz improvisation for music information retrieval. Doctoral thesis. Texas: University of Nort-Texas.
[13]   Toiviainen, P. and Krumhansl, C. L. 2003. Measuring and modelling real-time responses to music: The dynamics of tonality induction. Perception, Vol (32): 741-766.
[14]   Tseng, Y. 1999. Content-based retrieval for music collections. SIGIR, 172-182.
[15]   Uitdenbogerd, A. 2002. Music Information Retrieval Technology. Doctoral thesis. Melbourne: Royal Melbourne Institute of Technology.

RURAL AND URBAN DISPERITIES IN FULL ROUTINE IMMUNIZATION COVERAGE FOR UNDER-5 CHILDREN IN NIGERIA: A MARKOV CHAIN ANALYSIS

Phillips Edomwonyi Obasohan
President of the Institute for Operations Research of Nigeria (INFORN)
Executive Committee member of African Federations of Operations Research Societies (AFROS)
Department of Liberal Studies, College of Administrative and Business Studies, CABS, Niger State
Polytechnic, Bida Campus, Nigeria
E-mail: philiobas@yahoo.com

Abstract: The World Health Organization has earlier projected that full immunization of children should reach 90% at the national level by 2010. However, in spite of the huge resources committed to actualizing this target, Nigeria remains one of the top 10 countries in the world with full immunization uptake by her children less than 50%. The most worrisome of this is that wider disparities exist in full routine immunization coverage for children in rural and urban areas. This paper uses Markov Chain model to analyze the data of disparities between these areas and establishes that the gaps may persist for a very long time except intervention gear towards closing up the gaps are put in place.

Keywords: Immunization, Disparities, Markov Chain Application, Rural, Urban

1 INTRODUCTION

Globally, Immunization of children from Vaccine Preventable Diseases (VPDs) has remained one of the most important public health interventions [1,2]. In the last 10 years, there had been so much progress attained in Immunization coverage for children worldwide [3,4,5]. In spite, an estimated 21.8 million infants are still not been covered by routine Immunization services [3], with greater percentage of this from sub-Saharan African countries. In Nigeria, VPDs are attributable for 22 % of childhood deaths amounting to over 200,000 deaths per year [6,7]. For instance, in a recent national survey conducted in 2013 by National Population Commission/ICF Macro, indicated that the full routine Immunization coverage for Nigerian children was 25.3% [8], well short of the 2015 goal of 90% [5]. Restrepo-Méndez et al in [5] had observed that the progress which is commonly been made in Immunization coverage are often expressed in terms of national and regional mean values, whereas many underlying disparities among and within countries go unobserved or not been reported. In Nigeria findings from Nigerian Demographic and Health Survey data have revealed alarming wide disparities in full routine Immunization coverage between rural and urban children. The country in the recent times has shown pro-urban differences of not less than 20 percentage points [5]. There are many factors that have accounted for this gap in Immunization coverage and are well established in literature [2,9]. These include poor access to health care centres, poor educational levels in rural areas, low economic power and poor accessibilities to these localities [5,10,11]. However, research to understand the future behaviour of these disparities has not been fully established. The aim of this study therefore is to analyze the extent of the disparities that existed in full routine Immunization coverage between rural and urban Nigerian children under the age of five years with the view of projecting if the disparities can close up on or before year 2030 (the year United Nations intend to achieve 17 Sustainable Development goals) using Markov chain model. More specifically, the study determined

1. the trend in routine Immunization coverage between Nigeria's urban and rural children over the past twenty-three years for which national data are available
2. determine if this trend in disparities is independent of the period under consideration

3. given the historical data, to generate a probability that would predict the future trend of the disparities in full routine Immunization coverage between rural and urban Nigerian children
4. establish if there are chances of bridging these disparities that have existed in full routine Immunization coverage between rural and urban Nigerian children on or before 2030 (13 years from now).

The findings from this study will further add to the current knowledge regarding the immunization status of Nigeria Children in particular and in developing countries in general. In addition, the results may help as platform of having evidence based decisions for policy interventions that may reduce disparities in full routine Immunization coverage among children of different demographic/cultural settings.

2 MARKOV CHAIN MODEL SPECIFICATIONS

A process where the outcome of a given experiment can affect the outcome of the next experiment is called Markov chain [12]. In other words, a Markov chain process is a stochastic, mathematical model with transition probabilities that provide information about how to relate one stage of the process to the next [13]. However, in this study, the author had used discrete Markov Chain which has been defined in [14] as a stochastic processes with finitely many states on the nominal scale having the following properties:

- the number of possible outcomes or states is finite
- the outcome at any stage depends only on the outcome of the previous stage
- the probabilities are constant over time.

Symbolically, a discrete-time Markov chain with (finite or countable) state space Ω is a sequence $X_0, X_1, \dots$ of Ω - valued random variables such that for all states i,j, $k_0, k_1, \dots$ and all times n=0, 1, 2,...., then

$$P(X_{n+1} = j \parallel X_n = i, X_{n-1} = K_{n-1}, ..) = P(X_{n+1} = j \parallel X_n = i) = p_{ij}(n) \qquad (1)$$

where $p_{i,j}(n)$ depends only on the states i,j, and not on the time n (which in this study is expressed in number of years) or the previous states k_{n-1}, k_{n-2}. The numbers $p_{ij}(n)$ are called the transition probabilities of the chain [14,15,16]

Out of the set of states the process starts in one of the states and moves successively from one state to another then $X_n=\{x_{1n}, x_{2n}, \dots x_{rn}\}$. If the chain is currently in state x_i, then it moves to state x_j at the next step with a probability denoted by p_{ij} (n) [12]. In this study the author used a one-step transition Markov chain such that $P=|p_{i,j}(n)|$ is one-step transition matrix and $\sum_i^r p_{i,j}(n) = 1$ and $p_{i,j}$ (n) > 0 for all i and j, where r is the number of states used. Hence P will be the matrix of transition probabilities using the mathematical expression for steady state probability, y_x such that $Py_x = y_x$, hence $\lim_{t \to \infty} P^t y_o = y_x$, where y_o is the initial state vector [14]

$$\text{Then, } y_x = \frac{y}{\sum_{t=1}^r y} \qquad \text{for r=1,2,3} \qquad (2)$$

Our interest in this study is to consider the future behaviour of P whether it tends to a constant vector (steady state) as the period n tends to infinity. In this case, the premises under which this will occur are that:

- P must be a regular transition matrix (if for any integer m, all entries of Pm are strictly positives) of a Markov chain process
- That y is any state vector

Then as n approaches infinity, $P^m y$ approaches a fixed probability vector whose sum of its entries is 1 and all entries are positive [17]

3 SOURCE OF DATA FOR ANALYSIS

Data used for this study were extracted from the a national representative survey reports of Nigeria Demographic and Health Survey (NDHS) which is carried out every five years with the latest being 2013. This last report is the fifth in the series since 1990. The objective was to provide up-to-date information on health related matters that will assist policymakers and programme managers in evaluating and designing programmes and strategies for improving health and family planning services in the country [8].

The Nigeria Expanded Programme on Immunisation (EPI) mirrors the international recommendations of the World Health Organization. A child is considered fully vaccinated if she or he has received BCG Immunization against tuberculosis; three doses of vaccine to prevent diphtheria, pertussis, and tetanus; at least three doses of polio vaccine; and one dose of measles vaccine. These vaccines should be received during the first year of life [8]. However, in May 2012, Nigeria began the phased replacement of the diphtheria, pertussis, and tetanus (DPT) vaccine with the pentavalent vaccine, which contains more antigens (DPT, *Haemophilus influenzae* type B, and hepatitis B) [8]

4 ANALYSIS AND RESULTS

4.1 Objective 1:

Figure 1 reveals marked disparities in full routine Immunization status between urban and rural areas of Nigeria. The gap exhibited a tendency of closing up between 1999 and 2003, but began to widen out from 2008 to 2013 data. It is observed that trend in urban area has a u-shape indicating a fallen and rising up again.
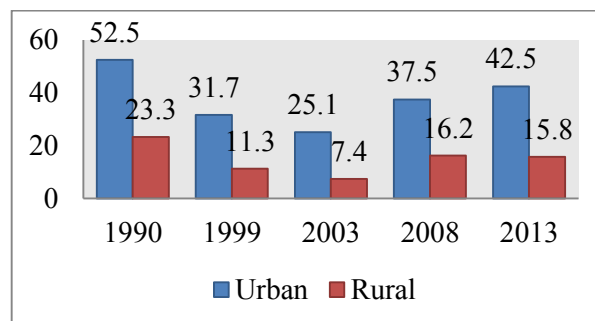
Figure 1: Bar Charts displaying the trend in full routine Immunization status of children under-5 years in rural and urban areas (1990 – 2013)

4.2 Objective 2:

Furthermore, we need the mathematical model to examine these disparities in more details so as to establish the objectives of our study. In the first instance, we need to determine the transition counts of the disparities (here denoted as d) and find out if they are independent of the time, n.

Table 1: Disparities in full Routine Immunization for Rural/Urban Nigerian Under-5 Children Computed from NDHS (1990 – 2013)

Years (n)	1990	1999	2003	2008	2013
Disparities (d)	29.2	20.4	17.7	21.3	26.7

That our disparities can be grouped into classes of our choice to define our states [14]; therefore, table 2 identifies these transition states and the d classes made up of 5% interval each for rural/urban disparities. The author's choice to use 5% interval is by convenience and on the premises that the data under consideration were collected in the interval of 5 years

Table 2: Transition states versus d classes of Disparities in Rural/Urban

States	X_1	X_2	X_3
d Classes	16 – 20	21 – 25	26 – 30

Using table 1 and 2, the author got the transition counts and matrix by determining the initial class and counts it movement from the state of that class to every other states.

Therefore, the second objective requires that we establish that disparities observed over time in full routine Immunization of rural/urban Nigerian under-five children is statistically independent of the periods. [13,18] have stated that we could consider the transition counts of disparities as contingency values and use chi square test. The chi square value of 5.005 obtained was significant at 95% confidence interval; therefore we conclude that the transition counts are independent of the time

4.3 Objective 3:

With respect to the third objective and using the mathematical equation 2 on tables 1 & 2, we have:

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{3}$$

by the nature of this matrix we are confident that the successive probabilities matrix will reach its steady state because considering a fourth iteration of P for instance, we have ,

$$p^4 = \begin{bmatrix} 0.5625 & 0.3125 & 0.1250 \\ 0.2500 & 0.2500 & 0.5000 \\ 0.6250 & 0.1250 & 0.2500 \end{bmatrix} \tag{4}$$

which is strictly a regular transition matrix as such will converge to a steady state [17].

4.4 Objective 4:

Now considering the last objective , the question here is whether Nigeria will be able to bridge the disparities in full routine Immunization coverage between under-5 rural and urban children by year 2030 (i.e., 13 years from now). If not, how long will it take her to accomplish this if the present process is not changed?

Consider the transition matrix after 13 steps, we have:

$$P^{13} = \begin{bmatrix} 0.50549316 & 0.24719238 & 0.24731445 \\ 0.49462891 & 0.25830078 & 0.24707031 \\ 0.49438477 & 0.24731445 & 0.25830078 \end{bmatrix} \tag{5}$$

Using eight digits decimal places, we observed that the transition matrix could not reach the steady state and therefore we conclude that Nigeria cannot bridge up the disparities that now exist in full routine immunization coverage for under-five years' children between rural and urban areas in the next 13 years.

Now, how long will it take Nigeria to close up this gap? By iterating this computation further, we observed that after the 54th iteration, the transition matrix entered into steady state:

$$P^{54} = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix} \qquad (6)$$

To establish that the steady state had not resulted before this step, we considered the 53rd steps if rounding off of decimal is not accommodated as recommended in [17] that we keep as many decimal places as possible to retain accuracy and we got:

$$P^{53} = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.50000001 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix} \qquad (7)$$

5 CONCLUSION

This study has revealed some important realities of full routine immunization processes in Nigeria:
- the disparities which had hitherto existed in full routine immunization coverage between rural and urban under-five Nigeria children will likely continue for more than year 2030
- the disparities will likely close up only after half a century if nothing deliberate is done to enhance routine immunization in the rural areas

6 LIMITATIONS OF STUDY

The interpretations of the results obtained from this study must be given bearing in mind the following limitations
   i. data for the analysis was only available for 5 periods which were collected on every five years, this would have yielded more accurate forecast if data were available for more periods and at shorter intervals
   ii. The mathematical model developed in this study did not take into consideration the effects some other potential factors (socioeconomic and demographic) which may have effect on bridging the gap in full routine immunization status between rural and urban under-five children.
To inculcate these socioeconomic and demographic factors it could be the subject of future investigation.

7 STUDY IMPLICATIONS/RECOMMENDATIONS

This study has a number of implications with respect to how the immunization processes have been conducted in Nigeria.
   i. In spite of the numerous programs that have been put in place over the years by Nigerian Government, for instance, various activities to attain the 2015 Millennium Development Goals (MDGs), the gap that existed between 2003 and 2013 showed an increasing rate. This study therefore is an eye opener that if meaningful achievement is to be attained in this direction, policies are needed to be directed consciously at closing up these gaps
   ii. The model which was developed revealed that these gaps will likely not close up before the end of 2030 (13 years from now), those concern in the implementations of immunization processes will need to double their efforts at the rural areas

iii. Some identified barriers to immunization processes in the rural areas should be addressed. For instance, rigorous campaign through advocacies using village heads and religious leaders; mass literacy campaign for rural women; making vaccines available at the rural areas with the provisions of solar refrigerators for preservations of the vaccines.

References

[1] Olumuyiwa OO, Alufohai EF, 2008. Meurice FP and Ahonkhai VI, Determinants of vaccination coverage in rural Nigeria *BMC Public Health*, 8:381 doi:10.1186/1471-2458-8-381

[2] Antai D. 2011. Rural-Urban inequalities in childhood immunization in Nigeria: The Role of Community Contexts; *African Journal of Primary Health Care and Family Medicine,* Vol 3, No 1

[3] Fulfilling the health agenda for women and children: The 2014 report. Geneva: World Health Organization; 2014. Available from: http://www.countdown2015mnch.org/documents/2014Report/ Countdown_to_2015-Fulfilling%20the%20Health_Agenda_for_Women_and_Children-The_2014_Report-Conference_Draft.pdf accessed 22/03/2017.

[4] Immunization summary. A statistical reference containing data through 2013. New York: United Nations Children's Fund; 2014. Available from: http://www.who.int/immunization/monitoring_ surveillance/ Immunization_Summary_2013.pdf accessed 22/03/2017.

[5] Restrepo-Méndez , MC, Barros AJD, Wong KLM, Johnson HL, Pariyo G, Franca GVA, Wehrmeister FC & Victora CG. 2016. Inequalities in full immunization coverage: trends in low- and middle-income countries. *Bulletin of the World Health Organization,*94:794-805B. doi: http://dx.doi.org/10.2471/ BLT.15.162172

[6] *United State Agency for International Development (USAID) for Africa, Immunisation Basics. Strengthening Routine Immunisation Services and Sustainable Financing for Immunisation 2009.* http://www.immunisationbasics.jsi.com/CountryActivities.htm. accessed 20/03/2017

[7] Adedire EB, Ajayi I, Fawole OI, Ajumobi O, Kasasa, S Wasswa P & Nguku P. 2016. Immunisation coverage and its determinants among children aged 12-23 months in Atakumosa-west district, Osun State Nigeria: a cross-sectional study. *BMC Public Health*, 16:905. DOI: 10.1186/s12889-016-3531-x

[8] National Population Commission and ICF Macro. 2014. Nigeria Demographic and Health Survey 2013. Abuja, Nigeria: National Population Commission and ICF Macro

[9] Obasohan PE, Anosike BU, Etsunyakpa MB. 2015. Determinant of Full Immunization Coverage and Reasons for its Failure for Children in Bida Emirate Area, Niger State, Nigeria. *Merit Research Journal of Medicine and Medical Sciences;* 3(10), 476-483

[10] Jani JV, De Schacht C, Jani IV, Bjune G. 2008. Risk factors for incomplete vaccination and missed opportunity for immunization in rural Mozambique. BMC Public Health. 05 16;8(1):161. http://dx.doi.org/10.1186/1471-2458-8-161

[11] Mitchell S, Andersson N, Ansari NM, Omer K, Soberanis JL, Cockcroft A. 2009. Equity and vaccine uptake: a cross-sectional study of measles vaccination in Lasbela district, Pakistan. *BMC Int Health Hum Rights*. 10 14;9 Suppl 1:S7. http://dx.doi.org/10.1186/1472-698X-9-S1-S7

[12] https//www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter11.pdf

[13] Gordon SP, Gordon FS, Tucker AC, Siegel MJ (2004) functioning in the realworld: A precalculus experience. New York: Pearson Education.

[14] Moody VR, DuCloux KK. 2014. Application of Markov Chains to Analyze and Predict the Mathematical Achievement Gap between African American and White American Students. J Appl Computat Math 3: 161 doi:10.4172/2168-9679.1000161

[15] http://galton.uchicago.edu/~lalley/Courses/312/MarkovChains.pdf

[16] Gambrah PP, Adzadu Y, 2013. Using Markov Chain To Predict The Probability Of Rural And Urban Child Mortality Rates Reduction In Ghana. *International Journal of Scientific & Technology Research,* Vol 2 Issue 11, 73-78

[17] Application to Markov Chains: Do you Like to See the Future? http://aix1.uottawa.ca/~jkhoury/ markov.htm

[18] Billingsley P. 1960. Statistical Methods in Markov Chains. Santa Monica, CA: RAND Corporation.

EMPIRICAL EVIDENCE(S) OF HUMAN CAPITAL INVESTMENTS AND NATIONAL WELFARE IN EU COUNTRIES

Snježana Pivac, Željana Aljinović Barać and Ivana Tadić
Faculty of Economics Split, University of Split
Cvite Fiskovića 5, 21000 Split, Croatia
snjezana.pivac@efst.hr, zbarac@efst.hr, itadic@efst.hr

Abstract: The paper explores the contribution of human capital investments in human capital intensive industries to the national welfare through the growth of gross domestic product in 28 European Union countries during the period of five years. It assumes that countries with greater level of human capital investments are wealthier, i.e. securing sustainable national growth. Research is provided through the sample of approximately 16,380 company-year observations for the 2011-2015 period. The research is done and the results are presented using appropriate panel data analysis technique. The main findings show that human capital investments have significance influence on the national welfare in EU countries.

Keywords: Human capital investments, EU countries, Panel analysis

1 INTRODUCTION

In last decades human resources (HR), and their value recognised through human capital (HC), have arisen to the most important organisational resource and capital. HR are valuable, exceptional and very difficult to imitate according to its specificities as knowledge, special experience, skills, abilities or emotional intelligence [4]. At the other hand, HC can be described as a set of knowledge, skills and abilities that one possesses and which creates and adds economic value to an individual, organisation or nation [10]. Finally, it is a stock of competences and knowledge gained through education and experience required for doing work and producing economic value [16].

Furthermore, human capital investments (HCI) usually are recognised as investments in employees through their additional education, training or workshops developing employees level of knowledge, skills or abilities, which will in future result in securing exceptional results on micro, but also on macro level. In accordance to this, the main issue of this paper is to elaborate the importance and linkage of HC and wider interests, recognised through national welfare. Till nowadays this topic was of great interest of many researchers investigating the relationship between HCI (measured through different educational characteristics, educational or training expenses or employees' income differences) and unemployment as independent variables and national welfare, usually gross national income (GNI) or gross domestic product (GDP) as dependent variables.

2 LITERATURE REVIEW

Since the middle of 20th century scientists and economists were arguing about the inequality in employees' incomes. The level of one's salary depends on investments in employees' training and education, but also personal working experience and promotions. Simply it can be stated it depends on HCI [10]. There are numerous studies proving that different educational types and stage efforts lead to different earnings between workers.

Many authors have been researching importance of HCI and its relation to economic growth, particularly to gross domestic product (GDP), gross national income (GNI) or gross national products (GNP). In this context it can be stated that development of human potential within particular countries contributes to its national development, differentiating poor from

rich countries. Despite, many south-eastern European countries are still lacking in human resource investments, investing more in physical capital instead of human. This is the reason why these countries are falling behind the most developed countries in the world [11]. The former authors researched the influence of HCI on GNP within 177 countries, whereas for HCI they chose three specific different indexes; the Human Development Index, the Gender Related Development Index and the Index of Press Freedom. It is important to mention that the results of statistical analysis on the sensitivity of variables for different countries show the highest influence of Human Development Index on GNP. A research provided within 28 EU countries explored contribution of HCI (measured using individual annual salary, perceived as higher according to greater educational investments) in HC intensive industries to the national welfare. Analysis confirmed that EU countries with high average cost of employee have higher GDP and GNI as well as lower unemployment rate [10]. Reviewing the literature, it is obvious that developing countries still have not received adequate attention as developed ones, while the main reason can be found in their different requirements of educational levels. Developed countries require higher level of education because those are technologically advanced, while certain nations can gain efficiency through imitation, which requires lower level of education. Bangladesh is an example of developing country, where was conducted research in order to reveal the relationship between HC stock and real GDP per capita in the period 1973-2004. Theoretically, HC stock is an important determinant of national income [1]. Findings of the research show that there is long term relationship between HC stock and real GDP per capita in Bangladesh. Also, there are authors who observe national welfare through the level of unemployment and try to reveal relations between HCI and unemployment rates. Research conducted in Germany and USA provided interesting results regarding HCI, investigating precisely short-term and long-term job-search oriented training programs [9]. Findings stress out that participation in short-term training reduces the remaining time in unemployment and increases job stability. Also, long-term training programs initially prolong the remaining time in unemployment, but once the program ends, participants exit to employment at faster rate than without training. Finally, those participants benefit from substantially more stable employment spells and higher earnings.

3 SAMPLE SELECTION AND VARIABLE DESCRIPTION

Based on theoretical background discussed above, research hypothesis is that countries with greater level of HCI are wealthier, indicating that investments in employees' knowledge, skills and abilities through salaries contribute to up growth of country's economic strength.

The data for the research are obtained from Bureau Van Dijk Amadeus database [6] and from World Bank's World Development Indicators database [18]. Annual financial reports of 19.3 million active companies were reviewed and companies were selected for the sample according to following criteria: (1) All legal entities paying the profit tax in the year 2015; (2) Company's main activity is human capital intensive industry, i.e. divisions 72 and 73 of NAICS 2012 classification (primary codes); (3) Company is located in one of the 28 EU countries; (4) Company's data are available for the five-year period (2011-2015).

Companies with missing or incomplete data are excluded. In this way, a relatively homogenous sample of 3,276 companies per year is provided. Companies from 22 EU countries are in the sample, because Greece, Cyprus, Lithuania, Luxembourg, Malta and Romania have no available data. Furthermore, to homogenize the data and the characteristic of the countries the sample is divided at subsample of 9 post-transition countries (Bulgaria, Czech Republic, Estonia, Croatia, Hungary, Latvia, Poland, Slovenia, Slovakia) and subsample of 13 developed countries (Austria, Belgium, Denmark, Germany, Spain, Finland,

France, United Kingdom, Ireland, Italy, Netherlands, Portugal, Sweden). Variable GDP per capita annual growth in percentage (GDPAG) quantifies national wealth based on location. GNI per capita annual growth in percentage (GNIAG) is the measures of national wealth based on ownership. Bough measures are defined as in World Bank methodology [18]. According to [8], the intensive use of HC accounts for increased productivity and technological growth that stimulates economic growth in terms of growth in GDP. Variable unemployment rate is expressed as percentage of labour force and it is state as recorded official unemployment for each country.

Variable that measure HCI is average cost of employee (AVCOSTE) and it is defined as annual salary divided by the number of employees for each company in each year. The individual annual salary is taken into consideration as individual HCI contributing to overall organisational success and national welfare, similar to the methodology in relevant previous research [3], [13], [15], [10] and taking into account the limitations of the available data. Also, variable profit per employee defined as profit before tax by the number of employees for each company in each year is used as proxy of organisational success.

4 RESEARCH RESULTS AND DISCUSSION

The data that include both crossection and time periods components of the analyzed variables are called the panel data, and the process is called the panel analysis. Dependent variables, approximated with GDP per capita growth and GNI per capita growth, are changed to the units of observation (by company) and by the time, so the evaluation of variables, which really determine the variables, are considered a more precise [2].

By observing average indicators for all companies in developed EU countries per years (figure 1) it has been noted that both national welfare variables have similar trends for all the periods, while the average cost of employee from 2011 to 2014 and profit per employee for all the periods show inversely proportional movements in comparison to national welfare measures.

Source: *According to data* http://www.amadeus.bvdinfo.com/; http://www.worldbank.org/

Source: *According to data* http://www.amadeus.bvdinfo.com/; http://www.worldbank.org/
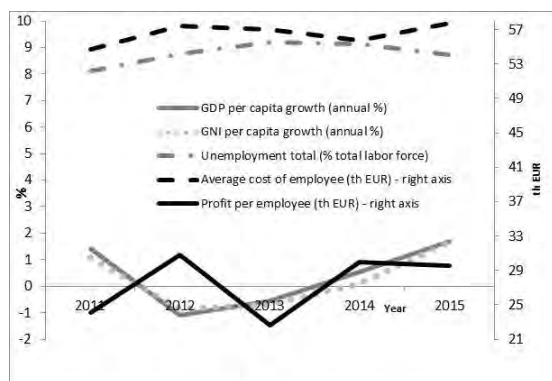
Figure 1: GDP per capita growth, GNI per capita growth, unemployment, average cost of employee and profit per employee of developed EU countries
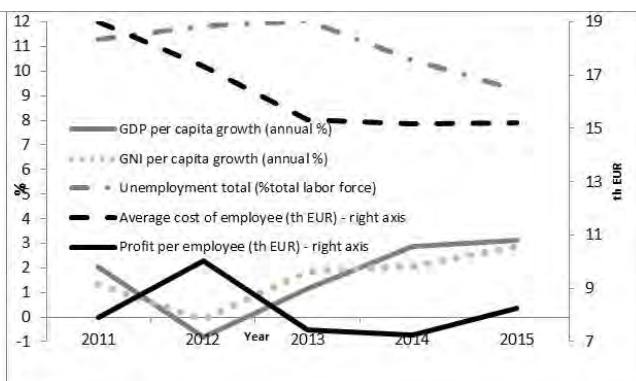
Figure 2: GDP per capita growth, GNI per capita growth, unemployment, average cost of employee and profit per employee of post-transition EU countries

The unemployment rate has inversely proportional trend in comparison with average cost of employee in 2015. For post-transition countries (figure 2) both GDPAG and GNIAG have similar trends for all the periods, while starting from 2012 the AVCOSTE shows inversely proportional movements in comparison with national welfare variables. The unemployment rate has inversely proportional trend in comparison with average cost of employee.

Comparing the data in the figure 1 and the figure 2 it can be concluded that AVCOSTE in the post-transition countries is on a much lower level (15-19 th EUR) by comparing it with developed EU countries (54-58 th EUR). The unemployment rate is on a lower level and the profit per employee is on a much higher level in the developed EU countries (22-30 th EUR) in comparison with the post-transition countries (7-10 th EUR).

Descriptive statistics of observed variables are shown in table 1. It can be seen that according to the mean values, the variation of the AVCOSTE in post-transition EU countries is much higher than in the developed EU countries [11]. The mean level and the variation of the profit per employee in developed EU countries are much higher than in the post-transition EU countries. Correlation coefficients between all independent variables were calculated. Their absolute values are lower than 0.20 which do not indicate the presence of a problem of multicollinearity in the models.

Table 1: Descriptive statistics of observed variables

Variables	Developed EU Countries					Post - transition EU countries				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
GDPAG	7,899	0.813	1.991	-3.638	7.910	8,481	1.579	1.815	-2.922	8.163
GNIAG	7,899	0.794	2.203	-4,131	8.126	8,481	1.531	1.740	-7.105	7.022
AVCOSTE	7,899	56.572	46.074	0,001	846.57	8,481	16.221	23.910	0.007	770.981
PRPEM	7,899	27.518	135.79	-89.036	6254.26	8,481	8.162	35.431	-90.921	1487.27
UNMP	7,899	8.782	1.807	4.100	16.707	8,481	10.981	2.897	5.046	26.300

Source: According to data http://www.amadeus.bvdinfo.com/; http://www.worldbank.org/

A further step in the empirical analysis is an estimate of the adequate panel data model. Selecting the non-adequate estimator can lead to dissimilar conclusions, so very often during the research authors use models estimations with more estimators [17]. After the implementation of the relevant tests for dynamic panel BB estimator two-step model with and without the robust option the models didn't satisfied the residuals autocorrelation conditions [5], [12]. According to relevant tests, the introduction of the random effect for each company was not justified and the model with a fixed effect was not suitable [14]. Between effects models for both post-transition and developed EU countries were estimated. The estimator of the between effects static panel model is unbiased and consistent. In this model, the influence of time component is lost due to the calculation of average values for each observation unit. This indicates that in this situation the multiple regression model estimate could be relevant. The following models were estimated:

$$\overline{GDPAG_i} = \mu + \beta_1\overline{AVCOSTE_i} + \beta_2\overline{PRPEM_i} + \beta_3\overline{UNMP_i} + \overline{\varepsilon_i}; \ i = 1,2,\dots,N \qquad (1)$$

where $\overline{GDPAG_i}$ are average values of dependent variables of a company i, (the appropriate models with $\overline{GNIAG_i}$ as dependent variable is estimated too), N is the number of units of observation. The independent variables are the relevant average values for each company i: $\overline{AVCOSTE_i}, \overline{PRPEM_i}, \overline{UNMP_i},$; μ is a constant; $\beta_1, \beta_2, \beta_3$ are parameters of the independent variables, $\overline{\varepsilon_i}$ is average value of error of a relation of a company i.

The results are shown in table 2. F-tests results show that all the models are significant at 0.01 level. In all the models significant and negative parameters are for the variables AVCOSTE and unemployment. Absolute values of the parameters show that the influence of AVCOSTE on the national welfare is higher in the post-transition countries and the influence of the unemployment rate on the national welfare is higher in the developed countries. It can be concluded about positive and significant influence of the profit per employee on the GDP per capita growth and the GNI per capita growth in post-transition EU countries, while there is no significant influence of the profit per employee on the national welfare in developed EU countries.

Variables	Developed EU countries (STAT-be)		Post-transition EU countries (STAT-be)	
	$GDPAG_{it}$	$GNIAG_{it}$	$GDPAG_{it}$	$GNIAG_{it}$
$AVCOSTE_{it}$	-0.0033**	-0.0035**	-0.0098**	-0.0102**
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$PRPEM_{it}$	-0.0001	-0.0001	0.019*	0.0018**
	(0.611)	(0.608)	(0.010)	(0.004)
$UNMP_{it}$	-0.4696**	-0.3911**	-0.1054**	-0.0743**
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
μ (cons)	4,6881**	3,8653**	3.0720**	2,6779**
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
N	7,899	7,899	8,481	8,481
R^2	0.4981	0.4340	0.4152	0.3457
F test (p-value)	<0.001	<0.001	<0.001	<0.001

Note: be-between estimator

Significances in parentheses

$* p < 0.05$, $** p < 0.01$

Source: Author's estimation according to data http://www.amadeus.bvdinfo.com/; http://www.worldbank.org/

5 CONCLUSION

In this paper it was investigated whether countries with greater level of HCI are wealthier. AVCOSTE in the post-transition countries are on a much lower level by comparing it with developed EU countries. As it was already stated within the literature review, developing countries (in this case post-transition countries) still do not provide sufficient investments in regard to HC in accordance to physical capital. Although, relationships between AVCOSTE as independent variable and GDP and GNI as dependent variables are weak, those are at the same time statistically significant and negative, meaning that lower AVCOSTE will cause higher level of GDP and GNI. This direction can be explained due to huge investments in HC that are undertaken nowadays. Namely, when companies are investing in HC, employees gain considerable level of knowledge and are capable to design, create, maintain and develop new technologies. Also, technology (in many different working positions) replaces employees themselves, what in change requires lower investments in their salaries and further investments, causing lower average cost of employee at the end. Although, profit per employee is at lower level in post-transitional countries than in developed countries, its influence is statistically significant in accordance to GDP and GNI. Post-transitional countries still provide lower HCI, but their each effort in HCI (because of their low starting point) terminates in greater profit per employees and finally higher level of GDP and GNI. Relationship between unemployment rates as independent variable and those dependent ones are also statistically significant and negative.

According to theoretic approach these relations were expected, because higher employment rate means higher employees efforts, energy and work activities (provided through HCI) which is recognised as small, but significant engine in creating economic growth and national welfare.

The main limitation of this study is the lack of the entry data concerning other relevant elements of the HCI. In order to create HCI variable, in majority cases, researches use employee education expenses (investments related to regular and additional education or any other type of individual or organisational education). Additional limitation can been recognised in a small number of different variables used, as well as short period of observation in regards to panel data analysis.

For further researches it is recommended to make panel data analysis about HCI and national welfare for each EU country separately. In this way conclusions will be more

concrete and precise, revealing the contribution of HCI in HC intensive industries to the national welfare.

References

[1] Ahmad, N. and French, J. J. (2011). Decomposing the relationship between human capital and GDP: an empirical analysis of Bangladesh. The Journal of Developing Areas, 44(2):127-142.

[2] Baltagi, B. H. (2008). Econometric analysis of panel data. 4th ed. Chester: John Wiley and Sons.

[3] Barcons-Vilardell, C, et al. (1999). Human Resources Accounting. International Advances in Economics Research, 5(3): 386-394.

[4] Belak, V., Aljinović Barać, Ž. and Tadić, I. (2009). Recognition and measurement of human capital expenditures – impacts on company's performance measurement. International Journal of Economics and Business Research, 1(2): 252-262.

[5] Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1): 115-143.

[6] Bureau Van Dijk (2016). Amadeus – a database of comparable financial information for public and private companies across Europe. Available at http://www.amadeus.bvdinfo.com [10 January 2016].

[7] Hadar, S. K. and Mallik, G. (2010). Does Human Capital Cause Economic Growth? A Case Study of India. International Journal of Economic Sciences and Applied Research, 3(1): 7-25.

[8] Ogunade, A. O. (2011). Human Capital Investments in the Developing World: An Analysis of Praxis. Kingston: University of Rode Island.

[9] Osikominu, A. (2013). Quick Job Entry or Long-Term Human Capital Development? The Dynamic Effecs of Alternative Training Schemes. Review of Economic Studies, 80: 313-342.

[10] Pivac, S., Aljinović Barać, Ž. and Tadić, I. (2016). Multivariate analysis of human capital investments and national welfare in EU countries. Proceedings of the ISCCRO – International Statistical Conference in Croatia, 1(1): 130-137.

[11] Požega, Ž. and Crnković, B. (2008). Impact of Human Capital on GNP Level. South East European Journal of Economics and Business, 3(1): 15-21.

[12] Stata, Data Analysis and Statistical Software. (2016). Dynamic panel-data (DPD) analysis. Available at: http://www.stata.com/features/overview/dynamic-panel-data/ [11 August 2016].

[13] Stiles, P. and Kulvisaechana, S. (2003). Human capital end performance: A literature review. Judge Institute of Management Paper, University of Cambridge.

[14] Škrabić Perić, B. (2012). Static panel data models: Case study of financial development in central and eastern European countries. In Aljinović Z. and Marasović B. (Eds.). Matematički modeli u analizi razvoja hrvatskog financijskog tržišta (173-199). Split: Faculty of Economics, University of Split.

[15] Tadić, I., Aljinović Barać, Ž. and Plazonić, N. (2015). Relations between Human Capital Investments and Business Excellence in Croatian Companies. International Journal of Social, Education, Economics and Management Engineering, 9(3): 745-750.

[16] Tadić, I. (2010). Human Capital Practices in Different Industries in Croatia. The Business Review Cambridge, 15(2): 239-246.

[17] Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge: MIT.

[18] World Bank (2016). World Development Indicators database. Available at: http://www.data.worldbank.org [10 January 2017].

AN OCCUPATIONAL RISK ASSESMENT METHOD FOR FIBER OPTIC CABLE INSTALLATION

Sukran Seker
Yildiz Technical University, Department of Industrial Engineering
Barbaros Street 34349 Istanbul, Turkey
E-mail: seker.sukran@gmail.com

Abstract: Occupational risk is assumed as the possibility of a worker suffering particular work-related injury. People in certain occupations or settings may face increased exposure to health hazards. Many people who install or maintain fiber optic cables do not take proper safety precautions to avoid the many hazards that can be caused by fiber optics. In this study a risk assessment approach is proposed for fiber optic cable installation. The proposed approach provides a quantitative framework for analyzing serious hazards and safety rules for fiber optic cable installation that can keep workers healthy and the work environment safe for all employees.

Keywords: Occupational risks, Fiber optic cable installation, Risk assessment.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 3:

Finance and Investments

DOES CVaR OVERCOME VaR ON THE CROATIAN STOCK MARKET

Zdravka Aljinović
University of Split, Faculty of Economics
Cvite Fiskovića 5, 21000 Split, Croatia
E-mail: zdravka.aljinovic@efst.hr

Andrea Trgo
KentBank d.d., Corporate banking department
Poljička cesta 26, 21000 Split, Croatia
E-mail: andreatrgo@gmail.com

Abstract: In the paper two well-known risk measurement methods Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) are applied on the Croatian stock market. The methods together with appropriate backtesting are applied on the sample of 29 stocks grouped into 8 sectors for the three different periods: 2006-2007, the period characterized with economic growth, the crisis period 2008-2009 and the post-crisis period 2013-2014, characterized with long-term economic stagnation in Croatia. The research confirms CVaR as the valid and appropriate risk measure for the Croatian stock market and gives the significant primacy to CVaR over VaR. Also, the insight to the riskiness of a particular sector through different periods is given.

Keywords: Conditional Value-at-Risk (CVaR), Value-at-Risk (VaR), Croatian stock market

1 INTRODUCTION ON VaR AND CVaR FEATURES

Last decades a lot of disturbances on financial markets all over the world, sometimes very strong, put into the focus of researchers and practitioners the market risk measurement and management. From the day of appearance within JP Morgan Bank, the risk measure Value-at-Risk (VaR) attracts a huge attention and with time it became one of the most controversial financial instruments, being at the same time very much criticised and very much in use [7]. VaR, defined as a statistical measure which assesses the risk of some asset or the whole portfolio expressed with one number – the worst estimated loss for a certain time horizon and a certain confidence level, due to its simplicity, applicability and universality, became very popular and widely used risk measure. Basel Committee for Banking Supervision in Basel II standards put VaR as the key factor in risk management and VaR became industrial standard for market risks measurement [4].

If the distribution of prices of an asset is given at the certain moment *t*, VaR represents the difference between the invested amount of money and the value which is not going to be failed in α% cases – the value which corresponds to the 1- α percentile of the distribution. One of the most common critics of the VaR says that with VaR we don't have any information about the values from the tail of the distribution – the values which exceed the value of VaR. Information about unexpected events important for the firm (small probability, high losses) is not included in VaR [11]. The risk measure which gives information about losses from the tail of distribution which exceed VaR is Conditional Value-at-Risk (CVaR). Actually, for a given time horizon and confidence level α, CVaR is defined as the conditional expectation of losses greater than VaR.

Maybe the biggest objection on the VaR is that it is not the coherent risk measure because it does not fulfil the subadditivity condition: $\rho(X + Y) \leq \rho(X) + \rho(Y)$, i.e. VaR of a portfolio is greater than the sum of VaRs of constituents of a portfolio [3], [5]. This can discourage a portfolio diversification and lead to the dangerous risk concentration. It can happen that well diversified portfolio requires more of regulatory capital than the worse diversified one. In the

paper [3] the concept of a coherent risk measure is presented and authors had shown that among three conventional risk measures: VaR, variance and CVaR, only CVaR is a coherent risk measure. This good features of CVaR, along with some other, are presented also in [1], [9], [12] and [13].

We can say that CVaR has superior mathematical characteristics over VaR; CVaR keeps good properties of VaR and overcomes its shortages. The main goal of the paper is to see how this two measures function on the Croatian capital market – if the theoretical dominance of CvaR is confirmed also in practice. The analysis is done for 29 stocks grouped into 8 sectors, for the three different periods, pre-crisis, crisis and post-crisis. This way, beside the insight of the CVaR and VaR performances we also have a picture about riskiness of a particular sector from the market. Knowing and understanding sectors' risk is very important for all investors and participants on the market, primarily from the aspect of possible adversely risk concentration in particular sectors. Identifying sectoral overconcentration is essential to managing portfolio credit risk [2].

The paper is organised as follows: this insight into the VaR and CVaR features is followed by the stocks sample presentation, together with the methodology and VaR and CVaR calculations and backtestings. The paper is concluded with the results together with the short analysis and conclusion.

2 DATA AND METHODOLOGY

2.1 Data

The first criterion for choosing the stocks sample was liquidity, since calculations are based on daily trading data and for precise and correct results the active trading is a prerequisite. According to this criterion, 41 stocks included in the Crobex Plus Index on the 14 August 2015 together with 12 stocks with turnover greater than 10 mil kn on the Zagreb Stock Exchange (ZSE) for 2014, are taken. Further, the necessity of having a good basis of trading data for all of the three observed periods and appearance of a stock splits for two stocks – meaning appearance of extreme values and consequently wrong calculations of sector's risk, lead us to the sample of 30 stocks divided into 9 sectors. Since in the Telecommunication sector only one stock remained, ERNT-R-A, it is decided to proceed without this one-stock sector. Finally, the sample for the analysis is consisting of 29 stocks from 8 sectors given in the Table 1. The sector division follows [8], due to the rather specific conditions on the Croatian capital market.

Table 1: The stocks sample divided into sectors

Sector	Stocks
Diversified (2)	ADRS-P-A, SNHO-R-A
Agriculture (2)	BLJE-R-A, CKML-R-A
Financials (2)	CROS-R-A, PBZ-R-A
Industry (4)	KOEI-R-A, ADPL-R-A, DDJH-R-A, PTKM-R-A
Construction (4)	DLKV-R-A, IGH-R-A, THNK-R-A, VKDT-R-A
Food and Staples Retailing (4)	KRAS-R-A, PODR-R-A, LEDO-R-A, ZVZD-R-A
Hotel-Management and Tourism (5)	ARNT-A, LRH-R-A, HUPZ-R-A, MAIS-R-A, TUHO-R-A
Transportation (6)	ATPL-R-A, LKPC-R-A, JDPL-R-A, LKRI-R-A, ULPL-R-A, LPLH-R-A

Daily trading prices for chosen stocks are taken for three periods. The first one is the pre-crisis period, 2006-2007, which is characterized with economic growth in Croatia. The second one from 2008 to 2009, is the period of a great financial crisis, where 2008 is characterized with huge fall in all activities on the ZSE followed with significant economic decline in 2009. The third, post-crisis period from 2013 to 2014 is a part of a long term stagnation period in Croatia. The number of trading dates for observed periods is 499, 502 and 499, respectively.

2.2 Methodology and Calculations

For VaR and CVaR calculations, the historical method is applied. It is very often used, and probably the simplest non-parametric method. The common characteristic of all non-parametric methods is that they use empirical distributions, while parametric methods assume a certain theoretical distribution. It is assumed that the trend of the latest changes of prices will be continued also in the future. Then the historical data is used for the risk evaluation in the near future. VaR and CVaR are calculated as percentiles of the empirical distribution, according to the chosen confidence level. There is no need for approximation of the distribution's parameters like volatility and correlation coefficients; the method is easy for implementation, only the historical rates of return are needed. It allows to describe non-normal distributions with fat tails which are rather often present among financial data [6]. All this make the historical method very suitable for implementation on the Croatian stock market.

For the VaR and CVaR calculations using historical method we follow the next steps:
- Collection of the stocks' prices classified into sectors for each trading day and each period;
- Calculation of the stocks' daily rates of return;
- Calculation of the sectors' daily rates of return: For example, for the Industry sector that means calculation of the rates of return of the portfolio consisting of the four stocks, all with the same proportion of 25%.
- Calculation of potential losses and/or gains for all stocks and appropriate sectors: For example, for the Industry sector we assume that we have an investment of 1 million kunas, 250.000 kn per each stock. By multiplying the rates of return with the appropriate investment, we have results – distribution of gains and losses for each stock and sector.
- Sorting of gains/losses results from the highest to the lowest value and calculation of VaR and CVaR; For example, for the Industry sector for the 2006-2007 period, from 499 sorted results for the confidence level of 95%, we take 5%(499)=24,95, i.e. 25 worst results, read the value of VaR and calculate the value of CVaR as the average of 25 worst results. For the Industry sector we got $VaR_{95\%}$=-21.158,12 and $CVaR_{95\%}$=-30.866,96.

Evaluation of the VaR and CVaR risk approximations is done through so called backtesting. The Basel Committee defines backtesting as the ex-post comparison of a model's risk approximations with the real daily changes of a portfolio's value through longer periods, or with the hypothetical changes based on the static positions [4]. So, the approximated values of VaR and CVaR should be compared with the real losses at the end of an observed period to see if there are cases where the portfolio suffered higher losses than those predicted by the model. Usually, backtesting inquires the frequency of exceedings of the approximated VaR. This failing rate should be in accordance with the applied confidence level. For example, if the daily VaR is approximated with the confidence level of 95%, the maximum of overruns for the year with 250 trading days is 5%(250) = 12,5. CVaR backtesting is more complex than the VaR backtesting and there is an opinion that it is one of the reasons why CvaR is not included in the Basel Committee's framework.

For the CVaR backtesting, the loss function ρ which compares the approximated value of CVaR with the real return r, in cases where r exceeds VaR, is settled [10]:

$$\rho = \begin{cases} r \; if \; r < VaR \\ 0 \; if \; r \geq VaR \end{cases} \tag{1}$$

The function ρ gives to each loos observation from the tail the weight 1, where the appropriate referent value is simply CVaR.

For example, for the Industry sector for the 2006-2007 period, backtesting is done for 80 portfolio's returns from 13 Oct 2007 to 31 Dec 2007. Each portfolio's return is compared with the sector's VaR calculated for the observed period with confidence level of 95%, and in this case that is -21.158,12. In the case when return is greater than or equal to the VaR value, the value 0 is given to the observation; otherwise its value is recorded. From 80 returns 5 of them exceed VaR, and from those five the mean is calculated; the value of the mean is -29.516,15. That value is compared with the referent value – CVaR calculated for the observed period with confidence level of 95% and in this case that is -30.866,96. Since 29.516,15<30.866,96 i.e. average of real losses is less than CVaR, it can be concluded that the risk measurement model for the 2006-2007 period is representative, applicable for the Industry sector on the Croatian stock market with precise risk evaluation.

For the same sample of the last 80 returns from the 2006-2007 period, the VaR backtesting is done. For the model acceptance, the maximum number of exceeds of VaR should be 5%(80)=4. Results of the VaR backtesting for the Industry sector, for the 2006-2007 period are presented in Figure 1.
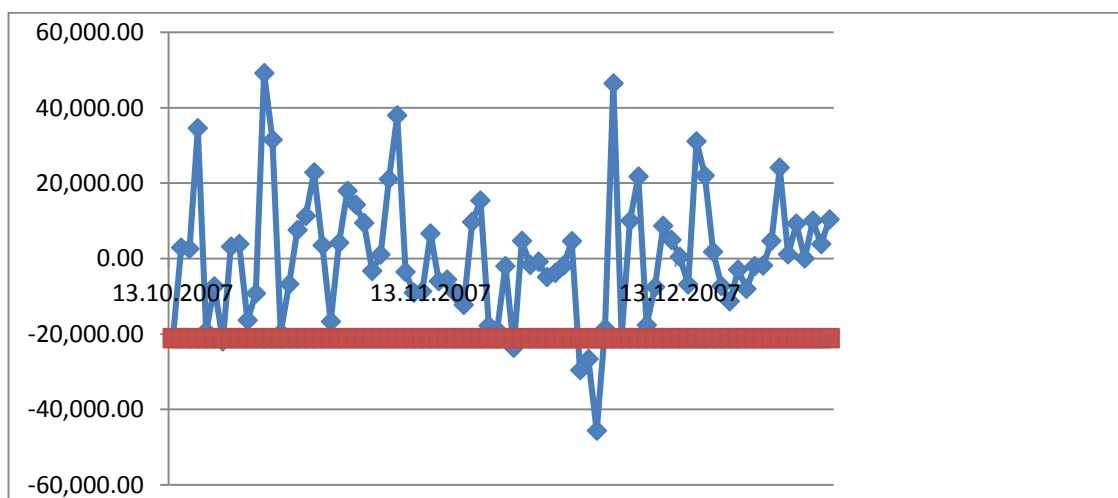
Figure 1: Results of the VaR backtesting fot the Industry sector for the 2006-2007 period, confidence level 95%

From Figure 1 we can see that the number of the VaR value overruns is 5, precisely for the dates 3 Dec, 2 Dec, 1 Dec, 23 Nov and 19 Oct 2007. Since the number of overruns is higher than the allowed maximum, 4, we cannot say that historical method for VaR evaluation for the Industry sector on the Croatian stock market for 2006-2007 period gives precise and reliable result.

3 RESULTS

Following the previously described methodology, CVaR and VaR calculations together with backtestings are done for all eight sectors and for all three observed periods. Backtesting results are given in Tables 2 and 3.

Sector / Period	2006-2007		2008-2009		2013-2014	
	Average value ρ	CVaR 95%	Average value ρ	CVaR 95%	Average value ρ	CVaR 95%
Hotel-Management and Tourism	-23.248,84	-22.080,46	0,00	-47.510,34	-19.786,28	-19.280,08
Food and Staples Retailing	-22.716,94	-25.058,82	-41.904,94	-55.066,54	-12.530,63	-15.122,48
Financials	-26.347,46	-27.429,24	0,00	-72.650,05	-21.707,20	-25.170,92
Construction	-42.258,14	-32.271,37	0,00	-83.276,05	0,00	-44.780,10
Transportation	-36.927,04	-31.719,01	0,00	-84.045,53	-22.834,10	-24.599,64
Industry	-29.516,15	-30.866,96	0,00	-64.010,66	-38.293,14	-36.846,59
Diversified	-23.719,19	-31.756,28	-53.285,33	-70.095,85	-21.224,85	-20.475,90
Agriculture	-26.623,11	-35.064,61	-51.826,11	-64.532,59	-35.397,25	-24.991,52

We can see that from the total number of 24 CVaR risk approximations on the Croatian stock market for eight sectors through three periods, in 18 cases CVaR is confirmed as representative and applicable. For the crisis period, 2008-2009, CVaR risk approximations were correct for all sectors. Equally good results for 2008-2009 period has also shown VaR backesting, but for other two periods results are worse, especially for the 2006-2007 period, as Table 3 shows. Cumulatively, VaR is confirmed as representative and applicable risk measure in 13 from 24 cases.

Table 3: VaR backtesting results

Sector / Period	2006-2007		2008-2009		2013-2014	
	Allowed overruns maximum	Real overruns number	Allowed overruns maximum	Real overruns number	Allowed overruns maximum	Real overruns number
Hotel-Management and Tourism	4	8	4	0	4	8
Food and Staples Retailing	4	9	4	1	4	1
Financials	4	9	4	0	4	4
Construction	4	8	4	0	4	0
Transportation	4	10	4	0	4	5
Industry	4	5	4	0	4	6
Diversified	4	5	4	2	4	10
Agriculture	4	4	4	3	4	4

In 2006-2007 period, VaR failed in all sectors, except in Agriculture sector where the border number of overruns is recorded. That is the single sector where VaR actually works well, regardless the period. CVaR is confirmed as absolutely representative and appropriate risk measure for sectors of Food and Staples Retailing and Financials, offering precise risk measurement for all observed periods. The riskiness of a particular sector varies significantly regarding the period; as it was expected, those sectors which are risky during the crisis are not so risky in pre or post crisis periods. But, there are some "rules" regarding riskiness of sectors

on the Croatian stock market: the least risky sectors regardless of the period are Hotel-Management and Tourism and Food and Staples Retailing, while the Construction sector is always the riskiest one. A further analysis can give a lot of interesting and useful insides and conclusions regarding sectors' risk on the Croatian stock market.

4 CONCLUSION

For now, we have an answer to the main question in the paper: CVaR overcomes VaR in adequate identification of risky sectors on the Croatian stock market.

Not only that CVaR has primacy over VaR in theory since it quantifies tail risk and fulfils the subadditivity condition; it is confirmed that CVaR simply works better, gives more reliable and precise results. Additionally, if we take into consideration that it is equally simple for calculation as VaR, there is really no obstacle for taking Conditional Value-at-Risk as a representative risk measure on the Croatian stock market.

References
[1] Acerbi, C., Tasche, D. 2002. On the coherence of expected shortfall. Journal of Banking&Finance, 26(7): 1487–1503.

[2] Allen, D.E., Kramadibrata, A.R., Powell, R.J., Singh, A.K. 2012. Identifying European Industries with Extreme Default Risk: Application of CVaR Techniques to Transition Matrices. World Review of Business Research, 2(6): 46-58.

[3] Artzner, P., Delbaen, F., Eber, J.M., Heath, D. 1999. Coherent Measures of Risk, Mathematical Finance, 9(3): 203-228.

[4] Basel Committee on Banking Supervision. 2004. International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Bank for International Settlements. http://www.bis.org/publ/bcbs107.pdf [Accessed 30/5/2017].

[5] Dowd, K. 2002. Measuring market risk. Chichester, New York: John Wiley and Sons

[6] Van den Goorbergh, R.W.J., Vlaar, P. 1999. Value-at-Risk Analysis of Stock Returns Historical Simulation, Variance Techniques or Tail Indeks Estimation?. De Nederlandsche Bank Staff Reports, No.40. https://www.dnb.nl/binaries/sr040_tcm46-146818.pdf [Accessed 30/5/2017].

[7] Hafsa, H. 2015. CVaR in Portfolio Optimization: An Essay on the French Market. International Journal of Financial Research, 6(2): 101-111.

[8] Jerončić, M., Aljinović, Z. 2011. Formiranje optimalnog portfelja pomoću Markowitzevog modela uz sektorsku podjelu kompanija. Ekonomski pregled, 62(9-10): 583-606.

[9] Krokhmal, P., Palmquist, J., Uryasev, S. 2002. Portfolio optimization with Conditional Value-at-Risk Objective and Constraints. The Journal of Risk, 4(2): 11-27.

[10] Letmark, M. 2010. Robustness of Conditional Value-at-Risk (CVaR) when measuring market risk across different asset classes. Master's thesis, Royal Institute of Technology, Stockholm. http://arc.hhs.se/download.aspx?MediumId=189 [Accessed 21/8/2017].

[11] Rootzen, H., Kluppelberg, C. 1999. A single number can't hedge against economic catastrophes. Ambio, 28(6): 550-555

[12] Uryasev, S. 2000. Conditional Value-at-Risk (CVaR): Optimization Algorithms and Applications, Financial Engineering News, Issue 14, February 2000: 1-5.

[13] Uryasev, S., Serriano, G., Sarykalin, S. 2008. Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization. Tutorials in Operation Research, INFORMS 2008. http://www.ise.ufl.edu/uryasev/files/2011/11/VaR_vs_CVaR_INFORMS.pdf [Accessed 30/5/2017].

CLASSIFICATIONS IN THE SYSTEM OF NATIONAL ACCOUNTS

Draženka Čizmić
Faculty of Economics & Business, University of Zagreb
Trg J.F. Kennedyja 6, Zagreb 10 000, Croatia
dcizmic@efzg.hr

Abstract: Economic statistics require different classifications for different purposes. Classifications of economic activities are designed to categorise data that can be related only to the unit of activity. Product classifications are designed to categorise products that have common characteristics. Economic statistics are required by different users for various types of analysis. The System of National Accounts (SNA) is a principal user and it has particular requirements.

Keywords: classification of economic activities, classification of products, functional classification

1 INTRODUCTION

In the study of economic phenomena, taking all elements into account simultaneously is not always possible. For the purposes of analysis, certain elements need to be chosen and grouped according to particular characteristics. Thus, all observations that are to be described in terms of statistics require systematic classification. Classifications partition the universe of statistical observations according to sets that are as homogeneous as possible with respect to the characteristics of the object of the statistical survey.

Economic statistics require different classifications for different purposes. Classifications of economic activities are designed to categorise data that can be related only to the unit of activity. Product classifications are designed to categorise products that have common characteristics. Economic statistics are required by different users for various types of analysis. The System of National Accounts (SNA) is a principal user and it has particular requirements.

In the context of a production approach of GDP, tables by industry and the input-output framework, use is made of two classifications: ISIC Rev.4 (NACE Rev.2) for economic activities and CPC Version 2.1 (CPA Version 2.1) for products by economic activities. Also, for an expenditure approach of GDP, CPC Version 2.1 (CPA Version 2.1), COFOG and COICOP are used.

2 CLASSIFICATIONS OF ECONOMIC ACTIVITIES

The International Standard Industrial Classification of All Economic Activities (ISIC) is the international reference classification of productive activities. Since the adoption of the original version of ISIC in 1948, the majority of countries around the world have used ISIC as their national activity classification or have developed national classifications derived from ISIC. The Statistical Commission[1] initiated reviews and revisions of ISIC in 1956, 1965, 1979 and again in 2000.

The fourth revision (ISIC, Rev.4) is the outcome of a review process that spanned several years and involved contributions from many classifications experts and users around the world. The structure of the fourth revision of ISIC was considered and approved by the Statistical Commission in March 2006. The objectives of the fourth revision of ISIC were formulated in terms of improving and strengthening its relevance and comparability with

---

[1] The United Nations Statistical Commission is the highest body of the global statistical system. It is the highest decision making body for international statistical activities especially the setting of statistical standards, the development of concept and methods and their implementation at the national and international level.

other classifications, while considering its continuity. The detail of the classification has substantially increased.[2]

The scope of ISIC in general covers productive activities, i.e., economic activities within the production boundary of the System of National Accounts.[3] These economic activities are subdivided in a hierarchical, four-level structure of mutually exclusive categories. The categories at the highest level are called sections, which are alphabetically coded categories.[4] The classification is then organized into successively more detailed categories, which are numerically coded: two-digit divisions; three-digit groups; and, at the greatest level of detail, four-digit classes.

The classification is used to classify statistical units, such as establishments or enterprises, according to the economic activity in which they mainly engage. At each level of ISIC, each statistical unit is assigned to one and only one ISIC code. The set of statistical units that are classified into the same ISIC category is then often referred to as an industry.

While ISIC was developed with a view to categorizing economic activities for national accounts and other economic analysis purposes, its use extends to data collection, tabulation, analysis and presentation for a variety of social and environmental applications.

ISIC does not draw distinctions according to kind of ownership of a producing unit, type of legal organization or mode of operation because such criteria do not relate to the characteristics of the activity itself. Therefore, a strict link between ISIC and the Classification of Institutional Sectors in SNA does not exist.

NACE is the European standard classification of productive economic activities. NACE is derived from ISIC, in the sense that it is more detailed than ISIC. ISIC and NACE have exactly the same items at the highest levels, where NACE is more detailed at lower levels.

NACE was developed in 1970. This first version of NACE suffered from two major drawbacks: 1) as it had not been established as part of the Community legislation, data were often collected according to the existing national classifications and then transformed into the NACE format by means of conversion keys, which did not produce satisfactorily comparable data; 2) as NACE Rev. 1970 had not been developed within a recognised international framework, it offered poor comparability with other international classifications. A working group promoted by Eurostat with representatives of Member States developed a revised version of NACE, called NACE Rev.1. Starting from the structure of ISIC Rev.3, details were added to reflect European activities that were inadequately represented in ISIC. In 2002, a minor update of NACE Rev.1, called NACE Rev. 1.1, was established. NACE Rev. 1.1 introduced a few additional items and changes to same titles. The Regulation establishing NACE Rev. 2 was adopted in December 2006.

The structure of NACE is described in the NACE Regulation as follows: 1) a first level consisting of headings identified by an alphabetical code (sections), 2) a second level consisting of headings identified by a two-digit numerical code (divisions), 3) a third level consisting of headings identified by a three-digit numerical code (groups) and 4) a fourth level consisting of headings identified by a four-digit numerical code (classes).

National accountants have identified a need for two standard aggregations of ISIC/NACE categories to be used for reporting SNA data from a wide range of countries. The first, known as "high-level aggregation", aggregates the ISIC/NACE sections into 10 or 11 categories; the second, called "intermediate aggregation", aggregates divisions and is

---

[2] For service-producing activities, this increase is visible at all levels, including the top level, while for other activities, such as agriculture, the increase in detail has affected mostly the lower levels of the classification.

[3] A few exceptions have been made to allow for the classification of activities beyond the production boundary but which are of importance for various other types of statistics.

[4] The sections subdivide the entire spectrum of productive activities into broad groupings, such as "Agriculture, forestry and fishing" (section A), "Manufacturing" (section C), "Information and communication" (section J).

composed of 38 categories. The two aggregated structures are not an integral part of ISIC/NACE, but are fully integrated into their hierarchical structure.

Statistics collected by EU Member States involving classification by economic activity must be compiled according NACE or a national classification derived from it. The National Classification of Activities in Croatia, the 2007 version (NKD 2007) is based on the NACE Rev.2. The NKD 2007 up to the class level ensures comparability and exchange of data for the same business entities in the Member States of the European Union and other countries that implement NACE Rev.2, and up to the group level, the NKD 2007 ensures comparability with the ISIC Rev.4.

3 CLASSIFICATIONS OF PRODUCTS

The Central Product Classification (CPC) serves as the reference classification for all product classifications within the international system of economic classifications put in place by the United Nations. The CPC constitutes a complete product classification covering all goods and services.[5] The primary purpose of CPC is to classify the goods and services that are the result of production in any economy. This production is accounted for in the national accounts of countries and can be measured and analysed using the System of National Accounts.

The CPC originated from initiatives in the early 1970s to harmonize international classifications prepared under the auspices of the United Nations and other international bodies. The first version of the CPC, the Provisional Central Product Classification, was published in 1991. This version was superseded by the CPC Version 1.0, published in 1998. In that publication particular attention was paid to the elaboration of the services part of the classification. CPC Version 1.1, published in 2002, represented a further update, incorporating modifications due to recent changes in economies and sustained technological advancement. The CPC Version 2, released in 2008, again reflected recent changes in the character of outputs. The revision process for the CPC Version 2 was strongly tied to the process for the fourth revision of ISIC. The current edition, CPC Version 2.1, is the result of a scheduled review of the CPC structure and detail to ensure the classifications' relevance for describing current products in the economy.

The CPC was developed primarily to enhance harmonization among various fields of economic and related statistics and to strengthen the role of national accounts as an instrument for the coordination of economic statistics. It serves as an international standard for assembling and tabulating all kinds of data requiring product detail, including statistics on industrial production, domestic and foreign commodity trade, international trade in services, balance of payments, consumption and price statistics and other data used within the national accounts.

The overall set of products is subdivided into a hierarchical, five-level structure of mutually exclusive categories, facilitating data collection, presentation and analysis at detailed levels of the economy in an internationally comparable, standardized way. The categories at the highest level are called sections, which are numerically coded categories. The classification is then organized into successively more detailed categories, which are numerically coded: two-digit divisions; three-digit groups; four-digit classes; and, at the greatest level of detail, five-digit subclasses. The CPC classifies products based on the

---

[5] The 2008 System of National Accounts provides a definition of product. It states that goods and services (products) are the result of production; they are exchanged and used for various purposes, such as for inputs in the production of other goods and services, or as final consumption or for investment. In order to study transactions in goods and services in detail, the SNA uses the CPC.

physical properties and the intrinsic nature of the products as well as on the principle of industrial origin.

For national accounting purposes, it may be necessary to classify data at a different level of detail from that required for industrial statistics purposes. The CPC is not an asset classification. Assets are classified according to a separate classification within the System of National Accounts.

In its function as a "central" product classification, the CPC has a natural relationship with all classifications that provide a structure for the classification of products.[6] The CPC as a classification of products has a strong natural relationship with the classification of economic activities, ISIC.[7]

The Harmonized Commodity Description and Coding System (HS) is the international customs product classification drawn up by the World Customs Organisation for foreign trade. HS covers all products which can be the subject of an international transaction and which have a physical dimension.[8] It comprises about 5000 commodity groups, each identified by a six digit code, arranged in a legal and logical structure. The HS Committee prepares amendments updating the HS every 5-6 years.

Each subclass in section 0 to 4 of the CPC is defined as the equivalent of one heading or subheading or the aggregation of several headings or subheadings of the HS. The Harmonised System uses primarily the physical property criterion for classifying goods.

Standard International Trade Classification (SITC) is recommended for analytical purposes only. The classification system is maintained by the United Nations. The SITC is currently at revision four, which was promulgated in 2006. The structure of SITC follows a traditional order in which the main considerations are the materials used, the stage of processing and the end use. The scope of SITC Rev. 4 covers all goods classified in HS except monetary gold, gold coins and current coin. Basic headings are defined in terms of HS subheadings.

The Classification of Products by Activity (CPA) is the European version of the CPC, and the purposes it serves are in line with those of the CPC. Whilst the CPC is merely a recommended classification, the CPA is legally binding in the European Union. The CPA was elaborated in 1993, updated in 1996, 2002 and again in 2012. The Regulation establishing CPA version 2.1 was adopted in October 2014, entering into force 1 January 2015. While some sections of the CPA have been aligned to the CPC version 2.1 and explanatory notes have been reviewed, the overall characteristics of the CPA remain unchanged.

The structure of the CPA is described in the CPA Regulation as follows: a first level, comprising headings identified by an alphabetical code; a second level, comprising headings identified by a two-digit numerical code; a third level, comprising headings identified by a three-digit numerical code; a fourth level, comprising headings identified by four-digit numerical code; a fifth level, comprising headings identified by a five-digit numerical code; and sixth level, comprising headings identified by a six-digit numerical code.

The link between the CPA and NACE Rev. 2 is evident in the CPA code: at all levels of the CPA, the coding of the first four digits is identical with that used in NACE Rev. 2, with very few exceptions.

---

[6] Either of all products or a specific subset, such as transportable goods, services, energy products etc.
[7] Each subclass of the CPC consists of goods or services that are generally produced in a specific class or classes of the ISIC, Rev.4.
[8] HS does not cover services, but does include the physical "manifestations" of services (e.g. architects' plans, diskettes with software, etc.). Although the HS basically covers goods it also encompasses electricity. It also includes goods which are not produced, such as used equipment.

Although the CPA is the European counterpart of the CPC, it differs from the latter not only in that it is usually more detailed, but also as regards its structure. The CPA uses the criterion of economic origin according to the structure of NACE, while the CPC has a specific structure which primarily separates goods and services.

The CPA Regulations allow Member States to use a national version derived from the CPA for national purposes. Such national versions must fit into the structural and hierarchical framework of the CPA. The Classification of Products by Activities in the Republic of Croatia (KPD 2015) has been harmonized with the CPA Version 2.1.

The Combined Nomenclature is the classification used within the EU for the purposes of foreign trade custom tariffs and statistics and provides a degree of detail going beyond that in the HS. The CN was introduced in 1988. Headings in the CN are identified by means of an eight-digit numerical code, adding two digits to the relevant HS code. The CN is revised every year. Member States may insert subdivisions after the CN subheadings for national statistical purposes.

In Croatia since 1 January 2002, the 8-digit code of the Combined Nomenclature has been used in data collection and it is harmonised up to 6-digit code level with the Harmonised system. Since 1 January 2007 goods listed under export and import are classified, for statistical needs, according to the SITC Rev. 4.

PRODCOM is the abbreviation for the EU system of production statistics for mining and manufacturing. The product classification (PRODCOM list) is drawn up each year by the PRODCOM committee. The headings of the PRODCOM list are derived from the CN, but their code is a further breakdown of the CPA code. The PRODCOM list is therefore linked to, and thus consistent with the CPA.

The Croatian Nomenclature of Industrial Products (NIP) and the Nomenclature of Industrial Products for the Monthly Survey (NIPUM) are comparable to the PRODCOM classification.

4 FUNCTIONAL CLASSIFICATIONS

Since the CPC provides the product dimension to many of the SNA tables, the CPC can be related to the Classifications of Expenditure According to Purpose: Classification of the Functions of Government (COFOG), Classification of Individual Consumption According to Purpose (COICOP), Classification of the Purposes of Non-Profit Institutions Serving Households (COPNI) and Classifications of the Outlays of Producers According to Purpose (COPP). They are described as "functional" classifications because they identify the "functions" for which transactors engage in certain transactions.

The Classification of Individual Consumption According to Purpose (COICOP) is used to describe expenditures of private households in national accounts, Household Budget Surveys and the Consumer Price Index. As COICOP is one of basic classifications of the SNA, it follows the concepts and definitions of the SNA. The primary purpose of COICOP is to classify individual consumption of goods and services according to their main purpose. For example, COICOP shows household expenditure on food, health and education services all of which are important indicators of national welfare.

The Classification of the Functions of Government (COFOG) was developed by the former Statistical Office of the United Nations Secretariat mainly for use in the SNA. COFOG is more appropriate than ISIC for classifying government expenditures because the COFOG list of functions is more detailed than the ISIC list of activities. COFOG is designed for classifying current transactions, capital outlays and acquisition of financial assets by general government and its subsectors. COFOG is used to distinguish between collective services and individual consumption goods and services provided by government.

For describing and analysing the expenditure of private non-profit institutions serving households, COPNI is used. This classification is a somewhat reduced version of the classification for all non-profit institutions.

For describing and analysing the behaviour of producers, COPP can be used. COPP may provide information on the "outsourcing" of business services, i.e. the substitution of ancillary activities by purchases of corresponding services from other producers.

7 CONCLUSION

One of the basic requirements for statistical work is the existence of a recognised framework which can accommodate the vast range of statistical data available so that they can be presented and analysed in a meaningful way. Classifications provide that common language for both the compilation and the presentation of statistics.

Changes in economic structures and organisations, as well as technological developments, give rise to new activities and products, which may supersede existing activities and products. Such changes imply a constant challenge for the compilation of statistical classifications. The intervals between revisions must not be too long, since the relevance of the classification diminishes with time, nor must they be too short, since otherwise the comparability of the data over time is adversely affected.

Through the joint efforts of the United Nations and the European Union for the harmonisation of economic classifications, the structure and content of the revised NACE and the related product classifications of the European Union were developed to be consistent with the ISIC and the CPC.

References

[1] Central Product Classification (CPC) Version 2.1. (2015). Department of Economic and Social Affairs. Statistics Division. United Nations. New York. http://unstats.un.org/unsd/cr/downloads. [Accessed 03/10/2016]

[2] Commission Regulation (EU) No 1209/2014 amending Regulation (EC) No 451/2008 of the European Parliament and of the Council establishing a new statistical classification of products by activity (CPA) and repealing Council Regulation (EEC) No 3696/93. (2014). Official Journal of the European Union. http://eur-lex.europa.eu. [Accessed 05/10/2016]

[3] CPA 2008 introductory guidelines. (2008). Eurostat. Luxembourg. http://ec.europa.eu/eurostat/ documents. [Accessed 03/10/2016]

[4] European system of accounts 2010. (2013). Publications Office of the European Union, Luxembourg

[5] Harmonized System. (2016). https://en.wikipedia.org. [Accessed 04/10/2016]

[6] International Standard Industrial Classification of All Economic Activities (ISIC), Rev.4. (2008). Department of Economic and Social Affairs. Statistics Division. United Nations. New York. http://unstats.un.org/unsd/publication. [Accessed 03/10/2016]

[7] Klasifikacija proizvoda po djelatnostima Republike Hrvatske – KPD 2008. (2008). Državni zavod za statistiku Republike Hrvatske. http://narodne-novine.nn.hr. [Accessed 05/10/2016]

[8] Klasifikacija proizvoda po djelatnostima Republike Hrvatske 2015 – KPD 2015. (2014). Državni zavod za statistiku Republike Hrvatske. http://narodne-novine.nn.hr. [Accessed 05/10/2016]

[9] NACE Rev.2. Statistical classification of economic activities in the European Community. (2008). Eurostat. Luxembourg. http://ec.europa.eu/eurostat/documents. [Accessed 03/10/2016]

[10] Metodologija za statističku primjenu Nacionalne klasifikacije djelatnosti 2007. – NKD 2007. (2007). Državni zavod za statistiku RH. http://narodne-novine.nn.hr [Accessed 05/10/2016]

[11] System of National Accounts 2008. (2009). European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank. New York

ZAGREB STOCK EXCHANGE AND THE (A)SYMMETRIC EFFECTS OF NEWS

Mirjana Čižmešija
Faculty of Economics and Business, University of Zagreb, Department of Statistics
Trg J.F. Kennedya 6, 10 000 Zagreb, Croatia
E-mail: <mcizmesija@efzg.hr>

Petar Sorić
Faculty of Economics and Business, University of Zagreb, Department of Statistics
Trg J.F. Kennedya 6, 10 000 Zagreb, Croatia
E-mail: <psoric@efzg.hr>

Marina Matošec
Faculty of Economics and Business, University of Zagreb, Department of Statistics
Trg J.F. Kennedya 6, 10 000 Zagreb, Croatia
E-mail: <mmatosec@efzg.hr>

Abstract: This paper is a pioneer empirical attempt to discern the interrelationship between news and stock market developments in Croatia. Using Structured Query Language (SQL) manipulations, the authors extract a database of news articles from some of the most popular Croatian news portals (Jutarnji list, Večernji list, Poslovni dnevnik, 24 sata, Index.hr, and Dnevnik.hr). Corroborating the loss aversion hypothesis, it is found that negative news Granger-cause both CROBEX returns and market turnover, while the influence of positive news is not so important.

Keywords: CROBEX, stock market, news media, loss aversion, prospect theory, negativity bias

1 INTRODUCTION

The interrelationship between news media and various types of economic behaviour is empirically confirmed by numerous studies. The media seem to play an important role in determining people's economic attitudes [6], they affect people's voting preferences [15], and even influence macroeconomic conditions such as the inflation rate [11]. This paper concentrates on the possible influence of news media on stock market developments in Croatia.

To be specific, we add to the literature by several aspects. First, we form a unique database of web articles from the archives of the most popular Croatian news portals: Jutarnji list, Večernji list, 24sata, Index.hr, Dnevnik.hr, and Poslovni dnevnik. Second, upon a meticulous analysis of the observed articles, we discriminate between positive and negative news on stock market developments to search for possible asymmetries in the observed relationship. The motivation for this distinction is found in the prospect theory of Kahneman and Tversky [8], postulating that agents are loss averse. Negative news should in that context cause a more intensive shift of the target variable (stock market return or turnover) than positive news. Third, we introduce a composite indicator of investor optimism as the difference between the frequencies of positive and negative news reports on stock market tendencies.

We find that both CROBEX return and the turnover of the Zagreb Stock Exchange are far more driven by negative than by positive news. On the other hand, media outlets also seem to react much more intensely to decreasing stock prices and diminishing market activity than to any positive developments. The paper is structured as follows. Section 2 provides a brief literature review, Section 3 discusses the main data and methodological specificities, while Section 4 presents the obtained empirical results. Finally, Section 5 concludes the paper.

2 LITERATURE REVIEW

When examining the psychological factors driving the stock markets, researchers usually resort to two types of variables: investor optimism and media news. The two concepts are obviously quite interrelated. News should have an impact on stock prices and stock returns dynamics, while the stock market developments and changes in economic sentiment should also be positively correlated. These hypotheses are the subject of many scientific studies [4, 9, 1, 3, 2, 5].

Investor optimism is empirically quite easy to proxy by e.g. Business and Consumer Survey data. Akhtar et al. [1] analyse the effect of consumer sentiment announcements on the Australian equity market. They provide evidence that when consumer sentiment indicator is lower than in the previous month, a significant negative daily effect occurs on the Australian equity market. When the released consumer sentiment indicator is higher than last month, the equity market has no significant reaction.

A lack of information has a considerable effect on stock returns, and it also drives investors' and consumers' sentiment. For example, Chen [3] provides evidence that the lack of consumer confidence indeed has an asymmetric effect on stock market fluctuations. He found that the impact of consumer confidence is greater in bear markets than in bull markets. In addition, with the greater stock market pessimism, the probability of switching from bull to bear markets is higher. The author finds that greater market pessimism has as a consequence that the market stays in bear regime longer. On the other hand, a greater lack of confidence leads the stock market into the bear market.

On the other hand, the information set transferred to the investors through news media (especially the positive/negative tone of news articles) is of latent character and rather hard to quantify. That is why e.g. Schumaker and Chen [13] use machine learning to scrutinize the value of news media information for predicting stock market returns. The authors provide evidence that the financial news articles (published twenty minutes before the price reaction) provide considerably accurate forecasts of actual future price dynamics.

Further on, several authors recognize that the tone of news articles heavily reflects on stock price dynamics. Bad and good news do not have the same impact on stock return volatility changes. In many cases, bad news (also reflected through the negative economic sentiment) produce a significant negative effect on the equity market. Most people usually overreact to unexpected bad news [4].

Veronesi [14] uses a rational expectations equilibrium model to find that, during recessions, investors are pessimistic and uncertain about the future stock market dynamics, effectuating in higher market volatility. In the period of dominating bad news, future expected dividends tend to decrease. Then the risk-averse investors' require additional asset price reductions, and the price consequentially drops by more than it fundamentally should. Quite intriguingly, the stock markets' overreaction to bad news is even more accentuated in good times than in the bad ones. A reverse mechanism is activated when good news occur in bad times. The expected future dividends increase, but the final effect on the asset prices is rather weak. Similar results were found by Koutmos and Booth [9]. They investigated the transmission mechanism of price and volatility spillovers across the New York, Tokyo and London stock markets and found that volatility spillovers in the analysed three markets are much more marked with bad news then with the other news.

The interrelationship of psychological factors and Zagreb Stock Exchange developments is still a quite underexplored phenomenon. One of its rare thorough studies (although a purely descriptive one) is done by Ivanov [7], who concludes that the investors' emotional attitudes and salaries influence the price dynamics on the financial market. The study presented in this

paper takes the issue one step further and analyses media content as a determinant of stock market trends.

3 DATA AND METHODOLOGY

In order to extract data on the amount of positive, negative or neutral news about the Croatian stock market that appeared in the media during the last 15 years, Web Crawler and SQL manipulations were used. The used media base consisted of six Croatian newspapers: Jutarnji list, Večernji list, Poslovni dnevnik, 24 sata, Index.hr and Dnevnik.hr. At first, before any SQL manipulations, there were over one million articles (1.069.967) that had to be processed. To filter all the articles, an SQL code was created by defining key words related to the Croatian stock market – "ZSE", "stock" and "CROBEX". Applying the code, 35.799 articles containing the above words remained and needed to be thoroughly examined, covering the period from November 2002 to April 2017.

Although many articles did not include necessary facts about the Croatian stock market situation or were ambiguous, others were considered important for investors' confidence and decisions, as they specifically discussed stock prices, profit and dividend trends. Therefore, all the news from the articles that had undoubtedly been interpreted as positive were marked with a plus, negative were marked with a minus, while stagnating trends were labeled as neutral. The rest of the articles had not been taken into account for the analysis.

In the end, the article base consisted of 6223 news articles. Time series were generated in monthly frequencies, as news believed to have positive, negative or neutral effects were summed for each month. Afterwards, the proportions of positive (PLUS), negative (MINUS) and neutral (NEUT) news in total number of monthly articles were calculated, as well as their monthly average and indices. All the composite indices were made by standardizing positive and negative average proportions of articles and scaled to have the mean of 100 and standard deviation of 10. Finally, an optimism indicator (OPTIMISM) was created as the difference between positive and negative proportions of monthly articles and also standardized and scaled in the same way[1].

To determine the causal relationship between the above indicators and CROBEX, as well as regular stock turnover on the Zagreb Stock Exchange, all of the variables were seasonally adjusted using the ARIMA X12 method. All analysed variables except CROBEX and stock market turnover were proven to be stationary using the ADF test (results available on demand). The latter two series are thus examined as logarithmic differences (DLCROBEX and DLTURNOVER). The variables of interest are depicted in Fig. 1.

---

[1] Further analysis did not take into account articles with a neutral note as they appeared to be rather irrelevant.

a) PLUS

b) MINUS

c) NEUTRAL

d) OPTIMISM

e) DLCROBEX

f) DLTURNOVER
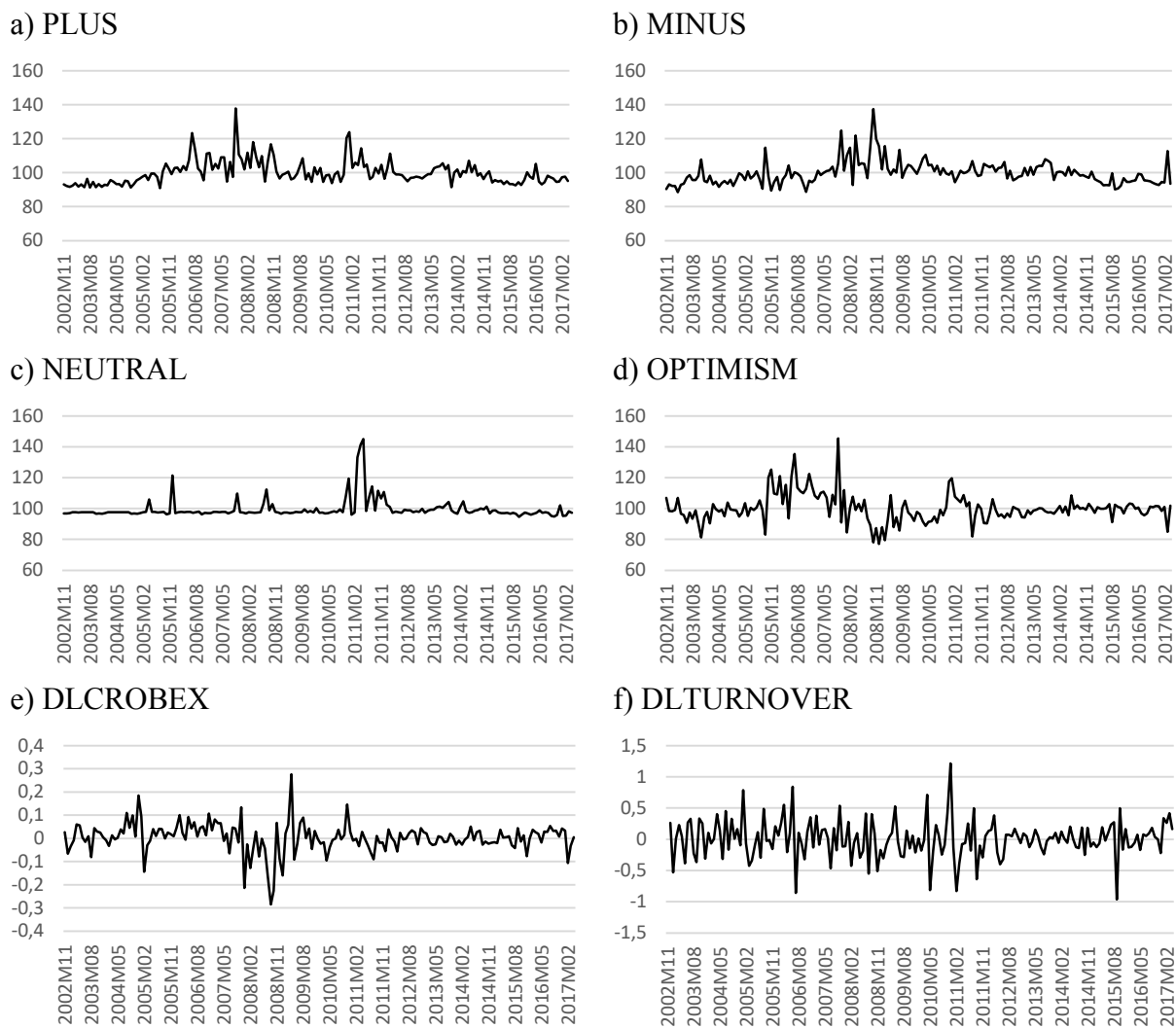
Figure 1: Graphs of analysed variables

4 EMPIRICAL RESULTS

This chapter presents and discusses the correlation matrix and results of the utilized econometric tests. Regarding the former, only correlations between CROBEX returns and negative news index, as well as between CROBEX returns and optimism indicator proved to be statistically significant; having also the expected sign (see Table 1). However, other correlation coefficients are not significant. That is not so surprising since this correlation matrix offers only a static view on the relationship between the variables. Thus, further steps in the analysis concern its dynamics.

Table 1: Correlation matrix

	PLUS	*MINUS*	*OPTIMISM*
DLCROBEX	-0.057	-0.254***	0.243***
DLTURNOVER	0.028	-0.012	0.043

Note: Coefficient significant at the level of significance: ***1%, **5%, *10%.

To statistically examine the causal relationship between the variables of interest, two multivariate and two bivariate reduced form VAR models were estimated. The lag order for each of the estimated models is chosen by the Schwarz information criterion. If the proposed lag order was not sufficient to eliminate residual autocorrelation (according to the LM autocorrelation test), additional lags were added successively. The final chosen lag orders are shown in Table 2.

Considering the fact that the applied White test has rejected the null of homoskedasticity for all the estimated models, heterogeneity was accounted for by including robust standard errors (HAC). The non-normality observed by the Jarque-Bera test is not considered to be a necessary condition for the validity of VAR modelling [10, 12].
Finally, the results of Granger Causality Test are presented in Table 2.

Table 2: Results of Granger Causality Test

Model	Lag length	Causal direction	Test statistic
Multivarate VAR model (DLCROBEX, PLUS, MINUS)	4	$MINUS \rightarrow DLCROBEX$ $PLUS \rightarrow DLCROBEX$ $DLCROBEX \rightarrow MINUS$ $DLCROBEX \rightarrow PLUS$	2.078* 0.903 4.182** 0.045
Bivariate VAR model (DLCROBEX, OPTIMISM)	2	$DLCROBEX \rightarrow OPTIMISM$ $OPTIMISM \rightarrow DLCROBEX$	6.179*** 1.226
Multivarate VAR model (DLTURNOVER, PLUS, MINUS)	7	$MINUS \rightarrow DLTURNOVER$ $PLUS \rightarrow DLTURNOVER$ $DLTURNOVER \rightarrow MINUS$ $DLTURNOVER \rightarrow PLUS$	6.132*** 2.854*** 1.584 3.699***
Bivariate VAR model (DLTURNOVER, OPTIMISM)	5	$DLTURNOVER \rightarrow OPTIMISM$ $OPTIMISM \rightarrow DLTURNOVER$	1.169 0.843

Note: Parameter significant at the level of significance: ***1%, **5%, *10%.

It is evident that there is a statistically significant bi-directional causal relationship between CROBEX returns and the negative news indicator, as well as between turnover and positive news indicator. Additionally, CROBEX returns cause optimism, and negative news index causes the turnover growth rate.

5 CONCLUSION

This paper presents the initial effort to scrutinize the interdependence between media news (the information set they offer to investors) and stock market developments in Croatia. Through discriminating between positively and negatively toned articles, we find that the prospect theory hypothesis of loss aversion is firmly established in the empirical data. The investors' response to negative news (both in terms of CROBEX returns and market turnover) is more pronounced than their reaction to positive news. A stronger link between the stock market and negative news is also corroborated by correlation analysis. The introduced optimism indicator is strongly linked to CROBEX returns, but no distinct leading characteristics are found in terms of market turnover. The last implication of the obtained results is the confirmation of the negativity bias. CROBEX returns feed into the tone of news articles only in the case of negative news. In line with the negativity bias, the media seem to over-blow negative news and have a strict focus on negative tendencies.

References

[1] Akhtar, S., Faff, R., Oliver, B., Subrahmanyam, A. (2010). The power of bad: The negativity bias in Australian consumer sentiment announcements on stock returns. Journal of Banking & Finance, 35: 1239-1249.

[2] Bollen, J., Mao, H and Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Sciense, 2:1-8.

[3] Chen, S. (2011). Lack of consumer confidence and stock returns. Journal of Empirical Finance, 18: 225-236.

[4] De Bondt, W. and Thaler, R. (1985). Does the Stock Market Overreact?. The Journal of Finance, XI(3): 793-805.

[5] Haq, S. and Larsson, R. (2016). The dynamics of stock market returns and macroeconomic indicators: An ARDL approach with cointegration, Master of Science Thesis INDEK 2016:59 ME211X, KTH. Stockholm: KTH Industrial Engineering and Management.

[6] Hester, J.B. and Gibson, R. (2003). The economy and second-level agenda setting: a time series analysis of economic news and public opinion about the economy. Journalism and Mass Communication Quarterly, 80(1): 73-90.

[7] Ivanov, M. (2008). Utjecaj psiholoških čimbenika na djelotvornost financijskih tržišta. Zbornih ekonomskog fakulteta Sveučilišta u Mostaru. 1, 7-30.

[8] Kahneman D. and Tversky A (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47: 263-292.

[9] Koutmos, G. and Booth, G. G. (1995). Asymetric volatility transmission in international stock markets. Journal of International Money and Finance, 14(6): 747-762

[10] Lanne, M. and Lütkepohl, H. (2010). Structural Vector Autoregressions with Nonnormal Residuals. Journal of Business & Economic Statistics, 28(1): 159-168.

[11] Lolić, I., Sorić, P. and Čižmešija, M. (2017). Disentangling the relationship between news media and consumers' inflation sentiment: the case of Croatia. Czech Journal of Economics and Finance, 67(3): 221-249.

[12] Lütkepohl, H. (2010). Vector Autoregressive Models. Economics Working Paper ECO 2011/30, European University Institute.

[13] Schumaker, R.P. and Chen, H. (2009). Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System. ACM Transactions on Information Systems, 27(12): 12:1-12:3.

[14] Veronesi, P. (1999). Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. The Review of Financial Studies, 12(5): 975-1007.

[15] Shah DV, Watts MD, Domke D, Fan DP, Shah MF (1999). News Coverage, Economic Cues, and the Public's Presidential Preferences, 1984–1996. The Journal of Politics, 61(04): 914-943.

PERIODIC AVERAGE NATIONAL REFERENCE RATE AS A NEW FINANCIAL STANDARD

Elza Jurun
University of Split/Faculty of Economics Split
21000 Split, Cvite Fiskovića 5, Croatia
E-mail: elza@efst.hr

Nada Ratković
University of Split/Faculty of Economics Split
21000 Split, Cvite Fiskovića 5, Croatia
E-mail: nada.ratkovic@efst.hr

Ivana Matić
Imex Bank Split
21000 Split, Tolstojeva 6, Croatia
E-mail: ivana.matic@imexbanka.hr

Abstract: In the focus of this paper is a new approach of financial competitiveness measurement. As the standard of this measurement, Periodic Average National Reference Rate (PANRR) is proposed. Authors conceive this standard as national replacement for international standards like EURIBOR or LIBOR. The average capital cost for national financial market is defined using monthly statistical reports published on the official Croatian National Bank website in the time horizon from 2009 to 2016. Finally, PANRR is established as the standard of financial competitiveness measurement.

Key words: Capital cost, National Reference Rate, Financial competitiveness, Periodic Average National Reference Rate

1 INTRODUCTION

Capital cost is the interest rate or rate of return at which capital is given for using at the capital market. Interest is a fee paid by the loan recipient to provider because equity is given in the present to be returned at some future time. As a representative of the all market interest rates movements at the credit market, the market interest rate is used.

Therefore, at the beginning it is necessary to define the methodology of calculating the National Reference Rate (NRR) which is the basis for the calculation of Periodic Average National Reference Rate (PANRR). This paper aims to promote PANRR as the new financial competition standard at the banking market. Also PANRR authors propose as the national replacement for international standards like EURIBOR or LIBOR.

Croatian Banking Association (CBA) has defined the reference rate of the average cost of financing Croatian banking sector – NRR. Starting from 2013 CBA has developed a transparent and simple methodology of NRR calculation for certain currencies using the only publicly published data of the Croatian National Bank (CNB). Based on the operations data of the banking sector CNB started to publish on their website "Indicators of credit institutions" [3].

NRR, therefore, represents the average funding sources cost of Croatian banking sector with respect to certain past period, other sources and relevant currency. The kinds of sources are considered to be retail deposits, corporate deposits from non-financial sector and other sources of banks funds. NRR, therefore, represents the average interest rate paid by the banking sector, in order to obtain the funds required for credit business. For these purposes three types of NRR are defined as following:

- NRR for deposits of natural persons (calculated for HRK and EUR) - NRR1.

- NRR for deposits of natural persons and non-financial sector (calculated for currency HRK and EUR) - NRR2.
- NRR for all major sources of funds from all natural and legal persons, including those from the financial sector (calculated for HRK, EUR, CHF and USD) - NRR3.

2 METHODOLOGY OVERVIEW

Each of the above NRRs is calculated for the period of 3 months, 6 months and 12 months and is marked with 3M, 6M and 12M. The abbreviation of each NRR contains the clear indications about funds comprised, the indication of the period for which it is calculated and specifies the currency. For example, quarterly NRR, formed on the basis of data on deposits of natural persons in HRK, will be marked as "3M NRR1 for HRK".

CNB in the second month of each quarter by the 15th day of the month, publishes aggregate data (covering all banks and savings banks licensed to operate in the Republic of Croatia) on interest costs on the main sources of funds during the previous quarter, as well as the states of the sources of funds to the end of each month of the quarter to which these interest expenses related (data divided on the means natural persons, natural persons and non-financial sector, as well as all natural and legal persons). Based on these data, the Croatian Banking Association (CBA) calculates the value of each NRR (according to the type of interest rate and currency period). The first business day after the day when the CNB publish aggregate data each NRR is published on the website of CBA.

Published NRR is valid as the reference rate for the harmonization of the variable part of the variable interest rate until the day before the day of following NRR publication.

NRR for the specific currency, determined by the scope of the sources of funds and the period which may be 3, 6 or 12 months, is calculated as follows:

The total interest cost of the banking sector is calculated, according to the particular currency and coverage of the fund sources that constitute the NRR, for the preliminary number of months corresponding to the length of the period. The average balance of the principal sources of funds in the previous period is estimated.

After the ratio of total interest cost and the average of the sources of funds are estimated, the ratio has to be divided by the total number of days in the previous period and multiplied by 365. It is the way to obtain the interest rate expressed on an annual basis and the resulting interest rate represents the NRR. NRR of each type is calculated according to the generic formula [4].

3 PERIODIC AVERAGE NATIONAL REFERENCE RATE

3.1 National Reference Rate as the instrument of Capital Cost measuring

Briefly speaking in the context of this paper Capital cost is the interest rate or rate of return at which capital is given for using at the financial market. NRRs which are already listed and defined in the previous chapter will be calculated for the period from 2009 to 2016 for the purposes of Capital cost measuring. To calculate these NRR, data base on interest costs of the main sources of funds in the balance sheet of banks and savings banks licensed to operate in the Republic of Croatia and the state of these sources are needed. Required data are taken from the aggregated statistical reports monthly published by the CNB on its website.

Coverage of resources that is taken into consideration is the Coverage 3 includes all natural and legal persons (including those from the financial sector) as well as all the major sources of funds (transaction accounts, all deposits, loans received and issued debt securities).

Coverage	HRK	EUR	USD	CHF
Quarterly	6M NRR3	6M NRR3	6M NRR3	6M NRR3
31.03.2009.	3.40%	3.55%	3.22%	3.11%
30.06.2009.	3.35%	3.51%	3.21%	3.09%
30.09.2009.	3.33%	3.41%	3.12%	3.15%
31.12.2010.	3.30%	3.36%	3.13%	3.14%
31.03.2010.	3.25%	3.33%	3.10%	2.95%
30.06.2010.	3.20%	3.23%	3.03%	2.91%
30.09.2010.	3.16%	3.20%	3.02%	2.99%
31.12.2010.	3.11%	3.16%	2.98%	2.84%
31.03.2011.	3.08%	3.21%	2.74%	2.70%
30.06.2011.	3.10%	3.19%	2.73%	2.63%
30.09.2011.	3.05%	3.15%	2.69%	2.51%
31.12.2011.	3.04%	3.10%	2.65%	2.48%
31.03.2012.	2.96%	3.02%	2.64%	2.43%
30.06.2012.	2.91%	3.10%	2.57%	2.38%
30.09.2012.	2.86%	3.09%	2.41%	2.13%
31.12.2012.	2.74%	2.92%	2.38%	2.36%
31.03.2013.	2.56%	2.89%	2.40%	2.34%
30.06.2013.	2.29%	2.86%	1.99%	2.28%
30.09.2013.	2.13%	2.75%	1.92%	2.24%
31.12.2013.	2.03%	2.65%	1.87%	2.25%
31.03.2014.	1.89%	2.55%	1.79%	2.20%
30.06.2014.	1.76%	2.49%	1.71%	2.11%
30.09.2014.	1.65%	2.43%	1.70%	2.12%
31.12.2014.	1.59%	2.34%	1.67%	2.11%
31.03.2015.	1.52%	2.23%	1.60%	2.03%
30.06.2015.	1.42%	2.12%	1.48%	1.87%
30.09.2015.	1.32%	2.00%	1.35%	1.69%
31.12.2015.	1.31%	1.89%	1.28%	1.79%
31.03.2016.	1.26%	1.79%	1.14%	1.53%
30.06.2016.	1.14%	1.66%	1.00%	0.62%
30.09.2016.	1.01%	1.48%	0.90%	0.58%
31.12.2016.	0.89%	1.36%	0.82%	0.51%

"EUR", "CHF" and "USD" include interest costs on the sources of funds in the currencies of the same name, as well as the HRK sources of funds indexed to these currencies. "HRK" does not include interest expense on funding liabilities indexed to foreign currency.

Before analysing data base at the annual level let us see the main results of the quarterly analysis. Table 1 shows the quarterly trends in NRR from 2009 to 2016. The average in quarterly analysis is defined as six month average and Coverage of funds is estimated as before mentioned Coverage 3. In the Table 1 it is marked as "6M NRR3". It may be noted that there is the downward trend of the NRR in the Croatian banking market for the whole researched period and it is expected to continue, due to favourable developments in the international markets. This is because the increase in the balance sheet of banks received deposits is greater than the increase of loans that banks will provide for their customers, i.e. the banks have excess funds that have no one to qualify for because of the reduced demand for loans.

Such a situation leads to the drop in interest rates on the main sources of funds in bank liabilities and leads to the decline on deposits. The reason why there is the decrease in interest rates on deposits of banks is the desire to maintain profit levels at the level of the previous year or to increase profit levels.

Due to the poor economic situation there are more and more loans that banks have to reserve. That leads to the fact that reservation costs increase as well as regulatory costs towards CNB.

These costs are regularly compensated by reducing salaries, bonuses and various other savings, including the most significant decrease in interest on deposits. All the pronounced facts are evident at the Figure 1.

6M NRS3

HRK 6M NRR3
EUR 6M NRR3
USD 6M NRR3
CHF 6M NRR3
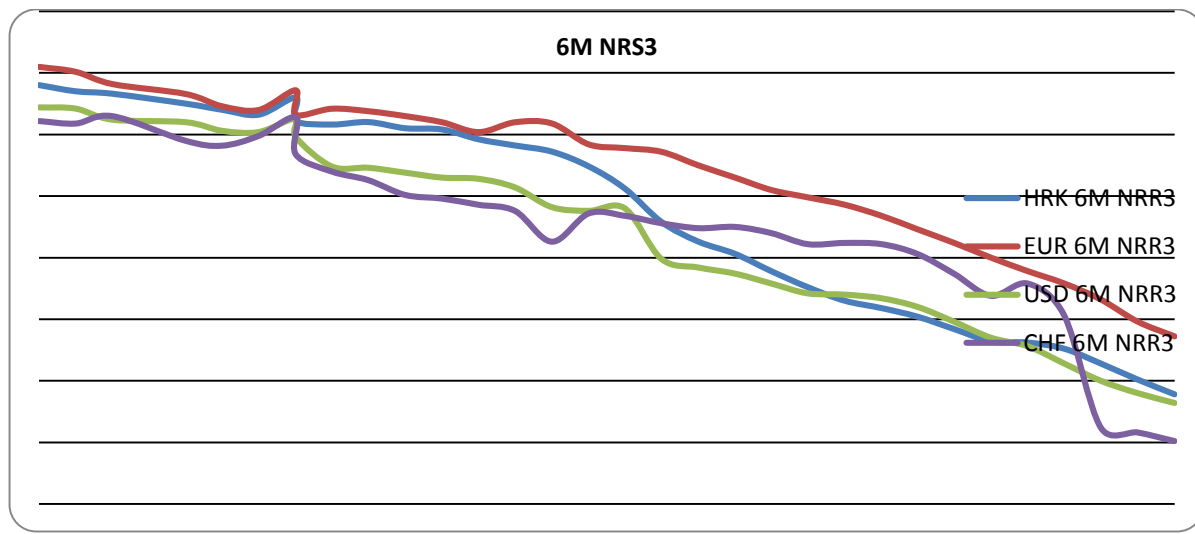
Figure 1: Trends of the National Reference Rates from 2009 to 2016.

Banks not only reduce interest on deposits, despite the increase in provisioning costs, but they also reduce the interest on loans due to the low demand for them. The deficient demand for loans is cause of pressure on the financial market to reduce interest spreads.

The cost of funding sources is increased by spread, country risk and other costs that the bank has to form loans interest rate. Reducing the NRR leads to the decrease in loans interests and thus reducing also results with attracting more and more bank's clients.

Recently Croatian admittance into the legal and institutional framework of the European Union is additional reason for the reduction of risks for domestic banks and international investors, so-called country risk. The interest rate decrease will be under influence of the movement of the European market interest rates where the EURIBOR has reached the historically minimum level.

3.2 Periodic Average Reference Rate methodology

This paper creates the model PANRR as a general standard for measuring the capital cost for all banks and other financial institutions at the territory of a country. Defining PANRR at the national level is necessary because each country has it's owns characteristics of the capital market and the banking system.

As it is practice in Croatian banking system, Croatian branches of the European banks give, for example, housing loans at the higher interest rates than the same bank branches give in the country where is the headquarter of the same banks association. In addition, support of building PANRR as the national standard is the fact that the model calculations involved

numerous actual specifics of the national capital market. Besides, the data dealing with the average capital cost for the fixed period and for the various currencies at the national capital market are much easier to get.

Furthermore, this paper develops the unique PANRR model for all currencies and for all time scopes. In order to calculate the annual unique NRR for all currencies and for all time scopes, average quarterly NRR3 has been estimated. As the first step it was necessary to calculate the structure of sources of funds in liabilities for each currency. The final results for all sources of funds and liabilities by all currencies from March 2009 to December 2016 are presented in Table 2 as well as the average quarterly NRR3 for the same observed period.

Table 2: The structure of sources of funds in liabilities by currency and average quarterly NRR3 from March 2009 to December 2016

Coverage	Sources of funds in liabilities				Average quarterly NRR3
Quarterly	HRK	EUR	USD	CHF	
31.03.2009.	0.3575	0.5590	0.0350	0.0440	3.45%
30.06.2009.	0.3564	0.5600	0.0321	0.0460	3.40%
30.09.2009.	0.3351	0.5833	0.0296	0.0464	3.34%
31.12.2009.	0.3204	0.6030	0.0295	0.0420	3.31%
31.03.2010.	0.3118	0.6045	0.0330	0.0448	3.26%
30.06.2010.	0.3180	0.5889	0.0367	0.0530	3.19%
30.09.2010.	0.3148	0.5939	0.0355	0.0500	3.15%
31.12.2010.	0.3112	0.5943	0.0355	0.0530	3.10%
31.03.2011.	0.3126	0.5971	0.0346	0.0498	3.11%
30.06.2011.	0.3201	0.5792	0.0354	0.0591	3.09%
30.09.2011.	0.3206	0.5798	0.0364	0.0570	3.05%
31.12.2011.	0.3232	0.5673	0.0394	0.0623	3.00%
31.03.2012.	0.3200	0.5698	0.0384	0.0635	2.92%
30.06.2012.	0.3185	0.5691	0.0382	0.0653	2.94%
30.09.2012.	0.3228	0.5646	0.0400	0.0633	2.90%
31.12.2012.	0.3137	0.5736	0.0430	0.0625	2.78%
31.03.2013.	0.3214	0.5693	0.0392	0.0620	2.71%
30.06.2013.	0.3137	0.5736	0.0391	0.0593	2.57%
30.09.2013.	0.3137	0.5736	0.0385	0.0613	2.46%
31.12.2013.	0.3137	0.5736	0.0357	0.0623	2.36%
31.03.2014.	0.3154	0.5697	0.0356	0.0442	2.21%
30.06.2014.	0.3154	0.5585	0.0353	0.0445	2.10%
30.09.2014.	0.3282	0.5574	0.0364	0.0408	2.04%
31.12.2014.	0.3285	0.5580	0.0371	0.0412	1.98%
31.03.2015.	0.3139	0.5443	0.0475	0.0306	1.83%
30.06.2015.	0.3211	0.5749	0.0467	0.0486	1.83%
30.09.2015.	0.3233	0.5887	0.0453	0.0343	1.72%
31.12.2015.	0.3298	0.6047	0.0492	0.0121	1.66%
31.03.2016.	0.3280	0.6036	0.0478	0.0111	1.57%
30.06.2016.	0.3395	0.5923	0.0483	0.0112	1.43%
30.09.2016.	0.3438	0.5890	0.0467	0.0115	1.27%
31.12.2016.	0.3535	0.5791	0.0482	0.0103	1.15%

At the next level of modelling PANRR as the national standard it is necessary to estimate the unique annual NRR. For this task the simple geometric mean of average quarterly NRR3 can

be used. Computing of the unique annual NRR using the geometric mean of the average quarterly NRR3 is shown below. Results cover the research period from 2009 to 2016.

Average annual NRR $_{2009}$ = $\sqrt[4]{0.0345*0.0340*0.0334*0.0331}$ = 3.37%

Average annual NRR $_{2010}$ = $\sqrt[4]{0.0326*0.0319*0.0315*0.0310}$ = 3.17%

Average annual NRR $_{2011}$ = $\sqrt[4]{0.0311*0.0309*0.0305*0.0300}$ = 3.06%

Average annual NRR $_{2012}$ = $\sqrt[4]{0.0292*0.0294*0.0290*0.0278}$ = 2.88%

Average annual NRR $_{2013}$ = $\sqrt[4]{0.0271*0.0257*0.0246*0.0236}$ = 2.52%

Average annual NRR $_{2014}$ = $\sqrt[4]{0.0221*0.0210*0.0204*0.0198}$ = 2.08%

Average annual NRR3 $_{2015}$ = $\sqrt[4]{0.0183*0.0183*0.0172*0.0166}$ = 1.76%

Average annual NRR3 $_{2016}$ = $\sqrt[4]{0.0157*0.0143*0.0127*0.0115}$ = 1.35%

Due to the impact of the prevailing trends at the financial markets at home and abroad, as might be expected, in the reporting period annual NRR continuously has declining trend.

4 CONCLUSION REMARKS

This paper aims to promote PANRR as the new financial competition standard at the banking market, something like a national replacement for international standards EURIBOR or LIBOR. Briefly speaking, in the context of this paper NRR is fundamental in defining Capital cost as the rate of return at which capital is given for using at the financial market.

By this research the whole PANRR methodology for time horizon 2009-2016 is carried out. This paper creates the model PANRR as the general standard for measuring the capital cost for all banks and other financial institutions at the territory of the country. The integral part of the paper is the case study of Croatia where the whole concrete modelling procedure for calculating average annual NRR is carried out over the time horizon from 2009 to 2016 year.

There is the downward trend of the NRR in the Croatian banking market for the whole researched period and it is expected to continue, due to favourable developments at the international markets.

For further research authors are planning to compare the results of implementing PANRR as the financial standard with the current situation of using classical financial standards. Defining PANRR at the national level is necessary because each country has it's owns characteristics of capital market and the banking system. Besides, the data dealing with the average capital cost for the fixed period and for the various currencies at the national capital market are much easier to get.

References

[1] Elton, E.J., Gruber, M.J., Brown, S.J. and Goetzmann, W. N. (2014). Modern Portfolio Theory and Investment Analysis, John Wiley & Sons.

[2] Foley,B. (2005): Tržište kapitala, (prijevod), Mate, Zagreb.

[3] HNB, Croatian National Bank, (2009-2017) "Statistical Reports".

[4] http://www.hub.hr/sites/default/files/2015metodologija_nrs.pdf[Accessed 30/05/2017]

[5] Jones, A., and Sufrin, B. (2014). Text, Cases, and Materials: EU Competition Law, 5th edt, Oxford University Press, Oxford.

[6] Krnic, B. (2014). Determinants of lending interest rates granted to companies in Croatia. Journal of Accounting and Management, Vol 4 (No 2), 1-20.

[7] Saunders, A., Millon Cornett, M., (2006): Financijska tržišta i institucije, Masmedia, Zagreb

[8] Šohinger, J., Galinec, D. (2013). Volatility of capital flows in emerging European economies: Lessons from Asia. Ekonomska misao i praksa, (No 1), 275-296.

AN EOQ MODEL WITH PARTIAL BACKORDERS UNDER FINANCIAL CONSTRAINTS AND MARKET TOLERANCE

Krommyda I.P.
Independent researcher
45333 Ioannina, Greece
E-mail: ikrommyd@gmail.com

Skouri K.
University of Ioannina, Department of Mathematics
45110 Ioannina, Greece
E-mail: kskouri@uoi.gr

Lagodimos A.G.
University of Piraeus, Department of Business Administration
18534 Piraeus, Greece
E-mail: alagod@unipi.gr

Abstract: We consider an extension to the classical EOQ model by incorporating features mostly pertaining to the recent economic crisis. On the supply side, there are financial constraints regarding replenishment, which first limit the available capital per replenishment and second impose the prepayment (either full or partial) of the quantity ordered. On the consumption side, the new concept of market tolerance is introduced, which effectively permits backorders to occur at no-cost for a predetermined time period. Analytical and computational results reveal the impact that each factor considered has on the system cost as well as the sensitivity of the system parameters.

Keywords: Single-echelon; Deterministic; Lost sales; Backlogging; Prepayment;

1 INTRODUCTION

The global financial crisis that we are experiencing is inevitable to have a significant effect on inventory management decisions. The classical EOQ model developed by Harris [3] implicitly assumes that the customer pays for the items at the time they are received. However, uncertainty, as a major consequence of the financial crisis, may urge suppliers in many cases to require full or partial prepayments for the ordered quantity (Taleizadeh et al. [10], Taleizadeh [11], Zhang et al. [14], Shah et al. [9], Tavakoli & Taleizadeh [12], Wu et al. [13]). The importance of prepayment is empirically highlighted by Ahn et al. [1], who show that the lack of trade credit and prepayment played an important role in the 2008 global trade collapse. An additional consequence of the financial crisis is budget limitations. Although inventory models under financial constraints have already been studied in the literature, a situation, like capital controls (as it happened in Greece in July 2015), makes the consideration of tough budget limitations absolutely necessary.

The above financial issues lead unavoidable an inventory system to stock out. Generally, the demand during stock-out period is regarded either as completely backordered, or completely lost, or partially backordered (i.e. a percentage of customers are willing to wait till the next arrival of stock while others cannot wait and will fill their demand from another source).The first model that added the assumption of partial backordering of demand during a stock-out period, to the classical EOQ model, was developed by Montgomery et al. [4]. A survey by Pentico and Drake [7] describes the deterministic models that have been formulated since then, which include not only partial backordering but a variety of other considerations. In all models with shortages a cost that is called backorders or backlogging cost is encountered. It is logical to assume that a financial crisis can also have an impact on customers' attitude, making the customer more tolerant to the occurrence of shortages either

due to solidarity or even because of reduction to retail outlets. Hence, we assume that there exists a specific time period during shortages, during which no backorders cost occur. After the end of this period a cost is charged per unit backordered, per time unit. This tolerance period could also be viewed as a credit period offered by the customer to the vendor, as we will see. We must note that the existence of customer tolerance during shortages is not an assumption that is limited to the situation of a financial crisis but could also be applied when dealing with a product distributed by a company with a prominent brand name. In such a situation, a customer could also be expected to show some tolerance in the occurrence of shortages being compensated by the fact that he will receive a trusted and desirable product.

In this study the classical EOQ model is modified in order to quantify the effect of the financial crisis on inventory management decisions. The following situations are combined: 1) the supplier requirement for full or partial order prepayments 2) the existence of budget limitations for inventory procurement 3) the manifestation of customer tolerance on inventory shortages due to reduced purchasing outlets.

The rest of the paper is organized as follows. In Section 2, we provide the necessary notations and assumptions used for the mathematical formulation of the inventory model, which is presented in Section 3. The solution procedure and optimal policy are presented in Section 4. A numerical example along with its sensitivity analysis is provided in Section 5. Finally conclusions are given in Section 6.

2 NOTATIONS AND ASSUMPTIONS

In this section we introduce the notation and operating assumptions underlying the analysis and used throughout the paper.

Notations:

D constant demand rate (items per time unit)

h inventory holding cost including capital cost (per item per time unit)

p bought-out items purchasing price (per item)

K ordering cost (per replenishment order)

b backorders cost (per item per time unit)

I_c interest rate for borrowing money (per time unit)

s lost-sales cost (lost profit + loss of goodwill per item lost)

B purchasing budget constraint (limiting total purchases per cycle)

T length of inventory cycle [a decision variable]

t_1 part of inventory cycle with non-zero inventory [a decision variable]

T_p fixed advanced payment period

T_t length of customers tolerance period

Q order quantity [decision variable, which is determined after determination of t_1 and T]

ω for convenience, we define $\omega = pI_c aT_p$, which can be interpreted as prepayment cost per

Assumptions

1. The demand rate is known constant, inventory replenishment rate is infinite.
2. Shortages are allowed and partially backlogged at a known fraction β, $0 \le \beta \le 1$; clearly, $\beta = 0$ and $\beta = 1$, represent complete lost sales and complete backlogging, respectively.
3. The supplier imposes that a fraction a (with $0 \le a \le 1$) of total purchasing cost is prepaid at order placement; clearly $a = 1$ represents full prepayment and $a = 0$ no prepayment (as in the classical EOQ model).

4. The buyer's budget for purchases (per inventory cycle) is constrained by a finite upper limit B.
5. There exists a finite time window (of length T_t), where the fraction of customers (β) who are willing to wait are rather tolerant to shortages; so, no backorders cost accruals occur in this period.
6. We define $\omega = pI_c aT_p$ as the prepayment cost per item and it is assumed that $s - \omega \geq 0$.

3 MODEL FORMULATION

In this section the system is modeled using as decision variables the part of the inventory cycle with non-zero inventory i.e. t_1, and the length of the inventory cycle, i.e. T. Notice that, using these decision variables, the order quantity, Q, can be easily determined.

According to assumption 3, a fraction a of the purchasing cost of the order quantity is prepaid upon order placement. The order quantity arrives after time period T_p and brings the inventory level up to Dt_1. Then the inventory level depletes due to demand and becomes zero at time point t_1. During time period $T - t_1$ shortages occur which are backlogged at a constant fraction β, while $(1 - \beta)$ is lost. This implies that the ordering quantity is $Q = D\left[t_1 + \beta(T - t_1)\right] = D\left[T - (1 - \beta)(T - t_1)\right]$. We assume that there exists a time period, T_t, during shortages, in which the customers show tolerance and no backorders cost occur.
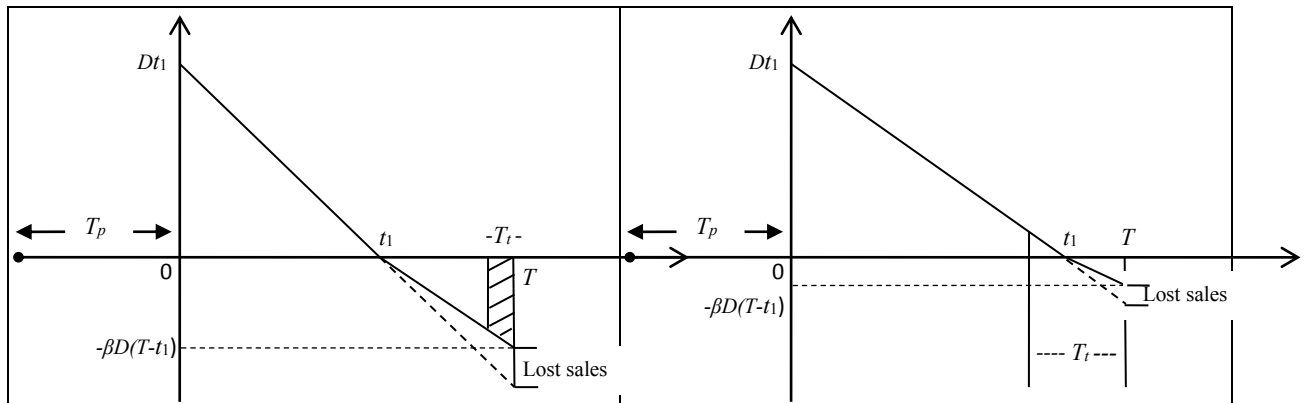
Figure 1a: The prepayment inventory model with shortages and customer tolerance on backlogging – Case A

Figure 1b: The prepayment inventory model with shortages and customer tolerance on backlogging – Case B

In the framework of the assumptions, the total cost of the inventory system, is comprised by:
The ordering cost: K

The inventory holding cost, including capital cost, during time period t_1: $\dfrac{hDt_1^2}{2}$

The capital opportunity cost during the time period, T_p, due to the prepayment:
$$pI_c aT_p Q = \omega D\left[t_1 + \beta(T - t_1)\right]$$
The lost sales cost during time period $T - t_1$: $s(1 - \beta)D(T - t_1)$

In order to model the backorders cost, two cases are considered, depending on the length of the shortage period $T - t_1$ and the customer tolerance period T_t.

In the first case (depicted in Figure 1a), we assume that the length of the period during which shortages occur is greater than the customer tolerance period, i.e. $T_t < T - t_1$, hence backorders costs occur during time period $T - (t_1 + T_t)$ and are equal to $\dfrac{b\beta D(T - (t_1 + T_t))^2}{2}$.

In the second case (depicted in Figure 1b), the customer tolerance period is greater or equal to the shortages period, $T - t_1 \leq T_t$, hence no backorders costs occur in this case.

Summarizing, the total cost for the two cases, per unit of time, is:

$$TC(T,t_1) = \begin{cases} TC_A(T,t_1), & \text{if } 0 \leq t_1 < T - T_t \\ TC_B(T,t_1), & \text{if } T - T_t \leq t_1 \leq T \end{cases} \tag{1}$$

where

$$TC_A(T,t_1) = \frac{K}{T} + \frac{hDt_1^2}{2T} + \frac{\omega D\left[t_1 + \beta(T - t_1)\right]}{T} + \frac{b\beta D(T - (t_1 + T_t))^2}{2T} + \frac{s(1 - \beta)D(T - t_1)}{T} \tag{2}$$

$$TC_B(T,t_1) = \frac{K}{T} + \frac{hDt_1^2}{2T} + \frac{\omega D\left[t_1 + \beta(T - t_1)\right]}{T} + \frac{s(1 - \beta)D(T - t_1)}{T} \tag{3}$$

Remark 1. The concept of tolerance during shortages can be viewed as analogous to the concept of trade credit offered by the supplier to the retailer. More specifically, let T_t denote either the tolerance period or the offered credit period, Y_1 the time period with positive inventory level and Y_2 the time period with no available inventory (β=1).

If trade credit is offered during a time period Tt, then the capital cost savings are: $pI_c\left(\dfrac{DY_1^2}{2} - \dfrac{D(Y_1 - T_t)^2}{2}\right)$ (Goyal [2]). Equivalently, when a tolerance period Tt during shortages exists, the cost savings are: $b\left(\dfrac{DY_2^2}{2} - \dfrac{D(Y_2 - T_t)^2}{2}\right)$.

4 THE OPTIMAL POLICY

In this section we determine the values of the decision variables (T, t_1) that minimize the total cost of the inventory system under the budget restrictions. Due to limited space the proofs are omitted. Notice that the feasible region is a compact set and the problem has always a minimum. The cost function, $TC(T, t_1)$, is not, in general, convex in (T, t_1), which is in line with the respective result in Montgomery et al. [4].

Proposition 1
If $2Kh - (s - \omega)^2(1 - \beta)^2 D > 0$, the cost function $TC(T, t_1)$ is differentiable and convex for every t_1, T.

In order to proceed with optimization the following problems will be solved:

$$P_1 : \begin{cases} \min TC_A(T, t_1) \\ s.t. \ \ pD\left[t_1 + \beta(T - t_1)\right] \leq B \\ \quad 0 \leq t_1 \leq T - T_t \end{cases} \qquad P_2 : \begin{cases} \min TC_B(T, t_1) \\ s.t. \ \ pD\left[t_1 + \beta(T - t_1)\right] \leq B \\ \quad T - T_t \leq t_1 \leq T \end{cases}$$

Let (t_1^*, T^*) denote the optimal solution of the problem, $(t_{1,A}^*, T_A^*)$ the optimal solution of P_1, $(t_{1,B}^*, T_B^*)$ the optimal solution of P_2. Consequently, the optimal values of the decision variables, (T, t_1), are given as $(t_1^*, T^*) = \arg\{\min\{TC_A(t_{1,A}^*, T_A^*), TC_B(t_{1,B}^*, T_B^*)\}\}$.

The next proposition gives the optimal policy of the problem when the budget constraint is satisfied.

Proposition 2

The optimal policy of the problem is as follows, provided that $pD\left[t_1^* + \beta\left(T^* - t_1^*\right)\right] \le B$.

1. If $2Kh - (s - \omega)^2 (1 - \beta)^2 D > 0$ and $T_t < \dfrac{B}{pD\beta}$, then $t_1^* = t_{1,A}^*$, $T^* = T_A^*$, where

$$t_{1,A}^* = \frac{b\beta\left(T^* - T_t\right) + (s - \omega)(1 - \beta)}{h + b\beta} \text{ and}$$

$$T_A^* = \sqrt{\frac{2K(h + b\beta) - (s - \omega)^2(1 - \beta)^2 D}{b\beta h D} + T_t^2 + \frac{2(s - \omega)(1 - \beta)T_t}{h}} \;.$$

2. If $2Kh - (s - \omega)^2 (1 - \beta)^2 D > 0$ and $T_t \ge \dfrac{B}{pD\beta}$, then $t_1^* = 0$, $T^* = \dfrac{B}{pD\beta}$.

3. If $2Kh - (s - \omega)^2 (1 - \beta)^2 D < 0$, then $t_1^* = T^* = \sqrt{\dfrac{2K}{hD}}$.

Remark 2. If the budget constraint is not satisfied, i.e. $pD\left[t_1^* + \beta\left(T^* - t_1^*\right)\right] > B$, the solution should be searched on the boundary i.e. on the budget constraint ($pD\left[t_1 + \beta\left(T - t_1\right)\right] = B$). Again, due to limited space the details are omitted.

Remark 3. According to Proposition 2 the optimal solution of the problem depends on the condition $2Kh - (s - \omega)^2 (1 - \beta)^2 D > 0$. Equivalent conditions have been derived by previous authors studying the EOQ with partial backorders (for example Montgomery et al. [4]), Rosenberg [8], Park [5] and Pentico and Drake [6].

The condition $2Kh - (s - \omega)^2 (1 - \beta)^2 D > 0$ can be modified to $\sqrt{2KhD} > (s - \omega)(1 - \beta)D$. A managerial interpretation of this is that the quantity $\sqrt{2KhD}$ describes the optimal cost for the EOQ problem without backlogging. On the other hand $(s - \omega)(1 - \beta)D$ describes the lost sales cost due to shortages (the quantity $\omega(1 - \beta)D$ is subtracted because it expresses the capital cost we would have paid if we had ordered the lost product quantity). Hence, depending on which of the two costs is greater we obtain the optimal solution from either the EOQ model or the model allowing shortages.

5 NUMERICAL EXAMPLE

In this section we present an example which depicts a bad case scenario with high holding costs, high interest rates and a low budget. For this numerical example the following data are used: p=20, h=5, I_c=0.1, D=100, s=22, b=8, β=0.7, α=0.5, K=500, B=4000, T_t=0.1, T_p=0.5. We note that for these parametric values we have $2Kh - (s - \omega)^2 (1 - \beta)^2 D = 839.75 > 0$ and $T_t < \dfrac{B}{pD\beta}$.

In Table 1 the optimal solution to the proposed model presented in section 3, as well as for the special case which is derived by setting T_t=0, T_p=0 and by assuming $B \to \infty$ (i.e. EOQ with partial backlogging and no budget constraint), is presented.

	t_1^*	T^*	$TC\left(t_1^*,T^*\right)$	Q^*
Proposed model	1.4	1.6	751.1	154.24
Special case	1.41	1.49	705.8	146.88

In order to highlight the effects of prepayment requirements, tolerance options and budget constraints to the optimal solution of the problem we use the data of the example and we compare the total cost and ordering quantities obtained by the proposed model to the ones obtained by assuming $T_t=0$, $T_p=0$ and $B\rightarrow\infty$. As it is obvious changes in ω, T_t and B do not affect the cost and ordering quantity of the second model, so it can be used as a benchmark to examine the impact of ω, T_t and B to the proposed model. From Figure 2 we observe the significant increase in total cost when ω increases. An increase is also observed in the ordering quantity which is always greater than that obtained by the classical EOQ with partial backlogging (Figure 3). From Figure 4 we observe that total cost decreases when T_t increases however the cost is stabilized when the budget constraint is attained. The increase in T_t leads to increase of the ordering quantity. Obviously, the order quantity is again stabilized when the budget constraint is attained (Figure 5). From Figures 6 and 7 it is observed that the available budget affects significantly the cost and order quantity.
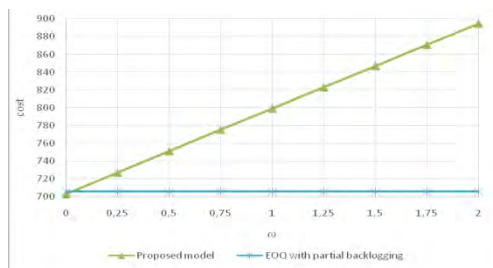
Figure 2: Cost comparisons of proposed model and EOQ with partial backlogging for different values of ω
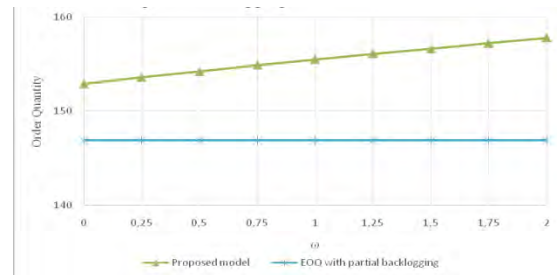
Figure 3: Order quantity comparisons of proposed model and EOQ with partial backlogging for different values of ω
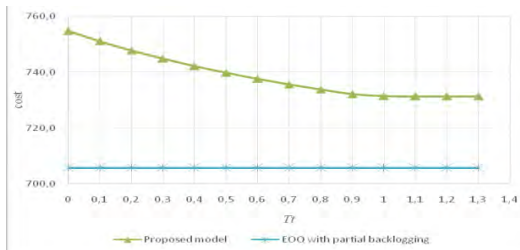
Figure 4: Cost comparisons of proposed model and EOQ with partial backlogging for different values of Tt
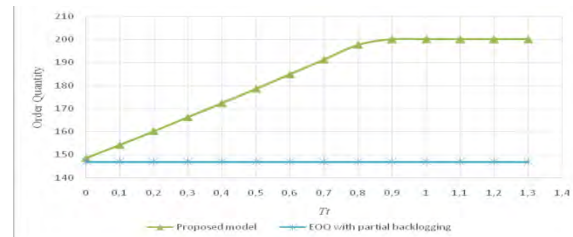
Figure 5: Order quantity comparisons of proposed model and EOQ with partial backlogging for different values of Tt
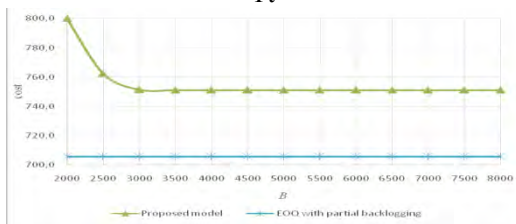
Figure 6: Cost comparisons of proposed model and EOQ with partial backlogging for different values of B
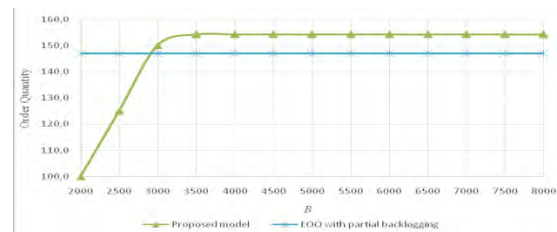
Figure 7: Order quantity comparisons of proposed model and EOQ with partial backlogging for different values of B

6 CONCLUSION

In this paper we examine the effect of financial constraints regarding replenishment, such as budget limitations on the available capital per replenishment and the imposition of prepayment (either full or partial) of the quantity ordered, to the classical EOQ with partial backorders model. On the consumption side, the new concept of market tolerance is introduced, which effectively permits backorders to occur at no-cost for a predetermined time period (the concept is actually analogous to payment credit). From the theoretical analysis, we observe, that the optimal solution depends on the value of the quantity $2Kh - (s-\omega)^2(1-\beta)^2 D$, which reflects the relation between the ordering cost, the holding cost and the lost sales cost. In order to highlight the theoretical results we present a numerical example along with sensitivity analysis. Further research directions could be the simultaneous consideration of trade credit offer and customer tolerance, the assumption of non constant demand and/or time dependent costs.

References

[1] Ahn, J.B., Amiti, M., Weinstein, D. E. 2011. Trade finance and the great trade collapse. American Economic Review: Papers and Proceedings, 101 (3): 298-302.

[2] Goyal, S.K. 1985. Economic order quantity under conditions of permissible delay in payments. Journal of the Operational Research Society, 36: 335-338.

[3] Harris, F.W. 1913.How many parts to make at once. Factory, The Magazine of Management, 10: 135–136.

[4] Montgomery, D.C., Bazaraa, M.S., Keswani, A.K. 1973. Inventory models with a mixture of backorders and lost sales. Naval Research Logistics Quarterly, 20 (2): 255–263.

[5] Park, K.S. 1982. Inventory model with partial backorders. International Journal of Systems Science, 13 (12): 1313–1317.

[6] Pentico, D.W., Drake, M.J. 2009. The deterministic EOQ with partial backordering: A new approach. European Journal of Operational Research, 194:102–113.

[7] Pentico, D.W., Drake, M.J. 2011. A survey of deterministic models for the EOQ and EPQ with partial backordering. European Journal of Operational Research, 214 (2): 179-198.

[8] Rosenberg, D. 1979. A new analysis of a lot-size model with partial backlogging. Naval Research Logistics Quarterly, 26 (2): 349–353.

[9] Shah NH., Jani MY., Chaudhari U. 2017. Optimal replenishment time for retailer under partial upstream prepayment and partial downstream overdue payment for quadratic demand. Mathematical and Computer Modelling of Dynamical Systems: 1-11.

[10] Taleizadeh, A.A., Pentico, D.W., Jabalameli, M.S., Aryanezhad, M. 2013. An economic order quantity model with multiple partial prepayments and partial backordering. Mathematical and Computer Modelling, 57: 311–323.

[11] Taleizadeh, A.A. 2014. An EOQ model with partial backordering and advance payments for an evaporating item. International Journal of Production Economics, 155: 185–193.

[12] Tavakoli S, Taleizadeh A.A. 2017. An EOQ model for decaying item with full advanced payment and conditional discount. Annals of Operations Research: 1–22

[13] Wu J, Teng J-T, Chan Y-L. 2017. Inventory policies for perishable products with expiration dates and advance-cash-credit payment schemes. International Journal of Systems Science: Operations & Logistics: 1-17.

[14] Zhang, Q., Tsao, Y-C., Chen, T-H. 2014. Economic order quantity under advance payment. Applied Mathematical Modeling, 38 (24): 5910-5921.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 4:
Location and Transport, Graphs and their Applications

BUS DRIVER SCHEDULING – HOW THE PROBLEM HAS CHANGED WITH IMPROVEMENTS IN COMPUTING

Sarah Fores
3 Woodlea Court, Meanwood, Leeds
LS6 4SL, UK
sarahfores@gmail.com

Abstract: Historically the bus driver scheduling problem was solved manually by experts in their field; in fact this is still the case in some companies. Even with many more limitations compared to today, the introduction of the computer to this field allowed alternative solutions to be produced more quickly. The early software was based on heuristics, using the expert knowledge observed in the manual scheduling process. Techniques such as integer linear programming further improved the solutions in terms of costs and time taken to produce schedules, especially when constraints changed and companies wanted to analyse 'what if' scenarios. The driver scheduling problem is NP-hard and this lends itself to other innovative techniques being explored, such as metaheuristics and constraint programming. Ultimately, though, a 'black box' software solution is naïve in that it is often impossible to truly reflect the problem and specific intricacies of a company without user involvement. The risk, then, is the vicious circle of reliance on a computerised system leading to fewer expert schedulers who truly understand what constitutes a 'good' solution, regardless of what the software produces. Computers therefore now have a role to play as a decision support tool, with graphical interface to help schedulers in their design as well as understanding of the constraints. The flexibility of mobile apps for drivers is also of interest to companies. The opportunities to improve the algorithmic core of the solver are perhaps now less important to companies who do not prioritise the saving of a small amount over having a system that is flexible and usable. Nevertheless algorithms should be further adapted as compute power continues to improve. There are now many commercial software systems for bus driver scheduling, some even claiming to give 'optimal' solutions, and most of the successful ones are still based on a set covering or set partitioning model using column generation to allow a vast search space to be explored. Using experience of a driver scheduling system developed at the University of Leeds as a basis we can explore some of the developments of the computerised systems and issued faced in real life scheduling.

Keywords: scheduling, integer linear programming, applications.

INNOVATIVE PROJECTS SCHEDULING WITH NON-RENEWABLE RESOURCES ON THE BASIS OF DECISION PROJECT GRAPHS

Helena Gaspars-Wieloch
Poznan Universtity of Economics and Business, Department of Operations Research
Al. Niepodleglosci 10, 61-875 Poznan, Poland
E-mail: Helena.gaspars@ue.poznan.pl

Abstract: In this paper we present a novel scenario-based DPG (decision project graph – combination of deterministic and stochastic networks) rule which takes into consideration possible alternative tasks (differing from other alternative activities in respect of times, non-renewable resources and sets of successors), possible scenarios and the decision-maker's attitude towards risk. The procedure is especially designed for totally new (innovative) projects where it is complicated to estimate probabilities of particular events since no historical data are available.

1 INTRODUCTION

In this contribution we would like to analyze the process of scheduling totally new projects, i.e. innovative projects (e.g. new product development project). In such circumstances the use of deterministic networks may be inappropriate as many factors concerning the project execution are not exactly known at the planning stage. Even the application of random variables describing activity durations, costs and resources along with a deterministic project structure might be hampered since, in the case of innovative projects, the likelihood (understood as frequency) cannot be known due to the lack of sufficient historical data. Therefore, in this research, we would like to refer to decision project graphs (DPGs), also called decision critical path method networks (DCPMN), which were originally proposed by [8, 21], designed for project planning and management, and related to the concept of multiple choices at alternative nodes when decision-making is of deterministic nature. In connection with the existence of many uncertain parameters describing the project, scientists started to develop DPGs under uncertainty [32, 38] where it is assumed that the probabilities of the occurrence of particular states of nature (scenarios, events) may be estimated. In this paper, we analyze a higher degree of uncertainty, i.e. situations where the likelihood cannot be calculated (lack of historical data concerning previous similar projects), but may be replaced by probability-like quantities. The new scenario-based DPG rule presented in the article takes into account possible scenarios and the decision-maker's attitude towards risk. It considers both uncertain times and non-renewable resources (liquid assets, raw materials, natural materials etc.). The procedure may support both reactive and proactive project management.

The paper is organized as follows. Section 2 presents DPGs and their application to innovative projects. Section 3 describes a novel DPG rule for totally new projects. Section 4 provides an illustrative example. Conclusions are gathered in the last part.

2 THE USE OF DECISION PROJECT GRAPHS FOR INNOVATIVE PROJECTS

Decision project graphs constitute a combination of deterministic (DAN, ang. Deterministic Analysis Network) and stochastic (GAN, ang. Generalized Analysis Network) networks. DPGs are similar to deterministic networks as they do not contain cycles (however they are merely suitable for simple projects with an explicitly defined technology [37]). On the other hand, they resemble stochastic networks since only some tasks (jobs, activities) considered in the graph are finally chosen and executed (with a non-negative probability), and additionally,

they are appropriate for new projects where changes are possible during the project execution [30]. DPGs contain deterministic and alternative nodes connected with subsets DET(N) and ALT(N), respectively. The former (related to *deterministic job sets*) are the beginning of deterministic tasks (all of them are supposed to be executed), the latter (related to *alternative job sets*) assume that at least one activity from a given set must be performed (note that after the final selection of some alternative tasks and the removal of the remaining ones the graph has to be still connected and can be converted to an ordinary deterministic graph), [35]. Particular alternative tasks may differ from other activities belonging to this set in respect of times, costs, resources and even sets of successors. DPGs can be presented by means of AON (ang. Activities on nodes) or AOA (ang. Activities on arcs) techniques. The first one uses nodes for jobs, and arcs for precedence relations. The second one represents tasks with the aid of arcs and events on the basis of nodes. Here, we are going to apply AOA. The DPG structure may be defined as follows [16]. Each graph contains a set of job sets $S=\{S_1, \ldots, S_i, \ldots, S_{n-1}\}$, where n denotes the number of all nodes (events) in the graph. Besides nodes belonging to subsets DET(N) and ALT(N), the network should consist of one end node (belonging to subset END(N)) with predecessors and no successors. The number of tasks connected with a job set S_i is equal to $k(i)$. We can distinguish three types of tasks: deterministic, alternative and dummy activities belonging to subsets DET(A), ALT(A) and DUM(A). Dummy activities (given by dotted arrows) are only used in the AOA technique. Their time is equal to zero since they just guarantee each necessary precedence relation in the graph [29]. In the literature, one can find many diverse scheduling optimization models based on DGPs, e.g. 1) the AOA model minimizing the project completion time where one or more activities belonging to a given alternative job set have to be executed [16]; 2) the model for project completion cost minimization with a desired project completion time [7, 8, 35]; 3) the model minimizing direct project costs subjected to a deadline constraint [19], 4) the time-cost trade-off model referring to dynamic programming [22].

DPGs have many advantages, but their original version with deterministic parameters does not meet the needs of innovative projects characterized by uncertain durations, costs and resources, and requiring project risk management skills. "Uncertainty" and "risk" are interpreted in different ways in the literature [11, 12, 13, 14, 15, 24, 25, 26, 28, 39, 41, 43]. Here, risk is related to the possibility that some bad or other than predicted circumstances will happen, while uncertainty involves all situations with non-deterministic parameters, hence decisions may lead to different consequences and the probability of scenarios is known or not. In the latter case some probability-like quantities may be often estimated and applied. The description of the project by means of uncertain parameters (interval values, scenarios, probabilities, fuzzy numbers etc.) has been already made by numerous researchers, see e.g. [2, 4, 16, 27, 31]. Uncertainty can be connected with activity characteristics (times, costs, resources) or the network logic (project structure). The most well-known procedure is PERT (Program Evaluation and Review Technique) which is a three-point estimation technique since it collects optimistic, the most likely and pessimistic duration estimates for all tasks [27]. PERT is used to find the expected length of the critical path. Monte Carlo simulation can also estimate the criticality of particular jobs and paths. On the other hand, among stochastic networks it is worth mentioning GERT (ang. Graphical Evaluation and Review Technique) and GERTS (ang. Graphical Evaluation and Review Technique Simulation) which allow a probabilistic treatment of both network logic and estimation of activity duration [33, 34, 44]. Due to some drawbacks related to methods mentioned above (e.g. independence of duration distributions, negligence of subcritical paths in PERT, identical duration distributions for repeated tasks within loops in GERT), further procedures were developed [10]. For instance [32] suggest relating uncertainty to different project scenarios rather than individual tasks since, in their opinion, uncertainty may be connected with diverse

sequences or combinations of events occurring or not occurring. The procedure enables computing expected criticalities of each tasks and analyzing projects with dependent activity durations. [9] propose a method for situations where activities fail during the project execution and [6] apply stochastic dynamic programming to incorporate the risk of activity failure and the possible pursuit of alternative technologies. Other interesting approaches combine stochastic project networks with net present value optimization or resource constraints [3, 5, 23, 36]. Now, let us see how uncertainty has been incorporated into DPGs. In CAAN (ang. Controlled Alternative Activity Network), proposed by [17], the lexicographical scanning and the use of stochastic and deterministic events allow the project manager to choose the optimal direction in the course of the project execution. GAAN (ang. Generalized Alternative Activity Network Model), developed by [18], applies lexicographical method and discrete optimization, considers three types of activities (deterministic, alternative stochastic and alternative deterministic) and diverse attitudes towards risk. CCANM (ang. Controlled Cyclic Alternative Network Model) [42] integrates CANN and the cyclic GERT (with loops and different logical relations).

Procedures presented above can be applied to project optimization when some factors concerning the project are uncertain. Nevertheless, they impose a concrete probability distribution or require the exact knowledge of the likelihood of the occurrence of each scenario, which, in the case of innovative projects, may be inappropriate. Objective and subjective probabilities might be extremely difficult to estimate for the project manager as historical data about similar or identical projects, already executed, do not exist, and the majority of circumstances accompanying a totally new project are unknown. Additionally, the optimization of an innovative project should be treated as a one-shot decision [20], and the effects of that project - as the occurrence of only one event (other events will not have the possibility to occur). Hence, according to [40], the mathematical probability understood as frequency cannot be computed in that case. Therefore, we would like to suggest a novel approach which does not require any information about the likelihood of particular events.

3 A SCENARIO-BASED DECISION RULE FOR INNOVATIVE PROJECTS

The assumptions adopted in the novel scenario-based DPG rule are as follows. The innovative project can be presented by means of a decision project graph (AOA). Four categories of job sets may occur in the network: 1) set of deterministic activities DET(A), 2) set of deterministic activities with scenarios DET(A)S, 3) set of alternative activities ALT(A), 4) set of alternative activities with scenarios ALT(A)S. All tasks from DET(A) and DET(A)S have to be executed. One task from each set ALT(A) and ALT(A)S has to be performed. The number of events for particular tasks from a given set ALT(A)S may be different. Durations and non-renewable resources (estimated by experts) are uncertain for activities from sets ALT(A)S and DET(A)S. For each activity from sets DET(A)S and ALT(A)S the project manager declares his/her coefficient of optimism (β) which satisfies conditions: $\alpha, \beta \in [0,1]$ and $\alpha + \beta = 1$, where α denotes the pessimism coefficient (α is close to 1 for extreme pessimists and β is close to 1 for radical optimists). The goal is to minimize the project completion time within available non-renewable resources (problem I) or to minimize non-renewable resources subjected to a deadline constraint (problem II). Activity durations and resources (materials, liquid assets for salaries, insurance, rental etc.) are treated as dependent criteria. All resources are expressed in monetary units. Limited resources can be used in any moment of the project execution. The obtained optimal solution is used merely for one project since after the execution of that project some parameters in a new similar project may change (scenarios, times, resources, network structure, attitude towards risk etc.).

The new DPG rule for innovative projects and problem I contain the following steps:

Step 1. Present the project network structure by means of a DPG and define all parameters concerning particular activities (activity type, time and resources).

Step 2. Declare the coefficient of optimism common for the whole project or separate for each task with scenarios (DET(A)S) or group of tasks with scenarios (ALT(A)S). Choose weights w^t and w^r for the time and resource criterion. The weights ought to describe the importance of each target and should satisfy: $w^t + w^r = 1$. Normalize time and resource evaluations, separately for each task with events (or group of tasks with events) according to:

$$e(n)_{ij}^y = \frac{\max\limits_{y=1,\ldots,s}\{e_{ij}^y\} - e_{ij}^y}{\max\limits_{y=1,\ldots,s}\{e_{ij}^y\} - \min\limits_{y=1,\ldots,s}\{e_{ij}^y\}} \quad or \quad e(n)_{ij}^y = \frac{\max\limits_{\substack{y=1,\ldots,s \\ z=1,\ldots,a}}\{e_{ij}^{yz}\} - e_{ij}^y}{\max\limits_{\substack{y=1,\ldots,s \\ z=1,\ldots,a}}\{e_{ij}^{yz}\} - \min\limits_{\substack{y=1,\ldots,s \\ z=1,\ldots,a}}\{e_{ij}^{yz}\}}$$

(1a), (1b)

where $e(n)_{ij}^y$ denotes the normalized time (t) or resource (r) evaluation for task $<i,j>$ and event y; z signifies a given alternative. Compute for each scenario the SAW (Simple Additive Weighting) value: $SAW_y = w^t \cdot t(n)_{ij}^y + w^r \cdot r(n)_{ij}^y$. For each task with events keep the scenario fulfilling Equation (2) and the time and resource connected with it (i.e. $t(\beta)_{ij}$ and $r(\beta)_{ij}$).

$$SAW(\beta)_{ij} = \beta \cdot (SAW_{max,ij} - SAW_{min,ij}) + SAW_{min,ij}$$

(2)

where $SAW_{max,ij}$ and $SAW_{min,ij}$ signify the maximal and minimal SAW values in the set of events connected with a given task. If such a scenario does not exist, find events $S^{min}_{y,ij}$ and $S^{max}_{y,ij}$ satisfying Equations (3) and (4) and calculate weighted average times $t(w)_{ij}$ and resources $r(w)_{ij}$ for each activity according to Equations (5)-(6) where $t^{min}_{y,ij}$ and $t^{max}_{y,ij}$ denote the activity time related to scenarios $S^{min}_{y,ij}$ and $S^{max}_{y,ij}$ (the interpretation of $r^{min}_{y,ij}$ and $r^{max}_{y,ij}$ is analogous). $SAW^{min}_{y,ij}$ and $SAW^{max}_{y,ij}$ signify SAW values connected with events $S^{min}_{y,ij}$ and $S^{max}_{y,ij}$. For step (3) replace $t(\beta)_{ij}$ $(r(\beta)_{ij})$ and $t(w)_{ij}$ $(r(w)_{ij})$ with t^*_{ij} (r^*_{ij}).

$$S^{min}_{y,ij} = \arg \min_{S_{y,ij}|SAW_{y,ij}<SAW(\beta)_{y,ij}} (SAW(\beta)_{y,ij} - SAW_{y,ij})$$

(3)

$$S^{max}_{y,ij} = \arg \min_{S_{y,ij}|SAW_{y,ij}>SAW(\beta)_{y,ij}} (SAW_{y,ij} - SAW(\beta)_{y,ij})$$

(4)

$$t(w)_{ij} = t^{min}_{y,ij} \cdot \left(\frac{SAW^{max}_{y,ij} - SAW(\beta)_{y,ij}}{SAW^{max}_{y,ij} - SAW^{min}_{y,ij}} \right) + t^{max}_{y,ij} \cdot \left(\frac{SAW(\beta)_{y,ij} - SAW^{min}_{y,ij}}{SAW^{max}_{y,ij} - SAW^{min}_{y,ij}} \right)$$

(5)

$$r(w)_{ij} = r^{min}_{y,ij} \cdot \left(\frac{SAW^{max}_{y,ij} - SAW(\beta)_{y,ij}}{SAW^{max}_{y,ij} - SAW^{min}_{y,ij}} \right) + r^{max}_{y,ij} \cdot \left(\frac{SAW(\beta)_{y,ij} - SAW^{min}_{y,ij}}{SAW^{max}_{y,ij} - SAW^{min}_{y,ij}} \right)$$

(6)

Step 3. Solve the non-linear optimization model (7)-(16). If any changes concerning activity parameters occur before (proactive management) or during (reactive management) the project execution, include those modifications in the model and solve it.

$$t_n \to \min$$

(7)

$$R_u \cdot t_n + \sum_{<i,j>\in DET(A)} r_{ij} + \sum_{<i,j>\in DET(A)S} r^*_{ij} + \sum_{<i,j>\in ALT(A)} r_{ij}x_{ij} + \sum_{<i,j>\in ALT(A)S} r^*_{ij}x_{ij} \leq R$$

(8)

$$t_j \geq t_i + t_{ij} \qquad <i,j> \in DET(A)$$

(9)

$$t_j \geq t_i + t^*_{ij} \qquad <i,j> \in DET(A)S$$

(10)

$$t_j \geq (t_i + t_{ij}) \cdot x_{ij} \qquad <i,j> \in ALT(A)$$

(11)

$$t_j \geq (t_i + t_{ij}^*) \cdot x_{ij} \qquad <i,j> \in ALT(A)S \qquad (12)$$

$$\sum_{<i,j> \in S_i} x_{ij} = 1 \qquad i \in ALT(N) \qquad (13)$$

$$t_j \geq t_i \qquad <i,j> \in DUM(A) \qquad (14)$$

$$t_i, t_j \geq 0 \qquad i = 1,..n-1; j = 2,...,n \qquad (15)$$

$$x_{ij} \in \{0,1\} \qquad <i,j> \in ALT(A) \qquad (16)$$

where: n – number of nodes in the graph; t_n – project completion time (continuous variable); R_u – unit resources connected with the whole project (not with particular activities), e.g. insurance, management, taxes; rent; r_{ij} (r^*_{ij}) - resources used to execute activity $<i,j>$; x_{ij} – binary variable connected with alternative activity $<i,j>$ (it is equal to 1, when this task is done; otherwise it is equal to 0); R – available resources; t_i, t_j – times of event i and j (continuous variables); t_{ij} (t^*_{ij}) – time of activity $<i,j>$ (parameter); ALT(N) – set of alternative nodes (alternative job sets); DUM(A) – set of dummy activities. When more than one activity from a given job set is supposed to be performed, then Equation (13) can be replaced with another constraint [35]. Note that the optimal value of variable t_i belongs always to interval $[t^I_i, t^{II}_i]$, where t^I_i and t^{II}_i denote the earliest and the latest time of event i [1]. In the case of problem II, it suffices to transform formula (8) into an objective function and to modify the end of Equation (7) by introducing inequality $\leq$ and T (deadline).

4 ILLUSTRATION

Let us illustrate the procedure presented in Section 3 for problem I. Data concerning a fictitious project are gathered in Table 1. Figure 1 presents its structure (step 1). The unit project cost is equal to 2 thousands of Euros per day and available resources amount to 115 thousands of Euros. Relations between activities are described by means of *FS* (finish to start, i.e. when the predecessor is completed, the successor can start). We assume (step 2) that the project manager is a moderate pessimist ($\beta=0.35$ for activities B and C, and $\beta=0.45$ for jobs G, H, I). Weights w^t and w^r equal 0.4 and 0.6. Times and resources are normalized in Table 1 (in brackets). *SAW* values are as follows. For activity B: 0.86, 0.88, 0.84, 0.28, 0.0; for C: 0.6, 0.66, 0.46, 0.4; for G: 0.52, 0.0, 1.0; for H: 0.56, 0.29, 0.12, 0.74; for I: 1.0, 0.89, 0.39.

Table 1: Activity parameters. Source: prepared by the author.

Tasks	Set	Time (in days)	Resources (in thousands of Euros)
A	DET(A)	10	1
B	DET(A)S	S_1.3(1.0); S_2.6(0.82); S_3.10 (0.59); S_4.18(0.12); S_5.20(0.0)	S_1.2(0.77); S_2.1.8(0.92); S_3.1.7(1.0); S_4.2.5(0.38); S_5.3(0.0)
C	ALT(A)S	S_1.13(0.0); S_2.4(0.9); S_3.6(0.7); S_4.3(1.0)	S_1.3(1.0); S_2.3.5(0.5); S_3.3.7(0.3); S_4.4(0.0)
D	ALT(A)	7	2
E	DET(A)	4	5
F	DET(A)	9	4.5
G	DET(A)S	S_1.11(0.5); S_2.17(0.0); S_3.5(1.0)	S_1.5.2(0.53); S_2.6(0.0); S_3.4.5(1.0)
H	ALT(A)S	S_1.12(0.87); S_2.20(0.62); S_3.30(0.31); S_4.15(0.78)	S_1.3(0.35); S_2.4(0.06); S_3.4.2(0.0); S_4.1.8(0.71)
I	ALT(A)S	S_1.8(1.0); S_2.10(0.94); S_3.40(0.0)	S_1.0.8(1.0); S_2.1.3(0.85); S_3.2(0.65)
D_1	DUM(A)	0	0
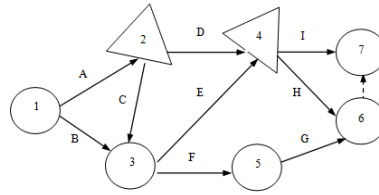
Figure 1: Decision network project (example). Source: prepared by the author [16].

$SAW(\beta)_{ij}$ for particular tasks with scenarios equal 0.31, 0.49, 0.45, 0.4, 0.66, respectively. Now, we ought to find scenarios $S^{min}_{y,ij}$ and $S^{max}_{y,ij}$ for each activity since there is no event satisfying Equation (2): B – S_4 and S_3; C – S_3 and S_1; G – S_2 and S_1; H – S_2 and S_1; I – S_3 and S_2. Weighted average times and resources are as follows: B – 17.55 and 2.46; C – 7.55 and 3.55, G – 11.81 and 5.31, H – 16.68 and 3.59, I – 22.33 and 1.61. Due to page limitations we do not present the optimization model (step 3). It contains 14 constraints, 1 objective function, 7 continuous variables and 4 binary variables. The optimal solution obtained by means of SAS/OR can be represented by the following results: t_1=0; t_2=10; t_3=17.55; t_4=22.55; t_5=22.05; t_6=39.23; t_7=39.23; x_C=x_{23}=0; x_D=x_{24}=1; x_H=x_{46}=1; x_I=x_{47}=0, which signifies that the estimated project completion time is equal to almost 40 days (objective function) and that the project manager should select and perform alternative activities D and H. For this plan estimated resources equal 102.31 thousands of Euros. Let us remember that the schedule described above has been generated for a moderate pessimist. For another project manager the suggested solution would be probably different.

5 CONCLUSIONS

The contribution presents a novel scenario-based decision project graph rule especially designed for innovative projects. Nevertheless, the procedure can also be applied to any other project where the project manager does not want to make use of historical data or expects some significant changes in comparison to previous similar projects. The approach refers to the SF+AS method presented in [13] due to the use of the scenario forecasting stage. The bi-criteria (time and non-renewable resources) method is assisted with a non-linear optimization model containing continuous and binary variables. The model can be easily solved with the use of SAS/OR, MiniZinc, CPLEX or "R". The main advantages of the procedure are as follows. Firstly, it allows taking into consideration uncertain information concerning activity times and non-renewable resources. Secondly, it gives the opportunity to consider more than three events for each task (in contradiction to PERT) and it allows inserting a different number of scenarios for particular activities. Thirdly, it only uses probability-like quantities on the basis of the DM's attitude towards risk measured by the coefficient of optimism (the probability estimation is not required) and it enables analyzing different kinds of decision makers (risk-neutral, risk-seeking, risk-aversion). Fourthly, it allows considering four categories of activities (deterministic/alternative structure; deterministic/scenario times), which is useful in problems taking place in a practical environment. Fifthly, it is attractive for passive project managers since only the coefficient of optimism (which can even be common for each activity with scenarios) and weights for both criteria are supposed to be declared. Other parameters are estimated by experts. In the future it would be desirable to formulate a scenario-based DPG model considering both limited renewable and nonrenewable resources.

References

[1] Anholcer, M., Gaspars-Wieloch, 2011. Efficiency analysis of the Kaufmann and Desbazeille algorithm for the deadline problem. Operations Research and Decisions, 21(1): 5–18.

[2] Azeem, S.A.A., Hosny, H.E., Ibrahim, A.H. 2014. Forecasting project schedule performance using probabilistic and deterministic models. HBRC Journal, 10(1): 35–42.

[3] Benati, S. 2006. An optimization model for stochastic project networks with cash flows. Computational Management Sciences, 3(4): 271–284.

[4] Błaszczyk, P., Błaszczyk, T., Kania, M.B. 2013. Project scheduling with fuzzy cost and schedule buffers. Lecture Notes in Electrical Engineering, 170: 375–388.

[5] Creemers, S., Leus, R., Lambrecht, M. 2010. Scheduling Markovian PERT networks to maximize the net present value. Operations Research Letters, 38(1): 51–56.

[6] Creemers, S., De Reyck, B., Leus, R. 2013. Project planning with alternative technologies in uncertain environments. KU Leuven. Faculty of Economics and Business.

[7] Crowston, W.B. 1970. Decision CPM: Network reduction and solution. Operational Research Quaterly, 21(4): 435–452.

[8] Crowston, W.B., Thompson, G.L. 1967. Decision CPM: a method for simultaneous planning, scheduling and control of projects. Operations Research, 15: 407–426.

[9] De Reyck, B., Leus, R. 2008. R&D-project scheduling when activities may fail. IIE Transactions, 40(4): 367–384.

[10] Dodin, B. 2006. A practical and accurate alternative to PERT. Perspectives in Modern Project Scheduling. International Series in Operations Research and Management Science, 92: 3–23.

[11] Dubois, D., Prade, H. 2012. Gradualness, uncertainty and bipolarity: making sense of fuzzy sets. Fuzzy Sets and Systems, 192: 3–24.

[12] Gaspars-Wieloch, H. 2014. Modifications of the Hurwicz's decision rules. Central European Journal of Operations Research, 22(4): 779–794.

[13] Gaspars-Wieloch, H. 2015. A decision rule supported by a forecasting stage based on the decision maker's coefficient of optimism. Central European Journal of Operations Research, 23(3): 579–594.

[14] Gaspars-Wieloch, H. 2016a. Resource allocation under complete uncertainty – case of asymmetric payoffs. Organization and Management (Organizacja i Zarzadzanie), 96: 247–258.

[15] Gaspars-Wieloch, H. 2016b. Newsvendor problem under complete uncertainty: a case of innovative products. Central European Journal of Operations Research. http://dx.doi.org/10.1007/s10100-016-0458-3

[16] Gaspars-Wieloch, H. 2017. Innovative project scheduling with scenario-based decision project graphs. Contemporary Issues in Business, Management and Education'2017, http://dx.doi.org/cbme.2017.078

[17] Golenko-Ginzburg, D. 1988. Controlled activity networks for project management. European Journal of Operations Research, 37(3): 336–346.

[18] Golenko-Ginzburg, D., Blokh, D. 1997. A generalized activity network model. Journal of the Operational Research Society, 48(4): 391–400.

[19] Grudzewski, W. 1985. Badania operacyjne w organizacji i zarządzaniu [Operations research in organization and management]. Warsaw: PWN. (in Polish)

[20] Guo, P. 2011. One-shot decision theory. IEEE Transactions on Systems, Man and Cybernetics. Part A, 41(5): 917–926.

[21] Hastings, N.A.J., Mello, J.M.C. 1978. Decision networks. Chichester, New York: Willey.

[22] Hindelang, T.J., Muth J.F. 1979. A dynamic programming algorithm for Decision CPM networks. Operations Research, 27(2): 225–241.

[23] Igelmund, G., Radermacher, F.J. 1983. Preselective strategies for the optimization of stochastic project networks under resource constraints. Networks, 13(1): 1–28.

[24] Kaplan, S., Barish, N.N. 1967. Decision-making allowing for uncertainty of future investment opportunities. Management Science, 13(10): 569–577.

[25] Kmietowicz, Z.W., Pearman, A.D. 1984. Decision theory, linear partial information and statistical dominance. Omega, 12: 391–399.

[26] Knight, F. H. 1921. Risk, uncertainty, profit. Hart. Boston MA. Schaffner & Marx. Houghton Mifflin Co.

[27] Malcolm, D.G., Rosenboom, J.H., Clark, C.E., Fazar, W. 1959. Application of a technique for research and development program evaluation. Operations Research, 7(5): 646–669.

[28] Merigo, J.M. 2015. Decision-making under risk and uncertainty and its application in strategic management. Journal of Business Economics and Management, 2015(1): 93–116.

[29] Moder, J.J., Phillips, C.R. 1964. Project Management with CPM and PERT. New York: Reinhold Publishing Corporation.

[30] Neumann, K., Steinhardt, U. 1979. GERT network and the time-oriented valuation of projects. Lecture Notes in Economics and Mathematical Systems, 172.

[31] Okmen, O., Oztas, A. 2014. Uncertainty evaluation with fuzzy schedule risk analysis model in activity networks of construction projects. Journal of South African Institution of Civil Engineering, 56(2).

[32] Pollack-Johnson, B., Liberatore, M.J. 2005. Project planning under uncertainty using scenario analysis. Project Management Journal, 36(1): 15–26.

[33] Pritsker, A.A.B., Happ, W.W. 1966. GERT: Graphical evaluation and review technique — part 1. Fundamentals. Journal of Industrial Engineering, 17: 267–274.

[34] Pritsker, A.A.B. 1979. Modeling and analysis using Q-GERT Networks. 2nd Edition. Wiley.

[35] San Cristobal, M.J.R. 2015. Management science, operations research and project management: modeling, evaluation, scheduling, monitoring. Gower Applied Business Research.

[36] Sobel, M.J., Szmerekovsky, J.G., Tilson, V. 2009. Scheduling projects with stochastic activity duration to maximize expected net present value. European Journal of Operational Research, 198(1): 697–705.

[37] Spinner, M. 1981. Elements of project management: plan, schedule and control. Prentice-Hall, Englewood Cliffs, New Jersey.

[38] Thompson, G.L. 1968. CPM and DCPM under risk. Naval Research Logistics, 15(2): 233–239.

[39] Trzaskalik, T. 2008. Wprowadzenie do badan operacyjnych z komputerem [Introduction to operations research with computer]. (2nd ed.). Warsaw: Polskie Wydawnictwo Ekonomiczne. (in Polish)

[40] von Mises, L. 1949. Human action. A treatise on economics. The Ludwig von Mises Institute. Auburn. Alabama.

[41] von Neumann, J., Morgenstern, O. 1944. Theory of games and economic behavior. Princeton University Press. Princeton. New York.

[42] Voropayev, V.I., Gelrud, Y.D., Golenko-Ginzburg, D. 2013. Decision making in controlled cyclic alternative network projects with deterministic branching outcomes. PM World Journal II(IX).

[43] Ward, S., Chapman, C. 2003. Transforming project risk management into project uncertainty management. International Journal of Project Management, 21: 97–105.

[44] Wiest, J.D., Levy, F.K. 1977. A management guide to PERT/CPM with GERT/PDM/DCPM and other networks. (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

ESTIMATION OF THE EFFECTIVE NUMBER OF CENTERS WITH REGARD TO THE DISTANCES IN MUNICIPALITIES

Marta Janáčková
University of Žilina, Department of Applied Mathematics
Univerzitná 8215/1, 01026 Žilina, Slovak Republic
Marta.Janackova@fstroj.uniza.sk

Alžbeta Szendreyová
University of Žilina, Department of Mathematical Methods and Operations Research
Univerzitná 8215/1, 01026 Žilina, Slovak Republic
Alzbeta.Szendreyova@fria.uniza.sk

Abstract: In the paper, we discuss ways of estimating a suitable number of the Emergency Medical Service (EMS) centres in a serviced area to satisfy demand of public for emergency service. As there is no direct relation between service quality and cost of a center, we formulate a hypothesis based on graph of service quality function depending on increasing service centre number. The hypothesis is tested on benchmarks derived from current EMS systems running in self – governing regions of the Slovak Republic and compared to the current centre deployments, which developed in the self-governing regions during two recent decades. As the service quality evaluation depends also on the data model of the underlying transportation network, two different models are studied to verify the hypothesis. The first model considers zero distance inside a municipality whereas the second one estimates the inside distances by an average non-zero value. Based on the results, we will try to answer two questions. We wonder if the hypothesis is able to estimate the appropriate number of service centers. Secondly, we want to know, which type of the distance matrix of the transportation network enables better estimation of the suitable number of centres.

Keywords: emergency service system, effective number of service centres, distance inside municipality.

1 INTRODUCTION

A design of the public service system is a task of strategic decision-making. The decisions on the number and location of the service centers, as well as on the assignment of the customers to the service providers, are already well managed in present. The problem can be solved as a location-allocation problem [1], [5]. The radial approach [2] is also known and commonly used in the case of large networks. In our contribution, we deal with the design of the Emergency Medical Service (EMS) system, which is a special case of the public service system.

Supervision of the Emergency Medical Service (EMS) is in competence of the individual self-governing regions of Slovak Republic. Self-government of a region has usually contracts with several EMS providers. They decide whether the EMS centre in the region will be cancelled or, on the contrary, will be added. The aim of the appropriate location and of the appropriate number of the EMS centers is to ensure the availability of the EMS for the inhabitants in the region. Two conflicting requirements have to be taken into account in determining of the appropriate number of EMS centers. The centers are supposed to be enough to satisfy the live-saving requirements as well as possible. On the other side, the public resources are restricted and this limits the number of centers.

The basic parameter of the availability of the service is the distance between EMS centre (provider) and the customer (inhabitant). This influences creating a minimum system of centres. Even if occurrence of demand for service at given place is random, we assume that the demands are proportional to the number of inhabitants. With a given number of centers, it is possible to provide a service only at a certain level. Addition of one centre ever shortens

distances among some customers and the assigned EMS centre. If we assess the availability of the service by the average distance among EMS centre and the associated customer, then each addition of an EMS centre improves the total availability of service. A question remains, for which number of the centers we still consider such an improvement of the EMS availability so significant that it is acceptable to add another centre.

2 EFFECTIVE NUMBER OF THE CENTERS

Let us consider the task of optimal location of p centres, which provide inhabitants with service. In our case, each municipality is a candidate for centre location. The individual inhabitants (applicants for the EMS) are concentrated in the municipalities. The distances among municipalities are known. Each municipality can therefore be considered as a customer. The quantity of the requirements of each municipality is given by the number of its inhabitants. Minimization of the sum of the distances among inhabitants and assigned centers is the optimization criterion for the location of the centres.

The model ranges and coefficients have the following meanings. Symbol I denotes the set of possible facility locations (EMS centers), J denotes the set of customers (municipalities), d_{ij} is the distance between center $i \in I$ and customer $j \in J$, b_j represents the demand of the customer $j \in J$ (number of inhabitants in the municipality) and p is the given number of centers. The decision on locating or not locating a center i at a place $i \in I$ is modeled by a variable y_i, which takes the value of 1, if the center is located at place i and it takes the value of 0 otherwise. The decision on allocation of the customer from node $j \in J$ to the center at the place $i \in I$ is modeled by a variable z_{ij}. It takes the value of 1 if the customer j will be served from the center i and takes the value of 0 otherwise. The model of the problem takes the form:

$$Minimize \ \sum_{i \in I} \sum_{j \in J} b_j d_{ij} z_{ij} \tag{1}$$

$$Subject \ to \ \sum_{i \in I} z_{ij} = 1 \quad for \ j \in J \tag{2}$$

$$z_{ij} \leq y_i \quad for \ i \in I, j \in J \tag{3}$$

$$\sum_{i \in I} y_i = p \tag{4}$$

$$z_{ij} \in \{0,1\} \quad for \ i \in I, j \in J \tag{5}$$

$$y_i \in \{0,1\} \quad for \ i \in I \tag{6}$$

The problem is known as the weighted p-median problem. Solution of the problem (1)-(6) gives such deployment of p service centres, which minimizes the total travelled distance necessary for public demand satisfaction and, this way, maximizes service availability.

The objective function value (1) represents the sum of the travelled distances, if we assume that satisfaction of each unit of a demand b_j requires trip from the assigned service centre i to location j. This objective function value (OV) can be measured in kilometres but, if necessary, it is possible to use a matrix of the time availability among municipalities. Then, the objective value would be measured, for example, in minutes. Alternatively, OV can be given in person-kilometres or person-minutes etc.

When we subsequently solve the problem (1)-(6) for $p+1$ centres in constraint (4), the objective value obviously decreases. Experiments have shown that the objective values decrease depending on the increasing value of p, and the differences among objective values also decrease for two consecutive values of p. This means that each next added centre will always bring a smaller improvement, i.e. smaller benefit for the provided service. We return

to the question: How do we determine the measure of the effectiveness of adding centres, when no counter balance is given? The minimal value of OV can be achieved for $p=|I|$.

We pronounce the hypothesis on effective number of service centres in serviced region: Let p_{ef} be such a number of centres, for which the decrease of the objective value OV between p_{ef} and $p_{ef}+1$ is equal or close to the average decrease of the objective value on the given range $< p_{min}, p_{max} >$. The average decrease of the objective value for the given range is a slope of the line connecting maximal and minimal value OV on the given range $< p_{min}, p_{max} >$ (see Figure 1).
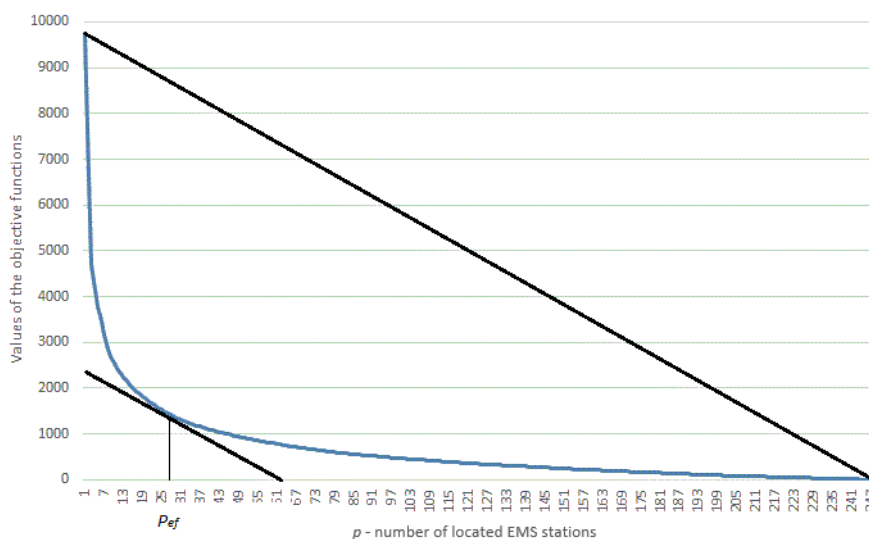
Figure 1: Dependence of the objective value OV on p (the number of located stations)

To determine the value of p_{ef}, we compute the average decrease of the objective value on the given range $< p_{min}, p_{max} >$ by the formula (7).

$$\frac{OV(p_{min}) - OV(p_{max})}{p_{max} - p_{min}} \tag{7}$$

Then we solve the problem (1)-(6) repeatedly for all p from the range $< p_{min}, p_{max} >$ until difference of the objective values for two consecutive p is less than or equal to the average decrease.

We try to verify the hypothesis by comparing the estimations of the effective number of centres with the current EMS systems running in the individual self-governing regions of the Slovak Republic. Furthermore, we want to find how the choice of distance matrix affects the estimated value of p_{ef}.

Usual way of distance matrix determination [3] is based on processing a road map of the considered region, when set of nodes and lengths of edges are obtained and these data are used to compute the distances among the individual nodes. The distances among the objects inside a node are considered to be zero. The number of inhabitants practically did not show up, when we used zero distances inside the municipalities and a service centre is located at this municipality [4]. This evoked us an idea to insert some non-zero distance value in diagonal entries. Therefore, we solve the task of the estimation of the effective number of centers in the given region by using two types of distance matrices in this paper.

3 EXPERIMENTS

We use the road networks of the regions in the Slovak Republic as benchmarks. In the first set of experiments, we use the distance matrix with zero values of diagonal entries corresponding to a distance between objects from the same municipality. In the second set of experiments, the distance between two objects from the same municipality is changed in this way: zero distance for a municipality i is replaced by value of 1 kilometre in case, when an hospital is in the municipality. The distance is increased to the value of 2 up to 3 kilometres in several municipalities with the biggest number of inhabitants or with the large area. The adjustment of the diagonal entries was applied to 88 municipalities from 2916 nodes (total number of municipalities of the Slovak Republic).

As concerns values of p_{min} and p_{max}, these boundary values have a significant effect on the calculation of the average decrease of the objective values for the selected range. Our effort is always to find such parameters of the task, so that the tested hypothesis the best approximates the current situation. Therefore, we decided to test the estimation of p_{ef} for both types of above-mentioned matrices and for two types of ranges. We get four groups of experiments in this way.

The value of p will take all possible values from the basic range, i.e. $p_{min} = 1$ and $p_{max} = |M|$, where M is the number of the municipalities in the given region. In the case of the second type of the range, we try to exclude the unlikely values of p. According to the documents of the institutions that provide the EMS, one EMS centre should service about 30000 inhabitants. Therefore, we make comparative calculations on the reduced range. We set the boundary values of p_{min}^* and p_{max}^* according to the number of inhabitants of the given region. We determine p_{min}^* so that one centre belongs to 40 000 inhabitants and p_{max}^* so that one centre belongs to 20 000 inhabitants.

We solve the weighted p-median problem in cycle with increasing value of p from the given range for each region of the Slovak Republic and for each group of experiments. We use system XPRESS [6] to solve exactly the weighted p-median problem. The system XPRESS solves the linear programming problem in the Mosel environment. Depending on the number of the municipalities, the cycle runs several minutes (the smallest BA region) up to several hours (PO region). In each cycle of experiments, we find such p_{ef} (one or more), in which the actual decrease of the objective value best approximates the average decrease of the objective value. We calculate the average and the actual decreases of the objective value from the optimal solutions by using MS Excel and then we determine the boundary values for estimation of p_{ef}.

The results, obtained for the distance matrix with zeroes on the diagonal (distances among the objects in the same municipality are equal to the value of zero) and for the basic range $<1, M>$, are shown in the table 1.

Table 1: Comparison of the estimated p_{ef} and current p for the distance matrix with zeroes at the diagonal and for the basic range $<1, M>$

Region	p_{max}	OV ($p = 1$)	Average decrease	Boundaries of p_{ef}	Current number of centers
BA	87	77465	900.76	17-18	15
BB	515	321002	624.52	30-34	36
KE	460	308257	671.58	33-35	32
NR	350	245843	704.42	28-30	27
PO	664	423786	639.20	32-34	32
TN	276	209135	760.49	19-22	22
TT	249	189580	764.44	21-22	17
ZA	315	297830	948.50	21-22	29

We do not specify the value of the objective function (*OV*) for $p = M$ in the table 1, because the objective value is equal to the value of zero for the distance matrix with zero values at the diagonal and for the location of the EMS centre into each municipality. The results for the distance matrix that takes into account the transits across large municipalities and for the basic range are shown in the table 2.

Table 2: Comparison of the estimated p_{ef} and current p for the distance matrix taking into account transits in large municipalities and for the basic range <1, *M*>

Region	p_{max}	OV (p = 1)	OV (p = p_{max})	Average decrease	Boundaries of p_{ef}	Current number of centers
BA	87	77853	8308	808.66	16-17	15
BB	515	321153	9163	606.98	30-34	36
KE	460	308460	8953	652.52	31-34	32
NR	350	245886	9943	676.05	27-30	27
PO	664	425618	9161	628.14	32-35	32
TN	276	209176	8027	731.45	18-20	22
TT	249	190904	6960	741.71	21-22	17
ZA	315	297906	9426	918.73	21-22	29

In the last but one column of the tables 1 and 2 (**Boundaries of p_{ef}**), there is listed the boundary values for estimation of p_{ef}. In the last column in both tables (**Current number of centers**), there is introduced the number of the EMS centers, which are currently active.

We could mention similar tables also for the results, obtained for the reduced range < p_{min}^*, p_{max}^* >. The process of the calculation and of the comparison was the same as the one for the basic range <1, *M*>. For better clarity, only the important data are shown in table 3. These include numbers of inhabitants, calculated values of p_{min}^* and p_{max}^* for each region, as well as estimations of the boundary values for estimation of p_{ef}, both for the zero and non-zero diagonal matrices subject to the reduced range< p_{min}^*, p_{max}^* >.

Table 3: Comparison of the results for estimated p_{ef} for tasks with and without zeros at the diagonal of the distance matrix

Region	Number of inhabitants	p_{min}^*	p_{max}^*	Boundaries of p_{ef} for diagonal=0	Boundaries of p_{ef} diagonal≠0	Current number of centers (p_{real})
BA	6061	16	32	22-23	22-23	15
BB	6609	17	36	25-26	25-26	36
KE	7929	15	40	25	25-26	32
NR	6900	17	34	24-25	24-25	27
PO	8183	20	42	27	31-32	32
TN	5942	14	30	19	22-23	22
TT	5563	14	28	20-21	20-21	17
ZA	6911	17	35	24-25	24-25	29

Comparing the results obtained for different distance matrices, we can state that although the changes in matrices are small, they have practical significance. Taking into account the transits of the EMS vehicles among the objects of one municipality, i.e. the case of distance matrix with non-zero diagonal, the increase of the objective value $OV(p)$ can be observed (compare columns "**OV (p = 1)**" of tables 1 and 2). The increase of the objective value for p=1 was approximately from 0 % up to 0.7 % comparing the values, which were obtained for the matrices with zeroes at the diagonal. In the case of distance matrix with non-zero diagonal, the objective function values decreased slower for with increasing value of p. In

this case, the difference of objective values ($OV(p{=}1)$ - $OV(p{=}M)$)) decreased. Therefore, the average decrease of the objective value on the basic range $<1, M>$ was also reduced in comparison to the results of tasks with zeros at the diagonal.

The boundary values for estimation of p_{ef} subject to the basic range $<1, M>$ did not changed in four regions (BB, PO, TT, ZA), when the transits in the municipality were taking into account. The boundary values for p_{ef} were slightly reduced in other four regions. The estimation of p_{ef} got closer to the current number of the EMS centers in three cases and the estimation slightly deviated in one case (TN).

In the case of distance matrix with non-zero diagonal and reduced range $p \in < p_{min}^*, p_{max}^* >$, the boundary values for estimation of p_{ef} changed in three regions (BA, TN, KE). The estimation moved closer to the current number of centres in all these cases.

4 CONCLUSION

We had to answer the question, whether the tested hypothesis is suitable to determine the effective number of the located centers in case of p-median task and whether it is appropriate to take the distances among the objects in one municipality into account in similar tasks. We can answer to the second part of the question positively. The estimation of the value p_{ef} for the modified matrix was successful in 15 instances from 16 benchmarks.

The answer to the first question is more complicated. The estimations of the ranges for p_{ef} differ significantly by changing the values of p_{min} and p_{max} regardless of the type of the used matrix. We can state that an acceptable estimation of the value p_{ef} was achieved for the basic range only.

The use of the modified distance matrix, which takes into account the transit across big municipalities, turned out to be suitable for the further studies in this research area. The suitability of the using the hypothesis to estimate the effective number of the centers for the basic range should still be verify on other benchmarks.

References
[1] Janáček, J. & al. 2010. Design of the areas of large service systems. Žilina: EDIS.
[2] Janáček, J., Kvet, M. 2012. Relevant Network Distances for Approximate Approach to Large p-Median Problems. Operations research proceedings 2012: selected papers of the international conference on operations research: September 4-7, 2012. Hannover, Germany: pp 123-128.
[3] Janáčková, M., Szendreyová, A. 2016. An importance of the population density for the location of the Emergency Medical Service stations. In: Mathematical methods in economics: 34th international conference. Liberec, Czech Republic. pp 354-358.
[4] Janáčková, M., Szendreyová, A. 2015. Time-distance versus utility in the public service system design. In: SOR ´15 Proceedings of the 13th International Symposium on Operational Research. Bled. Slovenia. pp 446-451.
[5] Marianov, V., Serra, D. 2002. Location problems in the public sector. Z. Drezner ed., Springer. Facility location - Applications and theory. pp 119-150. Berlin.
[6] Xpress-MP Manual 'Getting Started'. 2005. Dash Associates, Blisworth.

Operational research model for crew scheduling and application

Morapitiye Sunil

Coauthor: Illés Tibor

musz.sunil@gmail.com, illes@math.bme.hu

Today resource scheduling problems are in the spotlight in the field of operations research, as these problems have several applications in everyday life as well. These problems can be considered to be assignment problems, such as the assignment of pilots and stewardess to designated flights, shop assistants to shifts and public transport (vehicles) to routes, which depends on specific restrictions such as days off and holidays. These restrictions are greatly dependent on the application, blemishing the flow network behaviour of the problem, hence the characterisation of these problems can only be achieved by integer programming, usually leading to an NP-hard problem. Therefore the solution of the problem is a real professional challenge.

My thesis is focused on optimising the public transport network of Budapest, the work was carried out in collaboration with T-Systems Hungary Ltd. Although the actual transport routes and timetables are rarely modified, it is essential to constantly develop improved schedules, which is still surprisingly done in Excel. On the one hand the problem associated with this practise is that it requires a huge amount of time and resources. On the other hand the following questions arise: How good (or optimal) is the solution obtained? How it is possible to prove the optimality of the solution?

The initial step was to turn the problem into a mathematical one, which can be categorized to be an NP-hard problem. Since there is no efficient algorithm for such problems, the main aim was to develop a model (including the use of heuristics), which can be used practically, giving an acceptable result for even larger systems within a reasonable timeframe.

Over the course of this lecture, one mathematical model will be thoroughly introduced, which outputs the number of drivers required and describes the shifts of the drivers as a function of the data provided by T-Systems Hungary Ltd, such as timetables and terms and conditions of the shifts. Due to the use of heuristics and limitation of running time the solution is not necessarily optimal. The model of the assignment of drivers to shifts was solved using the FICO-XPRESS programme much faster than experts. The advantages of our solution to others are the following:

- Develops several solutions similar in quality (close to optimal)

- Faster than experts' solution

- The task of experts shifts from solving the problem to analyse these similar solutions and to select the most optimal one

- Allows experts to work on formulating new criteria and goals

- The quality comparison between schedule variants on the basis of one more criteria (the assignment of drivers) emerges

EVALUATION OF THE INFLUENCE OF DIFFERENT PARAMETERS IN GNSS STATIC POSITIONING

Polona Pavlovčič Prešeren
University of Ljubljana, Slovenia
polona.pavlovcic@fgg.uni-lj.si

Abstract: In GNSS-processing, it is essential to use processing parameters properly. The paper aim is to analyze the advantages and disadvantages of using the cut-off angle and the satellite constellation, occupation duration and the sort of ephemeris in observation processing correctly. Some results proceeded from the same set of the carrier-phase observations, performed in nearly ideal conditions. Simulations of obstacles near GNSS receivers followed the increasing the cut-off angle at specific azimuths. Further goals included the effect of broadcast versus final precise ephemeris on the processed coordinates. Different quality of coordinates showed which parameters influence the quality of the coordinates drastically. The findings in the paper are valuable in developing solutions to support decisions of using the best parameters in an efficient GNSS-positioning.

Keywords: GNSS positioning, post-processing, cut-off angle, occupation duration, carrier-phase observations, ephemeris, quality of coordinates

1 INTRODUCTION

The complete cycle of GNSS coordinates determination includes acquisition of observations at the field as well as their processing either in real time or in post-processing. The final step is to acquire coordinates in the appropriate coordinate system [1]. To achieve the best possible positioning results from GNSS it is essential to perform observations in decent observation conditions by using good processing engine. While positioning, we should avoid sites obscured by nearby buildings and vegetation. Obstructions disable the reception of satellite signals, so the positioning is worsening and in some cases even blocked.

The first challenge here is to show different observation conditions influence the quality of positioning drastically. Besides, a special challenge in high-quality positioning is to use carrier-phase observations and their ambiguity assessment as integers [2]. In some situations, it can happen, that despite good observation conditions, ambiguities cannot be resolved. This can happen in an unpredictable space weather conditions ([3], [4]) or due to the bias resulting from bad synchronization of the transmitted (from the satellite) and generated (in the receiver) signal ([2], [5]).

In GNSS positioning, it is important to consider several factors according to the baseline length and the height difference between the baseline's ends ([6]). There is a variety of credentials to set parameters in GNSS post-processing properly. Software set them as default values by filters. We exclude epochs or observations failing the following quality checks to acquire the best promising results [7]. There are two ways to do this, either by receiver settings at the field or by filters in the processing software. For high-quality positioning, we should realize the best potential conditions or use an alternative way to set coordinates indirectly from other measurement techniques.

However, when using the default values, the awareness that worse measurement conditions at the field could happen only for a shorter period is significant. Computed positions depend on the occupation duration, where redundancy plays a prominent role in meeting the best accuracy. In case of longer baselines, obtaining high accuracy usually requires long observation sessions ([9], [10]). A number of authors studied GNSS static positioning in difficult conditions ([11], [12]). While using many types of instruments at the

baselines ends we should consider also the length of the processing baseline and vertical difference between the ends as well as by using broadcast or precise ephemerides [13].

In GNSS, we perform observations in real conditions, which are in most cases unpredictable. Therefore, there is a variety of topics to discuss subsequently after the position acquisition. The variety of the empirical tests gives us the insight into the problem under predicted conditions. Interesting results came from the investigation of the real-time kinematic positioning under different measurement conditions [14].

The focus throughout this study is to analyze the quality of coordinates by using several frequency's carrier-phase receivers. We performed GNSS measurements in the ideal conditions and further processed under simulated worsen conditions. The last goal is to stress which factors influence the positioning accuracy drastically. Investigations were intended to evaluate the influence of using different parameters in post-processing mode on the quality of the coordinates. In our study, we focused on the short baselines with a minimum observation duration of 30 minutes.

2 PROBLEM DESCRIPTION

It is a common fact, that we access better accuracy in GNSS positioning through redundancy. Observation duration and final accuracy of coordinates are mainly a function of a baseline length. There are several recommendations that baseline lengths should be short, for example, 5 km [15]. Longer baselines need more observations. However, some epoch we should exclude, especially if ([16], [17]):

- there is any observation from a satellite vehicle (SV) closer than 10° to the local horizon; we know the credential as setting the cut-off angle. In some cases, it is set also to a lower value, most commonly value used is 15°;
- any occupation duration with less than five GNSS satellites in common for stations where the baseline is computed;
- any epoch with a PDOP (Position Dilution of Precision) greater than 6 (sometimes the values is set to 7), and for vertical positions, any epoch with a VDOP (Vertical Dilution of Precision) greater than 6 (in some cases 7).

In some rare cases, the increasing the cut-off angle to about 20° can be even advantageous. Such situations come from the disturbed ionosphere. In other cases, we should be aware that the increasing the cut-off angle means the lower number of visible satellites in view and further higher values of *DOP* factors. They depend on the geometry of four optimal satellites in view and the receiver, which form the unit tetrahedron. Larger volume of the tetrahedron means smaller measurement conditions and worsened quality of coordinates.

For successful GNSS point determination, it is good to have the best possible conditions at the vicinity of a point. By knowing point's approximate coordinates, we can even simulate conditions in a number of visible satellites and *DOP* factors (Figure 1).

By changing the cut-off angles at different azimuths (left picture in Figure 1, where trajectories of visible satellites are presented), we simulate physical obstructions to reception of the satellite signal near point's location. Right pictures of the Figure 1 depict the number of available satellites (upper) and *DOP* factors for a certain period in the future (lower). We present satellites from different systems separately, namely GPS, GLONASS, GALILEO, BEIDOU and QZSS. When using the receivers, which enable us the reception of only one system (mostly GPS), the number of visible satellites is even smaller. Presently, GPS and GLONASS are most widely used systems for positioning, especially due to the GALILEO problems with satellites' clocks [19].

Such investigations and simulations are precious when trying to perform positioning in bad conditions. According to the prior estimated number of satellites and *DOP* factors we

can estimate, which observation time will be the most effective. We should be aware that any time of observation performance needs at least four visible satellites in view and good *DOP* factors, i.e. values below 6.
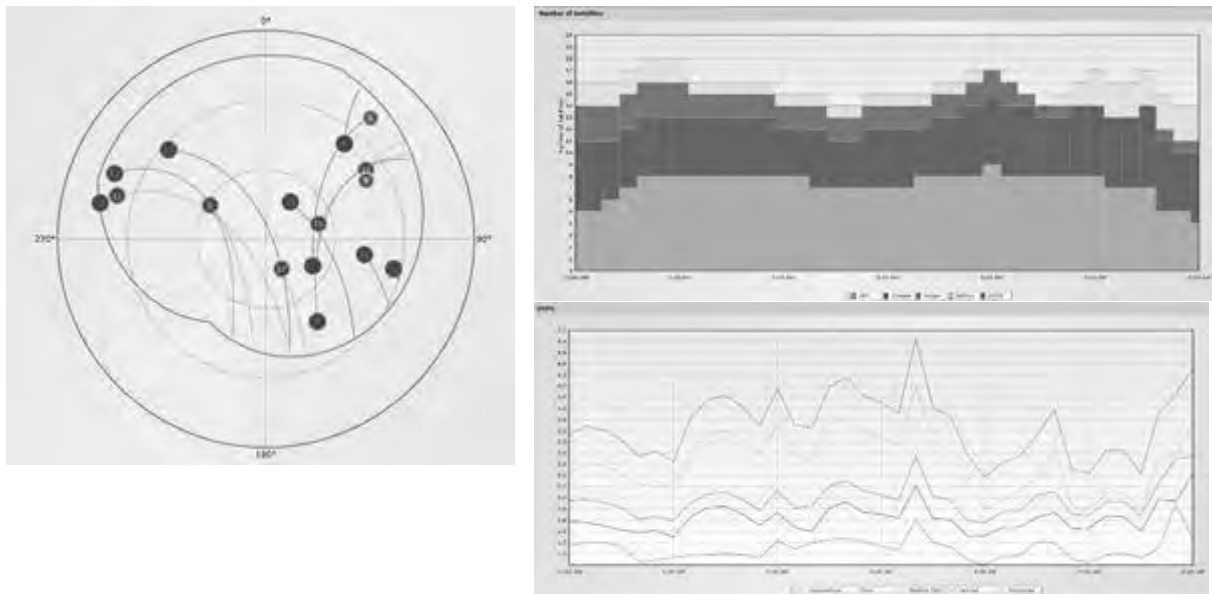
Figure 1: Prior prediction of GNSS surveying conditions by using Trimble online planning application [18].

In some cases, the possibility of GNSS positioning in worse conditions exists. We should be aware of a bad quality of coordinates in such conditions. Observation acquisition in good conditions and further simulation of worsen conditions give as the insight into the quality of estimated coordinates.

3 METHODOLOGY

Evaluation of the influence of using different parameters in GNSS processing followed a relative carrier-phase static positioning with a short baseline's length, approximately 4 km. All coordinates acquired initiated from the same set of two RINEX (Receiver Independent Exchange format) files. Further post-processing followed changing the parameters. Solutions distinguished according to different:
minimal cut-off elevation angle,
- application of ephemeris (broadcast or precise),
- shortening of the observation duration and
- interval of signal registration (1, 5 and 15 s).
In the further analysis, we used coordinates from optimal conditions as the reference. The analysis followed the comparison of northing (n) and easting (e) coordinates in the Slovenian realization of the coordinate system ETRS (D96/TM) as well as the comparison of ellipsoid heights (h). However, we presented results from the height component separately. The reason for that comes from the common fact, that in GNSS positioning, the accuracy of the height component is twice worsened comparing the horizontal coordinates.

4 RESULTS

The goal of experiments was to access and compare results in post-processing of relative static carrier-phase positioning based on the identical baseline of a short length.

Deviations in coordinates of the same set of RINEX files' processing, but by using different cut-off angles.

Cut-off angle	Δe [m]	Δn [m]	Δh [m]
0°	0.000	0.000	0.001
5°	0.000	0.000	0.001
10°	0.000	0.000	0.000
20°	−0.001	−0.001	0.000
25°	−0.003	−0.001	0.003
30°	−0.006	−0.005	0.019
35°	−0.004	−0.010	0.023
40°	−0.012	−0.007	0.063
45°	−0.011	−0.012	0.082

The results from Table 1 show that by using the cut-off elevation angle until 20° the acquired coordinates are mostly the same. By increasing the cut-off angle from 20° to 45° deviations in coordinates increase significantly. The increasing the cut-off angle means observation processing with a lower number of common satellites at the baseline ends. It also means the high barrier in the surrounding of the point restricts the number of satellites and therefore, contributes to a decrease of the geometric constellation of satellites available and further affects the position quality. Deviations in coordinates can reach from some millimeters to centimeter level at cut-of angle 45°. From Figure 3, we can see that the increasing cut-off angle can affect the heights even drastically. Deviations from some millimeters at the cut-off angle 25° to several centimeters at the cut-off angle 45° exist.
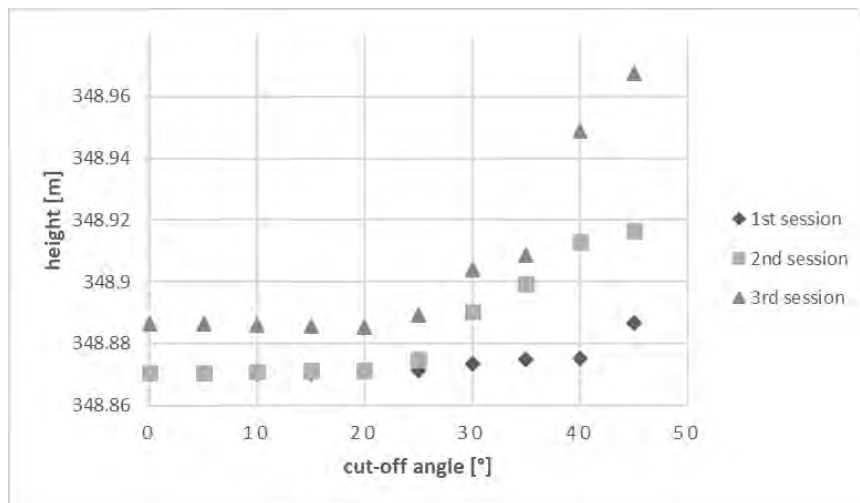
Figure 2: Deviations in heights by using different cut-off angles and consequently lower number of common satellites.

While using broadcast or precise ephemeris we got the same coordinates. This comes from the theoretical fact, that in relative positioning, many impacts cancel due to the double difference. The problem is more significant for longer baselines, which we did not take into the consideration in the present study.

In order to reach better results regard to RTK (Real Time Kinematic) measurements, we can use also fast static method. The main difference from static method is in shorter observation duration at the point. It is a common fact that GNSS redundancy leads to better results. Therefore, we investigated the question of meeting the same quality of coordinates

whether we perform shorter measurements with minimal interval registration. There are results from several processing steps that differed in occupation duration as well as the minimal time of signal registration in Table 2. We compared results, acquired from three points, named A, B and C. Static processing results from one-hour observation length with the 15-second interval of signal observation were the reference. Further, we used fast-static (15-minute) observations with 1 s, 5 s and 15 s intervals of registration.

Table 2: Deviations in coordinates of the same set of RINEX files' processing, but by using different duration occupation and intervals of signal registration

Point/deviations	Δe [m]	Δn [m]	Δh[m]
15-minute fast static observations, 15 s interval of signal registration			
A	−0.002	−0.006	0.002
B	0.002	−0.008	0.007
C	−0.002	−0.005	0.003
15-minute fast static observations, 5 s interval of signal registration			
A	0.000	0.003	0.006
B	0.003	0.006	0.007
C	0.002	0.005	0.003
15-minute fast static observations, 15 s interval of signal registration			
A	0.003	0.003	0.002
B	0.004	0.006	0.000
C	0.001	0.001	0.001

Results from Table 2 indicate there is no significant difference, whether we use shorter interval of signal registration in fast static measurements. We can conclude longer observation duration at the point has its advantages because different satellite constellation leads to better quality of coordinates.

5 DISCUSSION AND CONCLUSIONS

In the paper, we have presented results from different processing strategies for a short GNSS baseline, with the length of 4 km. Results proceeded from the same set of the carrier-phase relative measurements, performed in nearly ideal GNSS conditions. It means there were no obstacles at the receiver vicinity that could enable the GNSS signal reception. The results acquired are important because they address different aspects of processing strategies, which we usually do not take into the consideration. We can conclude that for short baselines, the usage of precise ephemeris is not beneficial. There is an important finding according to the site occupation duration and signal registration time. Simulations showed that occupation duration is more effective than setting the signal registration to a minimum. The finding obviously comes from the fact that different satellite constellation leads to the best possible results and because of that, static GNSS method with longer occupation duration still has its preferences. This is also the reason that several site occupations when performing kinematic observations is of a great importance.

References

[1] Sterle, O., Pavlovčič Prešeren, P., Kuhar, M. and Stopar B. 2009. Definicija, realizacija in vzdrževanje modernih koordinatnih sistemov. Geodetski vestnik, 56 (4): 679–694.

[2] Geng, J., Meng, X., Dodson, A. H. and Teferle, F. N. 2010. Integer ambiguity resolution in precise point positioning: method comparison. Journal of Geodesy, 84(9), 569–584.

[3] Sterle, O., Pavlovčič Prešeren, P. and Stopar B. (2013). Modeliranje ionosferske refrakcije za izboljšavo absolutnega GNSS-položaja s kodnimi instrumenti: priprava na 24. Sončev cikel. Geodetski vestnik, 57(2), 9–24.

[4] Jakowski, N., Mayer, C., Wilken, V., Hoque, M. M. 2008. Ionospheric Impact on GNSS Signals. Fisica de la Tierra. Special Edition: Ionosphere and its Influence on Positioning and Satellite Navigation, Vol. 20.

[5] Gao, Y. 2006. Precise Point Positioning and Its Challenges, Aided-GNSS and Signal Tracking. Inside GNSS, 1(8), 16–18

[6] Okorocha, C. V. and Olajugba, O. 2014. Comparative Analysis of Short, Medium and Long Baseline Processing in the Precision of GNSS Positioning. FIG Congress 2014 Engaging the Challenges – Enhancing the Relevance, Kuala Lumpur, Malaysia 16–21 June 2014.

[7] Dach, R., Lutz, S., Walser, P. and Fridez, P. 2015. User manual the Bernese GNSS Software, Version 5.2. Bern, Astronomical Institute University of Bern.

[8] Takasu, T. and Yasuda, A. 2009. Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. International Convention Center Jeju, Korea, November 4-6, 2009.

[9] Dawidowicz, K. 2012. GNSS satellite leveling using the ASG-EUPOS system services. Technical Sciences, 15(1), 35–48.

[10] El-Mowafy, A. 2011. Analysis of web-based GNSS post-processing services for static and kinematic positioning using short data spans. Survey Review, 43(323), 535–549.

[11] Bakuła, M. 2013. Study of Reliable Rapid and Ultra Rapid Static GNSS Surveying for Determination of the Coordinates of Control Points in Obstructed Conditions. Journal of Surveying Engineering, 139(4), 188–193.

[12] Dawidowicz, K. and Krzan, G. 2014. Coordinate estimation accuracy of static precise point positioning using on-line PPP service, a case study, Acta Geodaetica et Geophysica, 49(1), 37–55.

[13] Montenbruck, O., Steigenberger, P. and Hauschild, A. 2015. GPS Solutions, 19(2): 321–333.

[14] Pirti, A., Yucel, M. A. and Gumus K. 2013. Testing real time kinematic GNSS (GPS and GPS/GLONASS) methods in obstructed and unobstructed sites. Geodetski vestnik, 57(3): 498–512.

[15] Leica Geosystems. 2000. Geosystems General Guide to Static and Rapid-Static.

[16] The Connecticut Association of Land Surveyors, Inc. 2008. Guidelines and Specifications for Global Navigation Satellite System Land Surveys in Connecticut.

[17] Geodetska uprava RS. 2007. Tehnično navodilo za uporabo novega koordinatnega sistema v zemljiškem katastru.

[18] Trimble online planning. 2017. http://www.trimble.com/GNSSPlanningOnline [Accessed 26/03/2017].

[19] Inside GNSS. 2017. Gibbons Media & Research LLC. http://www.insidegnss.com [Accessed 26/03/2017].

Dynamic algorithm for
distance constrained labelings of graphs

Aleksander Vesel

University of Maribor, Faculty of Natural Sciences and Mathematics, Slovenia

Abstract: The Frequency Assignment Problem asks for assigning frequencies to transmitters in a broadcasting network with the aim of avoiding undesired interference. FAP includes a variety of graph theoretical specific subproblems, many of them are variations of the classical graph coloring which involve graph distance, i.e. a condition is usually imposed on the vertices that are given the same color. Some well known examples are the $L(2,1)$-labeling of graphs and the packing coloring of graphs. A dynamic algorithm approach for computing these two graph invariants is described in this paper.

Keywords: frequency assignment, distance constrained labeling, dynamic algorithm

1   Introduction

Efficient assignment of frequencies for wireless communication in different network scenarios has become an extremely important topic of recent research. The main reason is the enormous growth of wireless networks which causes the issue of efficiently using the radio spectrum - an expensive and scarce resource - very important. The wide class of frequency assignment problems (FAPs) asks for assigning frequencies (channels) from the available radio spectrum to the transceivers of the network, so that unconstrained simultaneous transmissions cannot cause interference. The objective is to minimize the used radio spectrum. While problems of these type arose some decades ago in radio and TV broadcasting, today they play a growingly significant role due to the wide diffusion of telephone and satellite communication.

In a broadcasting network, each transceiver is assigned a frequency channel for its transmissions. Two transmissions can interfere if their channels are too close. That means that even if two transceivers use different channels, there still may be interference if the two transceivers are located close to each other.

For the interference graph of the network such kind of problem can be seen as a variation of a vertex coloring problem. More formally, for a given graph $G = (V, E)$ modeling the wireless network, where vertices represent transceivers and edges between two transceivers represent possible interference, we consider a function from the set of vertices of G to the set of nonnegative integers such that the given separation constraints are satisfied.

In particular, Goddard at al. [4] introduced the concept of broadcast chromatic number. Later this model became known as the packing chromatic number.

A *k-coloring* of a graph G is a function f from $V(G)$ onto a set $C = \{1, 2, \ldots, k\}$ (with no additional constraints). The elements of C are called *colors*. Let X_i denote the set of vertices with the image (color) i. Note that $X_1, ..., X_k$ is partition of the vertex set of G into disjoint (color) classes.

Let $X_1, ..., X_k$ be a partition of the vertex set of G with respect to the following constraints: each color class X_i is a set of vertices with the property that any distinct pair $u, v \in X_i$ satisfies $d_G(u, v) > i$. Here $d_G(u, v)$ denotes the usual shortest path distance between u and v. Then X_i is said to be an *i-packing*, while such a partition is called a *packing k-coloring*. The smallest integer k for which there exists a packing k-coloring of G is called the *packing chromatic number of G* and it is denoted by $\chi_\rho(G)$.

A more general concept which can be noted already in [4] is later formally introduced in [5] as follows. For a nondecreasing sequence of integers $S = (s_1, s_2, \ldots)$, an *S-packing k-coloring* is a k-coloring c of $V(G)$ such that for every i, with $1 \le i \le k$, c is an s_i-packing. The *S-packing*

chromatic number of G denoted by $\chi_\rho^S(G)$, is the smallest k such that G admits an S-packing k-coloring.

Note that by setting $S := (1, 1, 1, \ldots)$ the S-packing k-coloring is the classical graph coloring with k colors, while for $S := (1, 2, 3, \ldots)$ we obtain a packing k-coloring.

Another important graph theoretical model is the concept of *distance constrained labeling* of graphs defined as follows. Let $k_1, k_2, \cdots, k_n$ be positive integers with $k_1 \geq k_2 \geq \cdots \geq k_n$. An $L(k_1, k_2, \cdots, k_n)$-labeling (we call it a graph labeling) for G is a function $f : V(G) \to \{0, 1, \cdots, \lambda\}$ such that for every $u, v \in V(G)$ and for all $1 \leq i \leq n$, $|f(u) - f(v)| \geq k_i$ if $d(u, v) = i$. The $\lambda_{k_1, k_2, \cdots, k_n}$-number for G, denoted by $\lambda_{k_1, k_2, \cdots, k_n}(G)$, is the minimum value λ such that G admits an $L(k_1, k_2, \cdots, k_n)$-labeling.

Among the variants of the frequency assignment problem, the $L(2, 1)$-labeling problem where adjacent vertices must be assigned colors of distance at least two apart and vertices of distance two must be assigned different colors has attracted most interest. More formally, a $L(2, 1)$-labeling of a graph G is a function f from the vertex set $V(G)$ to the set of non-negative integers C (called *labels*) such that

$$|f(u) - f(v)| \geq 2 \text{ if } d(u, v) = 1, \text{ and}$$
$$|f(u) - f(v)| \geq 1 \text{ if } d(u, v) = 2.$$

An k-$L(2,1)$-*labeling* is a $L(2, 1)$-labeling of G such that $C = \{0, \ldots, k\}$. An optimal $L(2, 1)$-labeling of G is a k-$L(2,1)$-labeling with k smallest possible. The largest label used by an optimal $L(2, 1)$-labeling is called the λ-*number* of G and denoted by $\lambda(G)$.

It is known that the packing k-coloring as well as the distance constrained labeling problem are **NP**-hard for general graphs [4, 6], and even for many special graphs such as planar graphs, bipartite graphs and chordal graphs [2, 3].

One way to approach NP-hard problems is to decompose a graph into smaller and simpler constituents by some basic operations and try to gain a global information from the constituents. For instance, one might try to decompose a graph with respect to one of the standard graphs products. Quite often, product graphs having cycles or paths as factors are important. A natural generalization of such structures is formed by rotagraphs that are in turn an important class of polygraphs.

In this paper we present a dynamic algorithm approach for computing distance constrained colorings of graphs that can be viewed as rotagraphs. The main idea is to build a function (an invariant) on a graph product from corresponding functions on its basic building blocks. Then the problem reduces to a search for a certain subgraph in an associated directed graph. In particular, if the function in question is rather local, one has to look for directed cycles.

2   Dynamic algorithm

Let $G = (V(G), E(G))$ be a graph. A *walk* is a sequence of vertices $v_1, v_2, \ldots, v_n$ and edges $v_i v_{i+1}$, $1 \leq i \leq n - 1$. A *path* on n vertices is a walk on n different vertices and denoted P_n. A walk is *closed* if $v_1 = v_n$. A closed walk in which all vertices (except the first and the last) are different, is a *cycle*. The cycle on n vertices is denoted C_n. For $u, v \in V(G)$, $d_G(u, v)$ or $d(u, v)$ denotes the length of a shortest walk (i.e., the number of edges on a shortest walk) in G from u to v. These definitions extend naturally to directed graphs.

Let $G_0, G_1, \ldots, G_{n-1}$ be disjoint graphs and $X_0, X_1, \ldots, X_{n-1}$ a sequence of sets of edges such that each edge of X_i joins a vertex of G_i with a vertex of G_{i+1} (indices modulo n). A *polygraph*

$$\Omega_n = \Omega_n(G_0, G_1, \ldots, G_{n-1}; X_0, X_1, \ldots, X_{n-1})$$

is defined in the following way:

$$V(\Omega_n) = V(G_0) \cup V(G_1) \cup \cdots \cup V(G_{n-1}),$$

$$E(\Omega_n) = E(G_0) \cup X_0 \cup E(G_1) \cup X_1 \cup \cdots \cup E(G_{n-1}) \cup X_{n-1}.$$

Polygraphs were introduced in chemical graph theory as a model for polymers, cf. [1].

Assume that for $0 \le i \le n-1$, G_i is isomorphic to a fixed graph G. Let, in addition, the sets X_i, $0 \le i \le n-1$, be equal to a fixed edge set X. Then we call the polygraph Ω_n a *rotagraph* and denote it $\omega_n(G; X)$. We will also say that $\omega_n(G; X)$ is a rotagraph with consecutive *fibers* $G_0, G_1, \ldots, G_{n-1}$. A *fasciagraph* $\Psi_n(G; X)$ is a rotagraph $\omega_n(G; X)$ *without* edges between the fibers G_{n-1} and G_0.

In the rest of this section we recall concepts and results that were introduced in [7] and are essential for the present work. For a graph G set

$$\mathcal{F}_q(G) = \{f : V(G) \to \{0, 1, \ldots, q-1\}\}.$$

A subset of $\mathcal{F}_q(G)$ will be called a *graph q-property*. If q will be clear from the context or not essential, we will shortly say a *graph property*.

Let $\mathcal{L}_q(G) \subseteq \mathcal{F}_q(G)$ be the set of functions f with the following property: Let $f \in \mathcal{L}_q(G)$, then if $uv \in E(G)$ we have $|f(u) - f(v)| \ge 2$, and if $d(u, v) = 2$ we have $|f(u) - f(v)| \ge 1$. Clearly, $\mathcal{L}_q(G)$ describes the admissible $(2, 1)$-labelings of G.

Let $\omega_n(G; X)$ be a rotagraph with consecutive fibers $G_0, G_1, \ldots, G_{n-1}$. Then the restriction of $f \in \mathcal{F}_q(\omega_n(G; X))$ to consecutive fibers $X_i, X_{i+1}, \ldots, X_{i+k}$ (indices modulo n) will be denoted f_i^{i+k}. We say that a graph property $\mathcal{P}_q$ is *hereditary* (for rotagraphs), if for any rotagraph $\omega_n(G; X)$ with consecutive fibers $G_0, G_1, \ldots, G_{n-1}$,

$$f \in \mathcal{P}_q(\omega_n(G; X)) \Rightarrow f_i^{i+k} \in \mathcal{P}_q(\Psi_{k+1}(G; X)); \ i, k = 0, 1, \ldots, n-1.$$

Note that $\mathcal{L}_q$ is hereditary property.

A graph property $\mathcal{P}_q$ is called *d-local* (for rotagraphs), $d \ge 1$, if for any rotagraph $\omega_n(G; X)$, $n \ge 2d+1$, with consecutive fibers $G_0, G_1, \ldots, G_{n-1}$, and any $f \in \mathcal{F}_q(\omega_n(G; X))$,

$$f_i^{i+d} \in \mathcal{P}_q(\Psi_{d+1}(G; X)), 0 \le i \le n-1 \Rightarrow f \in \mathcal{P}_q(\omega_n(G; X)).$$

Note that $\mathcal{L}_q$ is a 2-local property.

Let $\mathcal{P}_q$ be a d-local property, and $\omega_n(G; X)$ a rotagraph with $n \ge 2d+1$. We define a directed graph $D_d(G; X)$ as follows. Its vertices are the functions from $\mathcal{P}_q(\Psi_2(G; X))$, while its arcs are of two types: the first type arcs will be simply called *arcs*, and the second type arcs will be *d-arcs*. Now, in $D_d(G; X)$ make an arc from f to g if and only if f restricted to the second fiber of $\Psi_2(G; X)$ equals to g restricted to the first fiber of $\Psi_2(G; X)$. In addition, if $d \ge 2$, then for any directed path (consisting of arcs) of length $d-1$, say $f_1 \to f_2 \to \ldots \to f_d$, we make a d-arc from f_1 to f_d whenever the composition of $f_1, f_2, \ldots, f_d$ belongs to $\mathcal{P}_q(\Psi_{d+1}(G; X))$. In the particular case when $d = 2$ we interpret this as follows: If the composition of f_1 and f_2 belongs to $\mathcal{P}_q(\Psi_3(G; X))$ then we leave the arc from f_1 to f_2, otherwise we remove it.

Theorem 2.1 ([7]) *Let $\mathcal{P}_q$ be a hereditary, d-local property, and $\omega_n(G; X)$ a rotagraph with $n \ge 2d+1$. Then $\mathcal{P}_q(\omega_n(G; X)) \ne \emptyset$ if and only if $D_d(G; X)$ contains (not necessarily different) vertices $f_0, f_1, \ldots, f_{n-1}$ connected with arcs (f_i, f_{i+1}) and d-arcs (f_i, f_{i+d-1}) for $i = 0, 1, \ldots, n-1$ (indices modulo n).*

Corollary 2.2 ([7]) *Let $\mathcal{P}_q$ be a hereditary, d-local property, $1 \le d \le 2$, and $\omega_n(G; X)$ a rotagraph with $n \ge 5$. Then $\mathcal{P}_q(\omega_n(G; X)) \ne \emptyset$ if and only if $D_d(G; X)$ contains a directed closed walk of length n.*

3 Distance constrained labeling

We need the following definition.

The *strong product* of graphs G and H is the graph $G \boxtimes H$ with vertex set $V(G) \times V(H)$ and $(x_1, x_2)(y_1, y_2) \in E(G \boxtimes H)$ whenever

$x_1 y_1 \in E(G)$ and $x_2 = y_2$, or

$x_2 y_2 \in E(H)$ and $x_1 = y_1$, or

$x_1 y_1 \in E(G)$ and $x_2 y_2 \in E(H)$.

The strong product is commutative and associative, having the trivial graph as a unit.

3.1 $L(2, 1)$-labeling for strong products of paths and cycles

In order to compute the a $L(2, 1)$-labeling of the strong product of two cycles, we define a digraph $D_{n,k}$ as follows. Its vertices are the k-$L(2, 1)$-labelings of $C_n \boxtimes P_2$. Let $u = u_1 u_2$ be a vertex of $D_{n,k}$. Then u_1 and u_2 represent the k-$L(2, 1)$-labeling of $C_n \boxtimes P_2$ restricted to the first and second copies of C_n, respectively.

Let u and v be two vertices of $D_{n,k}$. Then $\widetilde{uv}$ denotes the labeling of $C_n \boxtimes P_3$ obtained by applying u_1, u_2, and v_2 to the consecutive copies of C_n. (Note that $\widetilde{uv}$ is not always a k-$L(2, 1)$-labeling of $C_n \boxtimes P_3$.) We make an arc from u to v in $D_{n,k}$ if and only if, the following two conditions are fulfilled:

(i) u_2 equals v_1 and

(ii) $\widetilde{uv}$ is a k-$L(2, 1)$-labeling of $C_n \boxtimes P_3$.

We now have the following result that follows from the results presented in Section 2.

Theorem 3.1 *[9]* $C_i \boxtimes C_j$ *admits a k-$L(2, 1)$-labeling if and only if $D_{i,k}$ contains a closed directed walk of length ℓ.*

The dynamic algorithm first generates all k-$L(2, 1)$-labelings of $C_i \boxtimes P_2$ which are the vertices of $D_{j,k}$. Since a main building block $C_i \boxtimes P_2$ is usualy relatively small, a simple method, e.g. backtracking, can be applied for this step. In the next step, the set of edges of $D_{i,k}$ has to be generated. A simple procedure for this step is described in [8].

The final step of the approach is the search for closed direct walks in $D_{i,k}$. In order to find these walks one can apply a matrix multiplication of the adjaceny matrix of $D_{i,k}$ or a breadth (depth) first search in $D_{i,k}$.

Note that a similar approach for the Cartesian product of cycles is described in [10].

3.2 Packing chromatic number of the strong product of two cycles

Our goal is to compute the packing chromatic number of the strong product of two cycles. We now define a directed graph as follows.

Let G be a graph. The vertices of D_k^G are all packing k-colorings of $G \boxtimes P_k$. Let u and v be two distinct vertices of D_k^G. Then f_{uv} denotes a k-coloring of $G \boxtimes P_{2k}$ such that u and v compose the respective packing k-coloring of the first and the second copy of $G \boxtimes P_k$ in $G \boxtimes P_{2k}$. Note that f_{uv} need not to be a packing k-coloring of $G \boxtimes P_{2k}$. However, uv is an arc in $D'_{n,k}$ if and only if f_{uv} is a packing k-coloring of $G \boxtimes P_{2k}$.

The following result is the basis for the dynamic computation of the packing chromatic number of $G \boxtimes C_m$.

Lemma 3.2 *Let n, t and k be integers. If G is a graph, $G \boxtimes C_{tk}$ admits a packing k-coloring if and only if D_k^G contains a closed directed walk.*

Proof. Suppose first that D_k^G contains contains a closed directed walk P. Let u, v and w be consecutive vertices of P. It follows that f_{uv} and f_{vw} are both packing k-colorings of $G \boxtimes P_{2k}$.

Let f_{uvw} be a k-coloring of $G \boxtimes P_{3k}$ such that u, v and w compose the respective packing k-colorings of the consecutive copies of $G \boxtimes P_k$ in $G \boxtimes P_{3k}$. Let x (resp. y) be a vertex of a copy of $G \boxtimes P_k$ that corresponds to u (resp. w). Since the distance between x and y in the graph $G \boxtimes P_{3k}$ that corresponds to f_{uvw} is at least $k + 1$, f_{uvw} is a packing k-coloring of $G \boxtimes P_{3k}$.

If $G \boxtimes C_{tk}$ admits a packing k-coloring denoted by f, let H_i, $i \geq 1$, be a copy of $G \boxtimes P_{ik}$ in $G \boxtimes C_{tk}$. By the definition of D'_k, the restriction of f to H_i corresponds to a walk of length i in D'_k. Since i can be chosen arbitrarily and the cardinality of $V(D_k^G)$ is finite, D_k^G cannot be acyclic. This completes the proof. $\square$

References

[1] Babić, D., Graovac, A., Mohar, B., and Pisanski, T. (1986). The matching polynomial of a polygraph. Discrete Appl. Math., 15: 11-24.

[2] Bodlaender, H. L., Kloks, T., Tan, R. B., and van Leeuwen, J. (2004). Approximations for λ-coloring of graphs. The Computer Journal 47: 193-204.

[3] Fiala, J., Golovach, P.A. (2010). Complexity of the packing coloring problem for trees. Discrete Appl. Math., 158 : 771–778.

[4] Goddard, W., Hedetniemi, S. M., Hedetniemi, S. T., Harris, J. M., and Rall, D. F. (2008). Broadcast chromatic numbers of graphs, Ars Combin. 86: 33–49.

[5] Goddard, W., Xu, H. (2012). The S-packing chromatic number of a graph, Discuss. Math. Graph Theory 32: 795–806.

[6] Griggs, J. R., Yeh, R. K. (1992). Labelling graphs with a condition at distance two. SIAM J. Discrete Math. 5: 586–595.

[7] Klavžar, S., Vesel, A. (2003). Computing graph invariants on rotagraphs using dynamic algorithm approach: the case of (2,1)-colorings and independence numbers. Discrete Appl. Math. 129: 449–460.

[8] Korže, D., Vesel, A. (2005). $L(2,1)$-labeling of strong products of cycles. Inf.Process.Lett. 94: 183–190.

[9] Shao, Z., Vesel, A. (2014). $L(2,1)$-Labeling of the strong product of paths and cycles. Sci. World J.

[10] Shao, Z., Vesel, A. (2015). $L(3,2,1)$-labeling of triangular and toroidal grids. Cent. Eur. J. Oper. Res. 23 (3): 659–673.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 5:
Machine Learning

Sample size for assessment of new feature's relevance in a given problem

Marko Bohanec
Salvirt Ltd.
Dunajska cesta 136, 1000 Ljubljana, Slovenia
E-mail: marko.bohanec@salvirt.com

Mirjana Kljajić Borštnar
University of Maribor, Faculty of Organizational Sciences
Kidričeva cesta 55a, 4000 Kranj, Slovenia
E-mail: mirjana.kljajic@fov.uni-mb.si

Marko Robnik-Šikonja
University of Ljubljana, Faculty of Computer and Information Science
Večna pot 113 , 1001 Ljubljana, Slovenia
E-mail: marko.robnik@fri.uni-lj.si

Abstract

In practical use of machine learning models users may add *ad hoc* new features to an existing classification model, reflecting their (changed) empirical understanding of a field. New features potentially increase classification accuracy of the model or improve its interpretability. We introduce a guideline for determination of sample size needed to reliably estimate the impact of a new feature. Our approach is based on the feature evaluation measure ReliefF and bootstrap-based estimation of confidence intervals for feature ranks. The results show that new features with high or low rank can be detected with a relatively small number of instances, but features ranked near the border of useful features need larger samples to determine their impact. We test our approach on qualitative business-to-business sales forecasting data.

Keywords − machine learning, feature ranking, feature evaluation

1   Introduction

In business practice, users of machine learning (ML) models are pragmatic about their effort to collect data describing a business process, for example selling into business-to-business (B2B) market segment. As mentioned in [3], users are upfront interested to learn how many historic cases are needed for the model to identify the most relevant features. For example, in [3] only $\simeq 1/3$ of the final data set would be needed to identify top three features with 80% certainty (if their rank within top 3 is not relevant).

When the data set is collected and the model built, optimized and in use, a new question arises from domain-expert users, adding new features *ad hoc* [5, p. 1159]: how many instances are needed to estimate the impact of a new, candidate feature? Here users try to minimize the effort needed, which in practice means that only a few dozen of instances could be available for an assessment of feature's impact. In this paper, we extend our previous research to answer this question. In our previous work [3], we analyzed the number of features and the number of instances needed to learn important features in a general business setting. Here we focus on reliability of ranks for new features given the context of an existing data set. We use a publicly available B2B sales forecasting data set [1] as a case study.

The rest of the paper is organized as follows. In Section 2 we introduce data set and calculate ground truth. In Section 3 we formalize the problem, and continue with experiments in Section 5. Our conclusions are put forward in Section 5.

2   Data set and ground truth

In this section, we introduce the data set and ground truth for feature ranks. We try to identify the median rank of a particular feature obtained from the random samples of size $|V|$. We use median instead of mean to obtain robust results.

As a use case we use a real world B2B sales data set [1] with 448 instances, 22 features and a class variable with two values. To form an optimization problem we need ground truth ranks of features which, for practical problems, are unavailable. We estimate the ground truth ranks of features $(a_1, ..., a_t)$, t being the number of features, we rank the features with a selected feature ranking algorithm on the complete data set using 10-fold cross-validation. In this paper, we use ReliefF feature evaluation [8], known for its robustness and ability to detect strongly dependent features. Figure 1 shows box-and-whiskers plots for all 22 attributes. The ranks of the most important features are stable, as indicated by low variance around median in box-and-whiskers plots. Similarly, the least performing features are consistently the last.
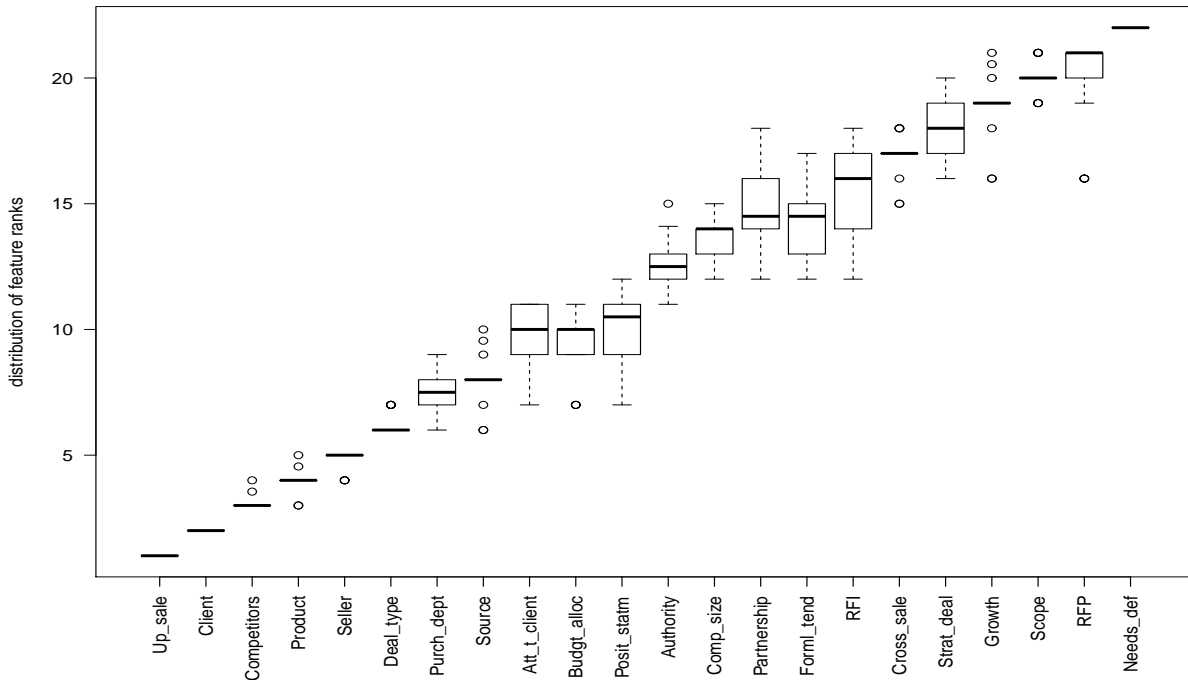
Figure 1: Feature ranks on the complete data set (ground truth), estimated with ReliefF using 10-fold cross-validation. Horizontal axis shows features and vertical axis shows distribution of their ReliefF ranks.

Next, we observe feature ranks as the number of instances increases. We start with a random sample of size 10 and increase the sample size to 150 in increments of 10. Each sample size is resampled 100-times from a complete data set. The distributions of obtained ranks for the feature with the strongest impact from Figure 1 (i.e. "Up_sale") for different sample sizes are reported in Figure 2. We see that this feature is consistently ranked among the best features even with very low number of instances. From sample size 30 this feature is ranked among top 5 features with high probability.

The rank distributions of the least performing feature "Needs_def" (ranked 22nd in Figure 1) are presented in Figure 3. The results show that this feature indicates a clear tendency to bottom
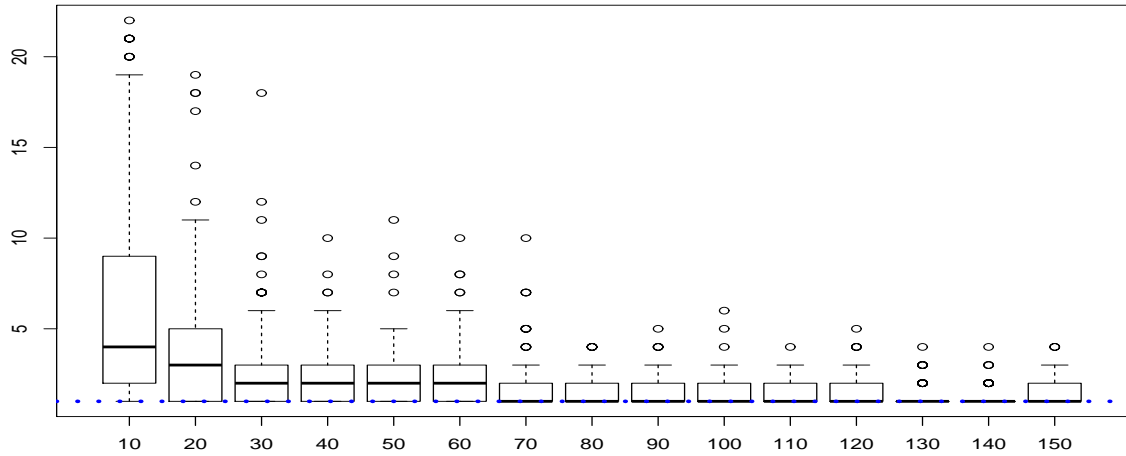
Figure 2: Distribution of ReliefF ranks for feature "Up_sale" ranked 1st for different sizes of estimation set (sampled directly from the full data set). Dotted line indicates true rank.

ranks from the smallest subset size on. From sample size 10, the vast majority of obtained ranks are larger than 10, as indicated by the notch of the box-and-whiskers plot.
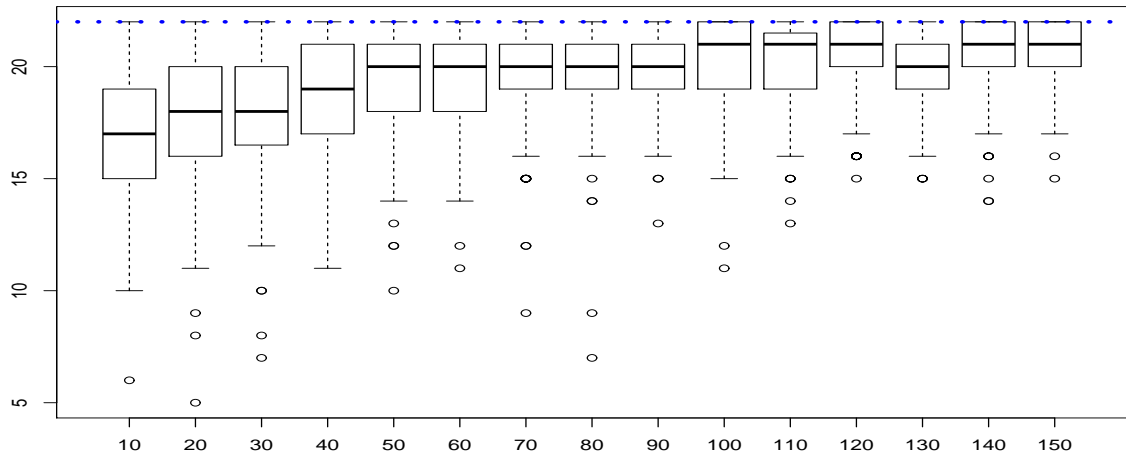
Figure 3: Distribution of ReliefF ranks for feature "Needs_def" ranked 22nd (last) for different sizes of estimation set (sampled directly from the full data set). Dotted line indicates true rank.

3 Formalization of the problem

We assume that we estimate features' impact within an existing data set. We evaluate the number of instances needed for a feature to reliably show its impact given that the ground truth is known. Therefore, our goal is to find the smallest size of a random subset of instances $|V|$, which assures that for a given feature a_i, ranked by function R, the rank of the feature computed on V is close to the the rank obtained on the complete data set:

$$|R(a_i^V) - R(a_i)| \le \varepsilon$$

We use the following notation: R is a ranking function, a_i represents feature i, $R(a_i)$ is the rank of feature a_i on the complete data set of size n, $R(a_i^V)$ is rank of feature a_i on the subset V of the data set, and $\varepsilon \geq 0$ is the tolerance of ranking error.

Eq. (1) determines the minimal size of a data set that assures with high probability (at least τ) that the rank of a given feature a_i is close to its true rank. We approximate the true ranks by ranking features on the complete data set.

$$|V|_i^{min} = \arg \min_{|V| \in [1,n]} [P(|R(a_i) - R(a_i^V)| \leq \varepsilon) \geqslant \tau] \tag{1}$$

For example, if we set $\tau = 0.95$, we expect that with 95% probability we will not make error larger than ε when estimating feature a_i from sample of size $|V|_i^{min}$ instead of from the complete data set. We expect to find sample sizes $|V|_i^{min}$ which will be robust to the variations in the randomly sampled training data for the given feature and the selected ranking function R. Discussion on stability of feature evaluation can be found in [6]. We use bootstrap sampling [4] to obtain confidence intervals (CI) for determination of $|V|_i^{min}$.

For practical use we propose two variants of Eq. (1). We are interested if a given feature might be useful in a predictive model, in this case its rank has to be lower than a prespecified rank threshold L. On the other hand, we are also interested if a given feature can be safely discarded from further consideration. In this case its rank has to be higher than a threshold H. Both cases are formalized below in Eqs. (2) and (3) and can be estimated with ranking function R applied to bootstrap samples.

$$|V|_i^L = \arg \min_{|V| \in [1,n]} [P(R(a_i^V) \leq L) \geqslant \tau] \tag{2}$$

$$|V|_i^H = \arg \min_{|V| \in [1,n]} [P(R(a_i^V) \geq H) \geqslant \tau] \tag{3}$$

4 Experiments

The aim of our study is to show a practical method how to estimate the number of instances needed for a new feature to reliably estimate its rank within an existing set of features and existing data set. Our procedure is as follows. For each feature we gradually increase the sample size $|V|$, randomly select a sample of this size from the full data set 30-times, and bootstrap each sample 500 times. The bootstraped samples are used with ranking function ReliefF and form a basis to calculate the median and confidence interval (CI) for each size. The collection of this estimates is illustrated with pseudo code in Algorithm 1. Actual experiments are run within R environment using libraries *caret* [7], *CORElearn* [9] and *ggplot2* [10].

Algorithm 1 Distribution of feature ranks for different number of instances
---
1: **procedure** SUBSETSIZES(parameters: data, numExperiments, initialSize, step)
2:     subsetSize = initialSize
3:     **while** subsetSize $\leq$ size(data) **do**
4:         **for** q in 1:30 **do**
5:             sampleData = Sample(data, subsetSize, replace = FALSE)
6:             **for** k in 1: 500 **do**
7:                 trialData = BootstrapData(sampleData, subsetSize, replace = TRUE)
8:                 trialRanks[k] = ReliefF(trialData)   ▷ get ranks for all features
9:             **end for**
10:            medianRanks[q] = median(trialRanks)   ▷ compute median ranks for all features
11:        **end for**
12:        Store medianRanks for current subsetSize
13:        subsetSize = subsetSize + step
14:    **end while**
15:    Return stored rank distributions for all sample sizes
16: **end procedure**

4.1 Results

In practice, users are providing instances of data in small chunks. In order to estimate feature impact we can use only these instances. To account for variance in the obtained sample provided by users we use bootstrap confidence interval estimation which uses sampling with replacement.

First, we analyze features in the existing data set to see what can we expect for new features. We are particularly interested in top performing features (which we want to retain) and least performing features (which we can safely discard). Our testing data set contains 22 features. Based on previous research [2, Figure 2], we know that 8 features can be sufficient for random forest classifier to reach satisfactory performance, therefore we set the threshold L to 11 (incorporating a safety band of 3 (this would correspond to $\epsilon = 3$ in Eq. 1)). We set the threshold for discarding the highest ranking features to 15. Results of experiments produce figures similar to Figures 2 and 3. From the distributions depicted with box-and-whiskers plots we can even visually determine how many instances are required to reliably recognize top ranked feature's and how many to reliably discard the features with high ranks.

Figure 4 shows results we obtained for all existing features. We simulated a scenario where a new feature described with 60 instances is provided. Based on that we can provide the following guidelines for a user with a given number of available instances describing a new feature. To estimate feature's rank, perform 500 repetitions of bootstrap sampling and feature evaluation with ReliefF. The median rank from bootstrap repetitions shall be recorded and compared with rank distributions of existing features using the same number of instances. Figure 4 gives an example of rank distributions for 60 instances. E.g., if the median rank of a new feature would be 5, the horizontal line passing the rank 5 reveals which features exhibited similar behavior with this number of instances. Based on that one can take one of the three decisions: a) if the obtained rank line crosses distributions of mostly top ranked features, retain the feature and use it in the model from that time onwards, b) if the rank's line crosses mostly distributions of least ranked features, discard the feature, or c) if neither a) or b) is true, postpone the decision and try to collect more data (depending on the effort, cost of data collection, etc).
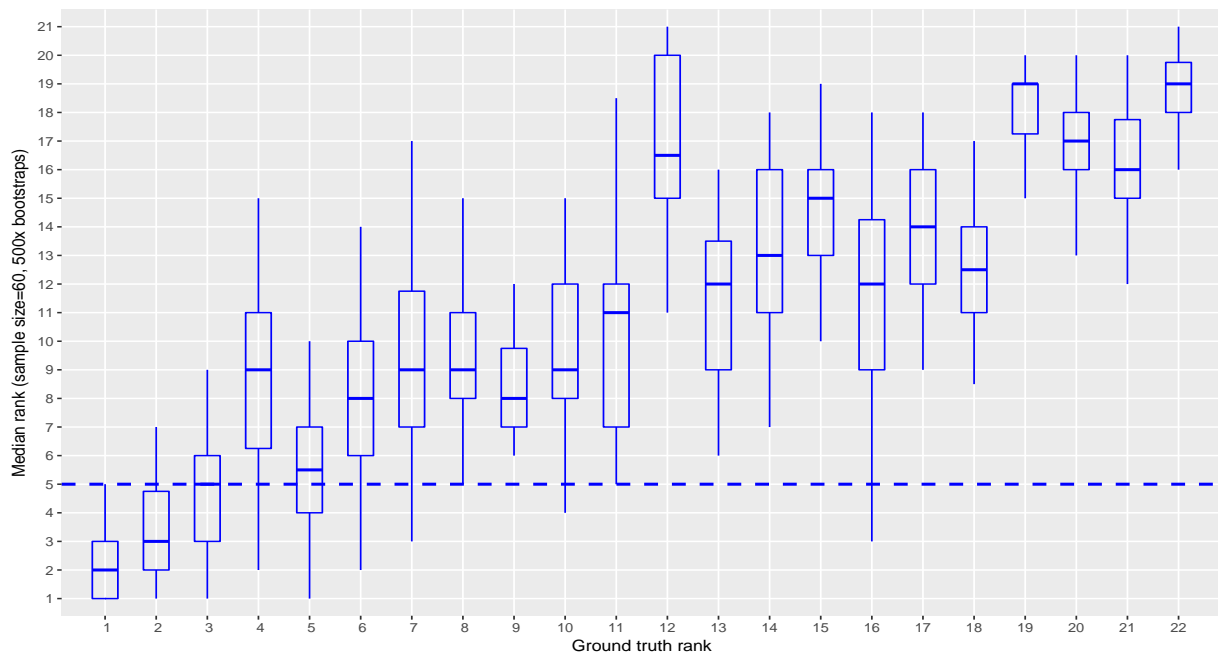
Figure 4: Rank of all features, based on sample size 60 with 500 bootstraps.

5 Conclusions

We address the problem of updates to the existing classification models as a result of changed problem understanding of domain experts. Experts consider to add various new features to the classification model and are interested to assess their potential impact with minimal data collection effort. For this purpose, we formalize a problem of minimal number of instances needed to reliably estimate the impact of new features added to the existing data set. We use the existing data set as a proxy for ground truth ranks.

The results on the analyzed B2B data set show that relatively low number of instances is required to determine impact of top performing features and least performing features. To generalize our findings the approach has to be tested for other problems and different fields. The proposed approach can be further enhanced with alternative ranking functions, sampling strategies, and reliability parameters.

Acknowledgement

References

[1] Bohanec, M. (2017). A public B2B data set used for qualitative sales forecasting research. http://www.salvirt.com/research/B2Bdataset/. [Online; accessed Avgust 2017].

[2] Bohanec, M., Kljajić Borštnar, M., and Robnik-Šikonja, M. (2015). Feature subset selection for B2B sales forecasting. In *13th International Symposium on Operational Research, Bled, Slovenia*, pages 285–290.

[3] Bohanec, M., Kljajić Borštnar, M., and Robnik-Šikonja, M. (2016). Sample size for identification of important attributes in B2B sales. In *16th International Conference on Operational Research, Osijek, Croatia*, page 133.

[4] Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*, volume 1. Cambridge University Press.

[5] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182.

[6] Kalousis, A., Prados, J., and Hilario, M. (2007). Stability of feature selection algorithms: a study on high-dimensional spaces. *Knowledge and information systems*, 12(1):95–116.

[7] Kuhn, M. (2017). A short introduction to the caret package. http:https://cran.r-project.org/web/packages/caret/vignettes/caret.pdf. [Online; accessed Avgust 2017].

[8] Robnik-Šikonja, M. and Kononenko, I. (2003). Theoretical and empirical analysis of ReliefF and RReliefF. *Machine learning*, 53(1-2):23–69.

[9] Robnik-Šikonja, M. and Savicky, P. (2017). *CORElearn - classification, regression, feature evaluation and ordinal evaluation*. R package version 1.51.2.

[10] Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

MODELLING ENERGY EFFICIENCY OF PUBLIC BUILDINGS BY NEURAL NETWORKS AND ITS ECONOMIC IMPLICATIONS

Adela Has
Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek
Gajev trg 7, 31000 Osijek, Croatia
adela.has@efos.hr

Marijana Zekić-Sušac[1]
Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek
Gajev trg 7, 31000 Osijek, Croatia
marijana@efos.hr

Abstract: Machine learning methods, such as artificial neural networks, have shown their success over statistical methods in previous research. However, they have not been exploited enough for the purpose of efficient prediction of energy efficiency. In the domain of public buildings owned by state, improving energy efficiency could significantly save the state budget. Therefore it is important to estimate the influence of characteristics of buildings and their interdependence in order to decide how to allocate resources for the reconstruction of public buildings. In this paper, methodology of neural network is used on the real dataset of Croatian public buildings covering the input space of 130 building attributes. After data pre-processing, two approaches of variable selection were used, based on statistical methods and sensitivity analysis. The most accurate model was selected, and economic implications of suggested model are also discussed. The results show that neural network methodology has the potential in predicting energy efficiency and estimating important features for classifying buildings.

Keywords: machine learning, artificial neural networks, high-dimensional data, energy efficiency, public buildings

1 INTRODUCTION

According to European Commission directives it is necessary to reduce greenhouse gas emissions, increase energy efficiency and use 20% of energy from renewable resources until 2020 [3]. Several institutions in Croatia were founded by the government in order to record and measure energy efficiency of buildings. The Agency for Legal Trade and Real Estate Brokerage (APN) manages the centralized information systems of energy efficiency in public buildings, while the Center for Monitoring Business Activities in the Energy Sector and Investments (CEI) aims to find solutions for improving the financial effectiveness of state companies in the energy sector. However, there is a lack of using intelligent data analysis in public buildings energy efficiency. The data reveal that for many public buildings in Croatia, the energy efficiency level is missing which aggravates the resource allocation for reconstruction measures. In order to deal with that issue, a model that will be able to estimate the energy efficiency level of a building based on construction data is suggested. The model is created using artificial neural networks (ANNs), a machine learning method that has shown success in prediction, classification, and association problems.

2 PREVIOUS RESEARCH

A brief overview of methods used to model energy efficiency of public buildings conducted by Zekić-Sušac [15] and Zhao and Magoulès [16] shows that some of the authors use purely statistical methods such as linear regression, time series analysis, probability density

---

[1] Corresponding author

functions, while others combine or competitively compare statistical methods with machine learning methods or use simulation modelling. If the variable selection is observed, the analysis revealed that most of the authors use building physical characteristics in addition to weather data to predict energy consumption, while some authors also use occupation data. Tsanas [14] analyzed the effect of eight input construction variables, such as relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution on two output variables, namely heating load (HL) and cooling load (CL). They used linear regression in comparison to random forest as a nonlinear non-parametric method, and concluded that nonlinear methods were more accurate. The potential of NNs as a design tool in services engineering of buildings is emphasized by Kalogirou [7]. Besides the methodology issue, there is the issue of accuracy in predicting energy efficiency. De Wilde [2] warns about a significant difference between predicted energy performance of buildings and actual measured energy use in the operational phase of buildings. He analyzes the performance gap and concludes that it can be overcome by a coordinated approach that includes model validation and verification, improved data collection for predictions, better forecasting, and change of industry practice. In our paper we try to overcome this gap by using a machine learning method on a pre-processed data, with variable reduction methods and a thorough cross-validation procedure. Economic implications of energy efficiency prediction were investigated by methods such as net present value (NPV), internal rate of return (IRR), payback analysis (PB), life cycle cost analysis, and marginal costs of energy efficiency [5][6][8][9].

3 METHODOLOGY

Prieto et al. [12] show that ANNs have proven its competitiveness in solving problems in simulators, implementations, and real-world applications for a number of years. Masters emphasize the ability of ANNs to approximate any nonlinear mathematical function [10], while Paliwal and Kumar describe its advantages over standard statistical methods in both regression and classification type of problems [11]. In this paper a most common type of ANNs is used, the multilayer perceptron (MLP) suggested by Werbos in 1974 and improved by Rumelhart et al. who introduced a backpropagation algorithm for error minimization [11]. A typical MLP consists of an input layer, one or more hidden layers, and an output layer. The input layer loads data from the input vector X of n elements with values $x_i \in R$, $i=1,2,...,n$, and randomly determined initial weights w_i usually from the interval [-1,1]. The weighted sum of all x_i values is forwarded from the input layer to a hidden layer, which uses an activation function to produce its output according to:

$$y_c = f\left(\sum_{i=1}^{n} w_i x_i\right) \tag{1}$$

The activation function f can be experimentally set as sigmoid, tangent hyperbolic, exponential, and linear, step or other [10]. The output layer computes the local error ε as the difference between the output y_c produced by the activation function and the actual output y_a. The error ε is then used to adjust the weights of the input vector according to a learning rule, usually the Delta rule [13]. The described process is repeated for another input vector in a number of iterations (epochs) until the minimum error is reached. Besides standard backpropagation algorithm suggested by [13], in this paper we additionally test the conjugate gradient descent and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithms [1]. In order to find the best ANN topology, the number of hidden units varied from 1 to 40, different activation functions were tested, while the training time is determined in an early-stopping

procedure which iteratively trains and tests the network on a separate test sample in a number of cycles, and saves the network which produces the lowest error on the test sample. In order to test the stability of ANN result, a 10-fold random subsampling cross-validation procedure is used.

4 DATA

In this research a real dataset obtained from The Agency for Legal Trade and Real Estate Brokerage (APN) in Croatia were used. It initially consisted of public buildings with 131 attributes describing their geospatial, construction, heating, cooling, meteorological and energy coefficients data. Due to a large portion of missing data, some data pre-processing procedures were necessary. After data cleaning and transformation, the sample consisted of 668 buildings for which the real energy efficiency level was available. This sample was used to build a prediction model with 74 input variables. The output variable was categorical and consisted of five categories representing energy efficiency level (AB, C, D, E, and FG). The variables used for modelling are presented in Table 1.

Table 1: Variables used for modelling

No.	Group of variables	Variable description
1.	Geospatial data	county, object region, type of object, object geo type, cultural heritage building
2.	Construction data	share of use of total building area, year of completion of construction, year of last restoration, flat gross floor area of building, useful area surface of building, object dim cooled area, object dim cooled surface area, object dim cooled volume area, number of floors, internal project temperature, share of windows surface
3.	Heating data	heated surface of the building, heated volume area of the building, installed power el. motor for pumps heat, type of heat pump, energy generating product, heating pump, total heat capacity of heat pump, total body heat radiator, total power body heat radiator, total body heat function oil, total power body heat function oil, total body heat other, total power body other, thermal power of heaters, primary heat sys using electrical heaters, installed capacity of electrical heaters, primary heating sys using split sys, installed electrical power of split sys heat, installed heat power of split sys heat, total heating power, factor of building shape f0, h1max. allowed coefficient of transmission heat loss per surface, transmission coefficient of heat loss, annual thermal energy needed 4heat
4.	Cooling data	object dimension of cooled area object dimension of cooled surface area object dimension of cooled volume area
5.	Meteorological data	air temperature
6.	Occupational data	number of employees, number of users, number of working days per week, number of working days per year, no of working hours per workday
7.	Energy coefficients of 9 specific parts of buildings	object construction coeff. trans d1,…,d9 object construction iso. thickness d1,…d9 object construction surface d1,…d9 object construction thickness d1,…d9 (d1=roof, d2=floor, d3=windows, d4=shades, d5=heated ceiling, d6=unheated ceiling, d7=external wall, d8=doors, d9=unheated wall)
8.	**Output variable**	**level of energy efficiency**

In order to obtain systematic training, testing and validation of NNs, the sample is divided into three subsamples, using equal distribution of output variables in the train (60% of cases) and test sample (20% of cases), while the rest of the cases is added to the validation sample

(20% of data). Since the instability of the result is one of the disadvantages of NN methodology, a 10-fold cross-validation procedure is performed, such that 10 random subsamples for training, testing and validation were generated and used to create 10 ANNs. The validation result of each ANN was captured and the average accuracy out of 10 validation samples is reported as the expected NN accuracy on future data.

5 RESULTS

In order to develop a successful ANN model for recognizing an energy efficient level of public buildings and identifying their most important characteristics, three models were created. The initial NN model included all available variables in the input space. After that, two variable selection procedures were conducted: (1) statistical chi-square test (p-value <0.1) and (2) sensitivity analysis. The initial NN model included 74 independent variables that describe characteristics of public buildings and climate data. The best initial ANN model produced the total classification rate of 43.34% which is very low. It is assumed that a large number of attributes brings some noise to the model, therefore the next step was to conduct variable reduction procedures.

5.1 Variable selection

The selection of variables in pre-modelling phase was performed by the *chi-square test*. On each of 74 input variables the *chi-square test* was conducted measuring the connection between an input variable and the output, and the predictors with *p-value* <10 % were selected. The selection process extracted 43 important variables that were used in the next step as input for a reduced NN model. Based on the initial NN model, a sensitivity analysis was conducted in the post-modelling phase during the 10-fold cross-validation procedure including 74 initial input variables. Higher sensitivity coefficient indicates a higher impact to the output. Only 15 variables had the sensitivity coefficient higher or equal to 1, and they belong to the groups of geospatial, construction, heating, cooling and meteorological data. Those variables with sensitivity coefficient higher than 1 were selected and used as the input variables for the second reduced NN model.

5.2 Reduced neural network models

In the next phase of modelling, the two reduced NN models were created: (1) with variables that were excluded by chi-square test, and (2) with variables excluded by the sensitivity analysis. The most accurate results of new NN models on the validation sample data are presented in Table 4.

Table 2: Results of reduced NN models based on variable selection process - validation sample

Model	NN structure	Activation function	Learning algorithm	Total classification rate - validation sample (%)	Average classification rate – 10-fold cross validation (%)
1 – Selected variables (chi-square - p value <0.1)	MLP 74-2-5	Logistic	BFGS	53.22	**40.60 (all categories)** *12.15 (AB), 56.79(C), 40.11(D), 9.07(E), 66.48(FG)*
2 – Selected variables (sensitivity analysis)	MLP 26-12-5	Tangent hyperbolic	BFGS	84.61	**75.11* (all categories)** *78.80 (AB), 79.26(C), 69.29(D), 65.16(E), 82.07(FG)*

In order to compare the obtained results, the statistical test of difference in proportions was conducted. The results show the ANN model based on sensitivity analysis is significantly more accurate than the ANN model based on the chi-square test as the criterion of variable selection (p=0.000, N=133). The most successful ANN model is able to classify public buildings according to their energy efficiency with the accuracy of 84.61%. The structure of the best ANN model is 26-12-5, meaning that it consists of 26 units in the input layer, 12 units in the hidden layer and 5 units in the output layer. The most accurate results were produced by the tangent hyperbolic activation function and BFGS learning algorithm. The average accuracy across all subsamples in the 10-fold cross-validation procedure was 75.11% and it represents the expected accuracy on new data. Besides the total classification rate on each subsample, separate classification rates for each category of the output variable were calculated. It can be noticed that the ANN model is more accurate in recognizing public buildings with lower energy efficiency level (FG). Additionally, the sensitivity analysis is performed on such reduced NN model, showing that variables that belong to the group of heating data, such as *type of heating pump*, *heated volume area of the building*, geospatial are the most influential in this reduced NN model, followed by the variables such as *type of building, building region* and construction variables such as *flat gross floor area of building, useful area surface of building*.

6 ECONOMIC IMPLICATIONS OF ENERGY EFFICIENCY OF PUBLIC BUILDINGS

Besides ecological, there are also economic reasons for increasing energy efficiency level of public buildings. There is an extensive number of public buildings that are energy inefficient – mostly older buildings with lower quality material such as isolation, quality of windows, roof etc. while energy price is increasing. The results is higher energy consumption of buildings and consequently higher energy cost for public buildings' users. Besides identification of most problematic areas of buildings and necessary improvements it is necessary to create financial projections and returns on investment in improving energy efficiency. Investments in energy efficiency need to be cost effective meaning that is important to estimate the energy and financial savings that would be generated by necessary improvements in building. For that purpose an extensive number of financial analyses, such as net present value (NPV), internal rate of return (IRR), payback analysis (PB), life cycle cost analysis, and marginal costs of energy efficiency can be used [5][6][8][9].Through simple methods such as *payback time*, the return on investment (ROI) can be calculated, and resources can be allocated to those buildings that are candidates for higher ROIs. Also, Kneifel [8] emphasize that the life cycle cost analysis (LCC) of buildings is able to determine if future operational savings justify initial investments in repairment. Economic benefits from energy efficiency investments in public buildings are multiple; lower energy and maintenance costs, increase property values, increase energy security, improved health and well-being of users, better allocation of revenues etc.[4]. Therefore, research efforts in creating the accurate model that will be able to predict future energy efficiency are of high economic interest.

7 DISCUSSION AND CONCLUSION

The paper deals with the problem of resource allocation aimed to increase energy efficiency of public buildings. For that purpose, a prediction model based on artificial neural networks is created on the real dataset of Croatian public buildings. Due to a large number of attributes, two variable reduction methods were used in order to improve the model efficiency. The

results showed that the NN model based on multi-layer perceptron with the variables selected by sensitivity analysis has produced the highest accuracy. The model can serve as a baseline for further research in this area, and is somewhat consistent to research in other countries showing that various groups of variables are important. In order to improve the model accuracy, more machine learning methods can be tested, and some other methods for variable reduction, such as principal component analysis, could be combined with NNs. It would be beneficial to conduct clustering in the pre-processing phase, which could enable the creation of separate models for each cluster. Such models have economic implications since they could serve as a support for estimating savings in reconstruction measures, and better allocation of state budget and other resources aimed to increase energy efficiency of public buildings.

References
[1] Dai, Y-H., 2002. Convergence properties of the BFGS algorithm, SIAM Journal of Optimization, Vol. 13, No. 3, pp. 693-701.
[2] De Wilde, P. (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. Automation in Construction, 41, 40-49.
[3] European Commission, Energy Efficiency Directive, https://ec.europa.eu/energy/en/topics/ energy-efficiency/energy-efficiency-directive, [18.6.2017.]
[4] European Investment Bank, The Benefits of Energy Efficiency, http://www.eib.org/epec/ ee/documents/factsheets-energy-efficiency-en.pdf, [18.6.2017.]
[5] Jackson, J. (2010). Promoting energy efficiency investments with risk management decision tools. *Energy Policy, 38*(8), 3865-3873.
[6] Jakob, M. (2006). Marginal costs and co-benefits of energy efficiency investments: The case of the Swiss residential sector. *Energy policy, 34*(2), 172-187.
[7] Kalogirou, S. A. (2006). Artificial neural networks in energy applications in buildings. International Journal of Low-Carbon Technologies, 1(3), 201-216.
[8] Kneifel, J. (2010). Life-cycle carbon and cost analysis of energy efficiency measures in new commercial buildings. *Energy and Buildings, 42*(3), 333-340.
[9] Martinaitis, V., Kazakevičius, E., & Vitkauskas, A. (2007). A two-factor method for appraising building renovation and energy efficiency improvement projects. Energy Policy, 35(1), 192-201.
[10] Masters, T., 1995. Advanced Algorithms for Neural Networks, A C++ Sourcebook, John Wiley & Sons, Inc., New York, USA.
[11] Paliwal, M. and Kumar U.A., 2009. Neural networks and statistical techniques: A review of applications, Expert Systems with Applications, Vol. 36, pp. 2–17.
[12] Prieto, A., Prieto, B., Martinez Ortigosa, E., Ros, E., Pelayo, F., Ortega, J., Rojas, I. (2016). Neural networks: An overview of early research, current frameworks and new challenges, Neurocomputing, in press, Available online 8 June 2016, ISSN 0925-2312, http://dx.doi.org/10.1016/ j.neucom.2016.06.014.
[13] Rumelhart, D. E., Hinton, G.E., Williams, R.J. (1989), Learning resentations by back-propagating errors, in Anderson, J.A. and Rosenfeld, E. (eds.) Neurocomputing: Foundations of Research, MIT Press, A Bradford Book, Cambridge, MA, USA.
[14] Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy and Buildings, 49, 560-567.
[15] Zekić-Sušac, M. (2017). Overview of prediction models for buildings energy efficiency, Proceedings Of The International Scientific Symposium „Economy Of Eastern Croatia – Vision And Growth", Anka Mašek Tonković (Ed.), Osijek, May 25-27, 2017, pp. 697-706.
[16] Zhao, H. X., & Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586-3592.

IT SECURITY GOVERNANCE AND MANAGEMENT BEST PRACTICES: ASSESING THEIR MATURITY IN A LARGE SPANISH COMPANY

Eloy Hontoria[a]**, Danijel Kovačić**[b]**, Wim Van Grembergen**[c]
[a] Technical University of Cartagena/Business Management**,** 30202 Cartagena, Spain
[b] MEDIFAS, Mednarodni prehod 6, SI-5290 Šempeter pri Gorici, Slovenia
[c] University of Antwerp/ Information Systems Management**,** 2000 Antwerpen, Belgium
eloy.hontoria@upct.es, kovacic.danijel@gmail.com, wim.vangrembergen@uantwerpen.be

Abstract: Organizations have to keep safety of one of their most value assets: Their information and technological components. Some firms are certificated in any IT Security standard like ISO 2700, but this is not enough to deal with IT Risks if the board is not convinced of the dangerousness for their resources. The objective of this research in progress is to present a methodology for the design of a steering tool focused on the board of the company to determine if the company is performing a suitable IT Security Governance to support IT Security Management and avoid IT Risks. The suitability of this steering tool will be checked through a case study of a large Spanish company whose IT Security Management is correct but board is not aware about IT Risks. The results of this case study may be used as a guideline for enterprises to assess their IT Security performance.

Keywords: IT Security Governance, IT Risks, Governing Body, Best Practices, ISO 27014.

1 INTRODUCTION

IT Security has become a top of mind among board members, executives and security professionals. In May 2017, a massive ransomware attack infected 230,000 computers in over 150 countries. It was not the first and unfortunately will not be the last one.

One of the most important assets a company has are the information and technological components, but a lack of an adequate governance over information stored, processed, or produced may have a significant negative impact like: unauthorized computer system use, unauthorized access to information that results in a loss of data, penalties related with non-compliance of legislation, damaged reputation that can take time and money to be rebuilt, etc.

On the other hand, benefits of good information security are not just a reduction in risk or a reduction in the impact if something goes wrong. Good security can improve reputation, confidence and trust from others with whom business is conducted, and can even improve efficiency by avoiding wasted time and effort recovering from a security incident [6]. Thus, information security governance requires senior management commitment, a security-aware culture, promotion of good security practices and compliance with policy. It is easier to buy a solution than to change a culture, but even the most secure system will not achieve a significant degree of security if used by ill-informed, untrained, careless or indifferent personnel [6].

An explanation to these attacks is the fact that in many firms, there is a disconnection between senior management and IT due to the view that IT exists solely to deliver day-to-day IT services [16].

Board should be concerned about the strategic importance of IT and the first step is to assess how are IT Risks (information confidentiality, integrity and availability) and resources managed (human and technical). The success of IT depends on the guidelines provided by the board, CEO, and other members of senior management, which should be communicated through the organization's strategic plan and structure.

Former approach is considered a top-down approach and the best option to govern IT Security, but: what happens in a bottom-up approach where IT Managers are accountable of

IT Security and are not supported by the board? To address this question a case study is conducted, where Company XX has a solid IT Security management (is certified on ISO 27001) but IT Risks are high.

This work is organized as follows: Section 2 and Section 3 are focused in IT Governance core concepts and ITSG principles and their importance in an organization. Section 3 describes the methodology established to carry out this work. In Section 4 a case study is deployed in order to assess ITSG in Company XX. Finally Section 5 summarizes key findings of this work and our conclusions.

2 IT GOVERNANCE

Today, IT is more critical to the business than ever [7], being considered in the past as a cost-only factor when they have to contribute to the achievement of the business objectives.

IT governance (ITG) is an area of corporate governance [17], [18] that is a responsibility of the board of directors [12] and executives [5]. ITG can also be defined in terms of "processes, structures and relational mechanisms in the organization that enable both business and IT people to execute their responsibilities in support of business/IT alignment and the creation of business value from IT-enabled business investments" [17].

But, companies should improve the integration of IT Governance with risks aspects. In our approach we want to show links between IT Governance and risk management and to attain this goal, best practices on IT Security Governance will set up the foundations.

3 IT SECURITY GOVERNANCE vs IT SECURITY MANAGEMENT

ITSG is an integral part of Corporate Governance [15], and is not the same as IT management because governing is not managing day-to-day activity but directing and controlling the organization, ensuring that shareholders' and stakeholders' desires are met [11]. In so doing, organizations ITSG will deliver value to the governing body and stakeholders and ensure that information risks are being adequately addressed [9].

Johnston and Hale [4] confirmed empirically that organizations that address their IT Security from the bottom up, and isolate the governance from the management of IT, have ineffective IT programs and can fall victim to internal and external cybersecurity attacks, in contrast to organizations whose ITSG programs have a proactive, top-down approach.

These companies have strategies to protect their assets based on incidents at the perimeter of the organization and at the same time they segregate information security from the top, thereby creating a division between the governance and the management of IT Security and the results of such disconnection can be disastrous [4]. Management implements information security components, such as policies and technical security measures, but to inculcate an acceptable level of information security culture, organizations must govern information security effectively by implementing all the required information security components [3].

With Security Governance at the top level of the firm, the board are providing the adequate care about security on all levels and there will be a high level of integration of data and IT Security in the business governance and decision making structure. An Information Security culture must emerge from the top with mandatory compliance by anyone, because organizations can implement security controls such as anti-virus programs, firewalls, and passwords but there is no sense in implementing these controls if users share passwords and connect through dialup to the Internet, by passing the firewall [3].

In order to summarize and according to ISO/IEC 27014:2013, Governance of Information Security needs to align objectives and strategies for information security with business objectives and strategies, and requires compliance with legislation, regulations and contracts.

It should be assessed, analyzed and implemented through a risk management approach, supported by an internal control system.

To assess the robustness of IT Security Governance in Company XX a methodology has been developed and it is shown in Section 4.

4 METHODOLOGY

In the first step of the methodology, a deep literature review in the academic and practitioner field will be carried out. The goal of this research is to find out a comprehensive set of best practices in ITSG a firm should comply in this field.

The goal of this second step is to apply an exhaustive method to filter previous set of best practices, regarding to have a reduced and selective list of them. To attain this goal a Delphy Study will be designed to validate and prioritize former set collected at the first step.

4.1 Best Practices on ITSG Literature Review

Primary objective of this paper is to validate and prioritize a set of best practices on ITSG to be checked furtherly in the pilot study. To attain this goal, our work started with a literature review both in the academic and practitioner field, being the main topic of interest the state of the art related with IT Governance, IT Security Governance, IT Security and IT Risks.

The result of the state of the art, has been a comprehensive set of best practices extracted mostly from:

a) *Reputed Standards, Frameworks, Universities:* **ISO/IEC 27014, COBIT 5 (ISACA),** Information Technology Governance Institute (ITGI), Software Engineering Institute (SEI) from Carnegie Mellon University [8].

b) *Practitioners organisations:* National Association of Corporate Directors (NACD) and National Cyber Security Summit Task Force.

c) *Researchers*

As a consequence of the literature review, a comprehensive list of 64 Best Practices were selected, but this should be filter regarding its manageability.

4.2 Delphy Study

To validate and prioritize former best practices collected on Section 4.1, the expert´s opinion in this field is indispensable. For this purpose Delphy research methodology can be very powerful. According to [10] "Delphi method is a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem".

Taking into account that this work is focused in analyzing the state of Information Technology Security Governance at the private sector, selected members from the Delphy Study belong to this kind of companies. In the first step of the Delphy Study, the objective was to reduce the initial list of 64 Best Practices to a manageable new one. As a result a new Best Practices list containing just 21 of the former set was obtained. The purpose of the second step of the Delphy Study was to prioritize the final set of 21 best practices. For this reason, experts were requested to determine the importance of any of the previous 21 Best Practice in regarding to attain a robust ITSG at a private company. Answers were marked in a scale from "0" up to "5", being "5" –Absolutely important- and "0" –Nothing important-.

4.3 Questionary

As a result of the Delphy Study a questionary was designed to be launched to Chief Information Officers (CIO´s). In former questionary, members were requested to answer at what extent a Best Practice was fulfilled in their companies.

The result of the Delphy Study is a questionary consisting in 21 Best Practices based on IT Security Governance which have been selected and prioritized as it is depicted in Table 1.

	IT SECURITY GOVERNANCE BEST PRACTICES	How much important is this Best Practice to perfom ITSG?					
		0	1	2	3	4	5
1	Understanding the criticality of information and information security to the organisation (Board have to understand why IS must be governed)						x
2	Board should understand implications of IT Risks to enterprise strategic objectives						x
3	Board should understand their responsibilities and roles with regard to IT Risks						x
4	Board must be aware about organisation's information assets and their criticality to ongoing business operations			x			
5	Board must validate the key assess they want to be protected and confirming that protection levels and priorities are appropriate			x			
6	Report to external stakeholders that the organisation practices a level of information security					x	
7	Ensure that business initiatives take into account information security issues						x
8	Place information security on the board's agenda					x	
9	To identify an enterprise level information Risks						x
10	To establish a security management structure to assign explicit individual roles, responsibilities and authority for managing that risk						x
11	Board must allocated appropiate investment and resources to information security efforts						x
12	Reviewing investment strategy in information security for alignment with the organisation strategy and risk profile and set priorities						x
13	Organisations should provide IS awareness, training and education to personnel						x
14	Should provide strategic oversight regarding information security						x
15	Identify information security leaders, hold them accountable and ensure support for them					x	
16	Neeeded of independent security audits						x
17	Notify executive management of the results of security audits that have identified IS security issues, and prioritize and initiate required corrective actions				x		
18	CEOs should have an IS evaluation conducted, review the evaluation results with staff, and report on performance to the board of directors	x					
19	Board must be provided periodically with the results of risk assessments and business impact analyses					x	
20	ITSG should ensure that IS policies and practices conform to legislation, commited business, contractal (internal o external) requirements						x
21	Set direction for the information security strategy, policies and and procedures based on risk assessments to secure information assets						x
		Note: "0" means Nothing Important and "5" is Absolutely Important					

4.4 Structures, Processes and Relational Mechanisms

Structures refer to committees and councils, as well as formal positions and roles for IT-related decision-making [13], [14]. Processes are focused on the implementation of IT management techniques and procedures in compliance with establishing IT strategies and policies [2]. The structures and processes need to be complemented by relational mechanisms [14] which are referred to the collaboration and active participation between corporate executives, IT management and business management [14]. Management leadership should be proactive in ensuring that the activities of IS are supported and understood at all organizational levels and aligned with organizational objectives [11].

These main components of ISG ensure that the confidentiality, integrity and availability of an organization's electronic assets are maintained all the time and information is never compromised, as well as cultivating and sustaining a security culture in the organization [1].

5 CASE STUDY

Company XX is a large Spanish firm certificated under the ISO 27001 Standard, which is a proof that its IT Security management is well performed. However, IT Security managers are concerned about IT Risks and their perception is that this risk is high, because there is not a real awareness in the firm about it. This enterprise has been chosen to prove our questionary to ascertain if the problem of this high Security Risk perception relies in the lackness of IT Security Governance mentality of its board.

For this purpose the questionary was launched to the Chief Information Officer (CIO) of the company and the results are shown in Table 2.

	IT SECURITY GOVERNANCE BEST PRACTICES	Is this Best Practice being fullfilled at your firm?					
		0	1	2	3	4	5
1	Understanding the criticality of information and information security to the organisation (Board have to understand why IS must be governed)		X				
2	Board should understand implications of IT Risks to enterprise strategic objectives		X				
3	Board should understand their responsibilities and roles with regard to IT Risks	X					
4	Board must be aware about organisation's information assets and their criticality to ongoing business operations	X					
5	Board must validate the key assess they want to be protected and confirming that protection levels and priorities are appropriate	X					
6	Report to external stakeholders that the organisation practices a level of information security	X					
7	Ensure that business initiatives take into account information security issues	X					
8	Place information security on the board's agenda	X					
9	To identify an enterprise level information Risks		X				
10	To establish a security management structure to assign explicit individual roles, responsibilities and authority for managing that risk		X				
11	Board must allocated appropiate investment and resources to information security efforts	X					
12	Reviewing investment strategy in information security for alignment with the organisation strategy and risk profile and set priorities	X					
13	Organisations should provide IS awareness, training and education to personnel	X					
14	Should provide strategic oversight regarding information security		X				
15	Identify information security leaders, hold them accountable and ensure support for them	X					
16	Neeeded of independent security audits						X
17	Notify executive management of the results of security audits that have identified IS security issues, and prioritize and initiate required corrective actions				X		
18	CEOs should have an IS evaluation conducted, review the evaluation results with staff, and report on performance to the board of directors	X					
19	Board must be provided periodically with the results of risk assessments and business impact analyses				X		
20	ITSG should ensure that IS policies and practices conform to legislation, commited business, contractal (internal o external) requirements		X				
21	Set direction for the information security strategy, policies and and procedures based on risk assessments to secure information assets				X		
		Note: Mark with an "X" if Best					
		Practice is fullfilled at your firm					

6 RESULTS AND CONCLUSIONS

Figure 1 shows the result of the questionary launched to Company XX.
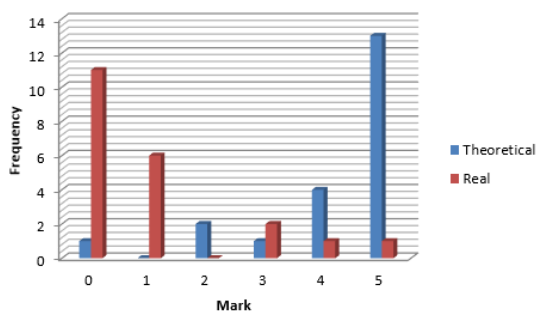
At a first sight, Company XX presents a low ITSG maturity level, due to the fact that out of 13 Best Practices' considered as "Absolutely Important", only one of them is fulfilled in this firm. Beside this important finding, out of 21 questions, 14 of them (67%) present a very high discrepancy (3 points of difference or more) between the importance given by the experts of the Delphy Study (theoretical) and the implemented actions, 3 of them (14%) present a moderate discrepancy (1 or 2 points of difference) and 4 (19%) coincide.

Additionally, we can confirm that in Company XX IT Security is addressed from the bottom up, and isolates the governance from the management of IT.

Third finding is that first six Best Practices which have a direct relation with the concerned of executive managers on IT Security are not accomplished by this enterprise.

According to these results and the literature review on IT Security Governance, we can conclude that Company XX is in a high IT Risk as it was defended by its CIO.

Although there is a lot of work for Company XX to avoid this scenario, our obligation as researchers is to give it some short suggestions on relation to Structures, Processes and Relation Mechanisms. A comprehensive action plan could be a drawback for this purpose.

One of the most important suggestions which is a key concept, derives from IT Governance and it is called "CIO on the board". This means that Chief Information Officer should be a member of the Decision Making Process at the top level of the company, transferring business decisions to IT Management and IT Risks. Two main consequences are expected to be reached with this action: To get a top-down approach on IT Security Governance and to aware executives in this field to split IT Security culture as a cascade through different areas of the company. To attain this goal, users must be formed by their superiors about its importance and consequences of non-compliance IT Security guidelines.

In the author´s opinion, the main value of this work is it use for privates companies to compare their IT Security Risks situation with Company XX and to ascertain if their assess are well protected. In order to summarize, the selected list of ITSG Best Practices' can also be a toolkit for assessing and implementing ITSG in the enterprise.

References

[1] Allen, J. (2005). Governing for Enterprise Security, Technical Note. Pittsburgh.

[2] Bowen, P. L., Cheung, M-Y. D., Rohde, F. H. (2007). Enhancing IT governance practices: A model and case study of an organisation's efforts. International Journal of Accounting Information Systems, 8, 191–221.

[3] Da Veiga, A., Eloff, J. H. P (2007). An Information Security Governance Framework," Information Systems Management, vol. 24, pp. 361-372.

[4] Johnston, A. C., R. (2009). Improved Security Through Information Security Governance, Communications of the ACM, 52 (1), 126.

[5] ITGI, IT Governance Institute (2003). Board briefing on IT Governance (2nd ed.). Rolling Meadows, IL: Author.

[6] ITGI, IT Governance Institute, COBIT® Security Baseline, USA, 2004, www.itgi.org).

[7] ITGI, IT Governance Global Status Report. (2006). http://www.itgi.org

[8] ISACA (2012). COBIT 5 for Information Security. IL, USA. Available at: www.isaca.org/cobit5info-sec

[9] ISO/IEC 27014 (2013). Governance of Information Security. Geneva: International Organisation for Standardization and the International Electrotechnical Commission.

[10] Linstone H. A., Turoff M. (2002). The Delphi Method, Techniques and Applications, Addison-Wesley, London.

[11] Love, P., Reinhard, J., Schwab, A. J., Spafford, G. (2010). GTAG Information Security Governance. The Institute of Internal Auditors, 134.

[12] Parent, M., Reich, B. H. (2009). Governing information technology risk. California Management Review, 51(3), 134–152.

[13] Peterson, R. (2003). Information strategies and tactics for information technology governance. In Van Grembergen (Ed.), (pp. 37–80). Strategies for information technology governance. Hershey, PA: Idea Group Publishing.

[14] Peterson, R. (2004). Crafting information technology governance. Information Systems Management, 21(4), 7–22.

[15] Rastogi, R., Von Solms, R. (2006). Information Security Governance-A Re-Definition, Security Management, Integrity, and Internal Control in Information Systems, 193, 223–236

[16] The institute of internal auditors (2012). Global Technology Audit Guide (GTAG®) 17: Auditing IT Governance.

[17] Van Grembergen, W., De Haes, S. (2009). Enterprise governance of information technology—Achieving strategic alignment and value. New York: Springer Science + Business Media.

[18] Wilkin, C. L., Chenhall, R. H. (2010). A review of IT governance: A taxonomy to inform accounting information systems. Journal of Information Systems, 24(2), 107–146.

FORECASTING DAILY PATIENT VISITS IN AN EMERGENCY DEPARTMENT BY GA-ANN HYBRID APPROACH

Engin Pekel

Yildiz Technical University, Department of Industrial Engineering, 34349, Istanbul, TURKEY
E-mail: <epekel@yildiz.edu.tr>

Muhammet Gul

Munzur University, Department of Industrial Engineering, 62000, Tunceli, TURKEY
E-mail: <muhammetgul@munzur.edu.tr>

Erkan Celik

Munzur University, Department of Industrial Engineering, 62000, Tunceli, TURKEY
E-mail: <erkancelik@munzur.edu.tr>

Abstract: An Emergency Department (ED) plays a crucial role in the health system by providing acute care for patients who attend hospital without prior appointment. An accurate forecasting of patient visits in EDs contributes to health care decision makers to better allocate ED human resources and medical equipment. In this paper, a hybrid genetic algorithm–artificial neural network (GA–ANN) approach is developed. The forecasting performance of the hybrid approach is obtained using the real-world data set collected from a public hospital in Istanbul, Turkey. The hybrid GA–ANN model is shown to perform well in terms of forecasting accuracy. In order to contribute to the current knowledge, this paper is a novel attempt of applying GA-ANN to model ED patient arrivals, and the results can be used to aid in strategic decision-making on ED resource planning in response to predictable arrival variations.

Keywords: forecasting, emergency department, patient visit, artificial neural network, genetic algorithm.

1 INTRODUCTION

Emergency department (ED) is one of the most important units of a hospital and is considered as heart of hospitals. The service quality through the EDs, expressed by waiting time and average total length of stay (ALOS), is significantly related to the patient visits [2, 11]. Increased patient visits at these departments may lead to prolonged waiting and overcrowding. In order to improve ED processes, a better setting between ED human resources and medical equipment and visits (patient demand) is essential. Accurate forecasting of ED patient visits can provide decision makers to arrange staffing policies so that the EDs can be better prepared for the coming visit variations [11]. There is an urgent need to develop a more realistic forecasting model to account for the anticipated ED visits to enable better strategic planning of ED resources and more effective staff arrangement. At this point, this study objects to use a hybrid approach to forecast daily ED patient visits.

Regarding the ED patient visit forecasting on long term (annual), medium term (monthly, weekly) and short term (daily, hourly), several studies have been done by various methods such as regression, time series and data mining [1, 3-6, 7, 10]. However, the current literature is satisfied enough with the work done up to now, the current works on ED patient visits forecasting cover mostly single method based models. However, in this study, ED patient visits have been forecasted based on a GA-ANN hybrid approach. We also included an environmental variable of maximum temperature and some calendar variables of day of the week, month of the year and holidays to the current study.

The remaining of the study are organized as follows: In section 2, the data and forecasting methods used in the study are introduced. In section 3, results of the application are provided. In the final section, the conclusion is presented.

2 MATERIAL AND METHOD

2.1 Data

Data was obtained from the hospital information system of a public hospital in Istanbul, Turkey. The time series spans from January 1, 2011 to December 31, 2012. The time series used for modelling is shown in Figure 1. At first, augmented dickey-fuller (ADF) test was employed to check the stationarity of the time series. The test shows that the time series is stationary. Clearly, there is no trend in the data as shown in Figure 1. However, autocorrelation function (ACF) and partial autocorrelation function (PACF) show significant autocorrelations. Therefore, it is important to investigate the forecasting performance of GA-ANN model.
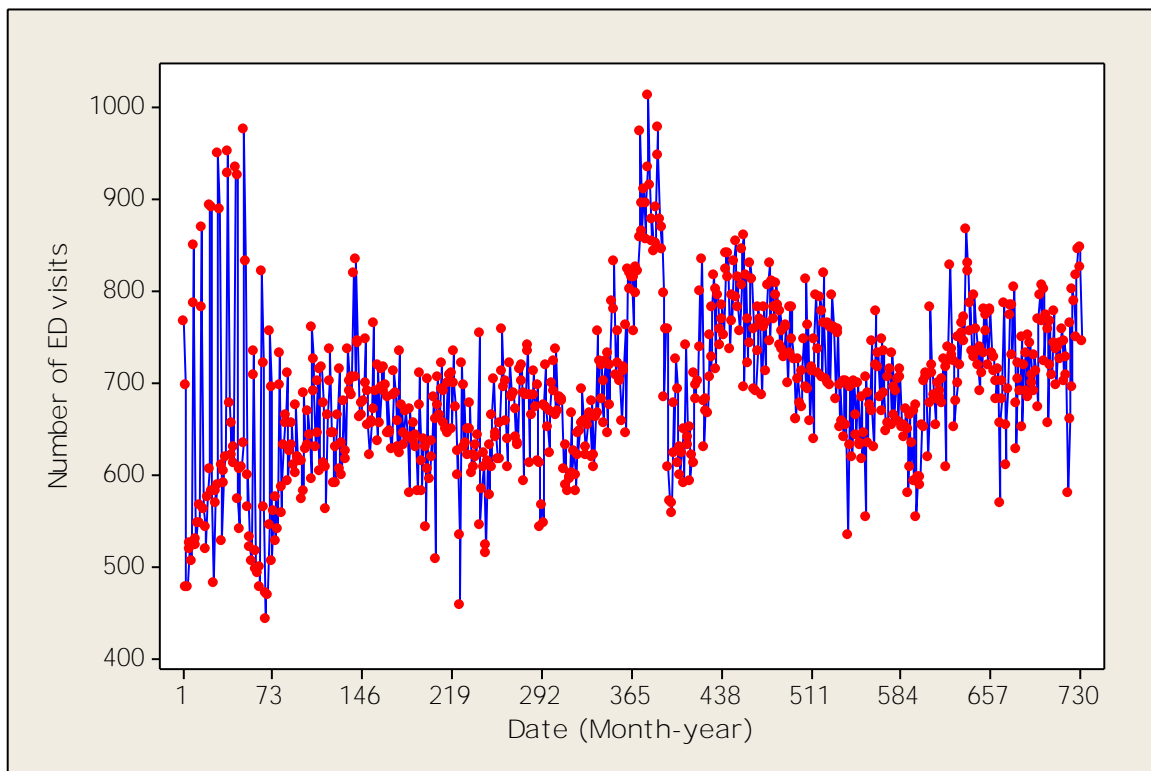
Figure 1: Real ED visits used for modelling

2.1 Methodology

This section describes the general background of GA-ANN which will be implemented in the proposed system.

GA is embarked on a set of solution pools and then the stage of natural selection is carried out by using the solution pools. In this stage, poor solutions run out and solutions, which have the highest quality, survive to reproduce. This stage is iterated until the optimal condition is satisfied. A hybrid GA-based neural network is basically a back-propagation network, with the only exception being that the weight matrix was acquired from performing the genetic operations under optimal convergence conditions [8-9].

Pseudo code of GA-ANN is illustrated in Algorithm 1. Initial weights are set off randomly in the first iteration and the output of each hidden neuron and the error are computed. Logarithmic sigmoidal transfer function and purelin transfer function are used as input and output functions, respectively. The updated weights are calculated with regard to GA by

applying parent selection, reproduction and mutation. Then, the next iteration is carried out with regard to the updated weights.

Algorithm 1. GA-ANN
1 **Initialize** weights ($w(t = 0)$)
2 t←0;
3**Compute** the output for every neuron by applying activation function
4**Compute** the error at the output
5**While** not_terminated() **do**
6 $w_p(t)$←Select_parents();
7 $w_r(t)$←Reproduction();
8 Mutate($w_p(t)$);
9 Evaluate($w_r(t)$);
10 $w_r(t)$←build_next_generation($w_r(t), w(t)$)
11**End while**
12 t←t+1;

---

3 RESULTS AND DISCUSSION

The proposed GA-ANN model is run with the combination of four different parameters, which are "population size", "crossover", "mutation" and "hidden neuron". Two values of population size are tested; 10 and 20. Ten different values for crossover (0.1, 0.2,…, 1.0), ten different values for mutation (0.1, 0.2,…, 1.0) and nineteen values for hidden neuron (2, 3,…, 20) are tested. In total, 3800 (=: $2x10x10x19$) different combinations (see Table 1) are tested to find the best (highest) R^2 for training. Because of page limitation, we did not present the whole R^2 values (except the best 10) for combinations.

Table 1: Combinations that are tested for GA-ANN hybrid approach

Population Size	Crossover	Mutation	Hidden Neuron
			2
			3
			4
			5
	0.1	0.1	6
	0.2	0.2	7
	0.3	0.3	8
	0.4	0.4	9
10	0.5	0.5	10
20	0.6	0.6	11
	0.7	0.7	12
	0.8	0.8	13
	0.9	0.9	14
	1	1	15
			16
			17
			18
			19
			20

The proposed GA-ANN have been run on a computer that has a 32-bit Windows 7 operating system, 2.4-GHz processor, and 16-GB memory. GA-ANN has been implemented in Matlab 7.12.

Table 2 shows the best 10 values of R^2 with regard to six parameters. The best combination has the highest R^2 value that equals to 0.8104 in the training stage. The result of R^2 can be

acceptable because the value is higher than 0.80. In this study, R^2 is used for the performance measurement instead of mean square error because it provides the correlation concurrently.

Table 2: The best 10 values of R^2 with regard to six parameters

Population Size	Crossover Rate	Mutation Rate	Hidden Neuron	R^2 for Training (Mean value)	R^2 for Testing (Mean value)	Standard Deviation for Training	Standard Deviation for Testing
20	0.8	0.3	19	0.8104	0.8011	0.0077	0.0117
20	0.4	0.5	12	0.8098	0.7927	0.0081	0.0063
20	0.6	0.6	20	0.8097	0.7209	0.093	0.0094
20	0.7	0.2	16	0.8089	0.7978	0.0070	0.0078
20	0.4	0.1	14	0.8078	0.7585	0.0106	0.0148
20	0.4	0.1	19	0.8077	0.7589	0.0214	0.0219
20	0.5	0.1	16	0.8072	0.7826	0.0219	0.0170
20	0.7	0.2	15	0.8071	0.7083	0.0111	0.0247
20	0.6	0.1	19	0.8070	0.7480	0.0112	0.0129
20	0.6	0.2	18	0.8069	0.7730	0.0110	0.0144

Figure 2 shows the architecture for the proposed GA-ANN. Transfer functions, which are used in the layer of GA-ANN, are logarithmic and purelin, respectively.
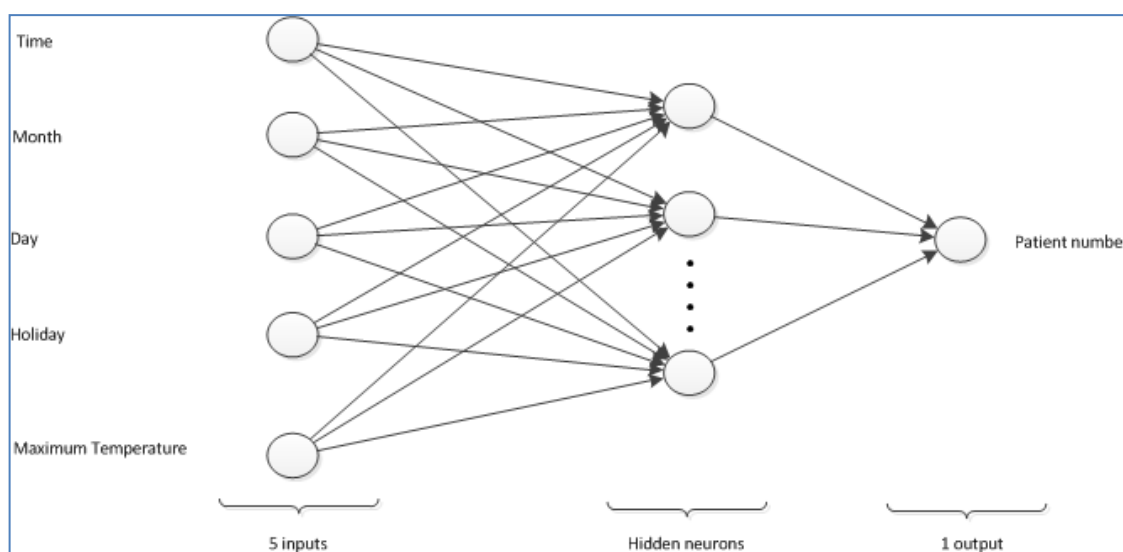
Figure 2: The architecture for the proposed GA based ANN

Various options such as population, fitness scaling, selection, reproduction, mutation, crossover and migration need to be specified as shown in Table 3. The parameters are tuned with respect to trial and error. Optimal condition of population size, crossover fraction, migration fraction and hidden neuron number are searched from 10 to 20, from 0.1 to 0.9, from 0.1 to 0.9 and from 2 to 19, respectively. Especially, when the number of population size and the hidden neuron is increased, computing time also raises.

Table 3: The optimal parameters of GA-ANN

	Options	Optimal conditions
Population	Creation Function	
	Population size	20
Fitness scaling		-
Selection		Stochastic uniform
Reproduction	Elite count	2
	Crossover fraction	0.80
Mutation		Gaussian mutation
Crossover		Heuristic crossover
Migration	Direction	Both
	Fraction	0.30
Algorithm settings	Initial penalty	10
	Penalty factor	100
Stopping criteria	Generations	1000
	Fitness limit	1.00e-8
	Stall generations	50
Hidden neuron number	Single hidden layer	19
CPU time (second)		624

4 CONCLUSION

The importance of forecasting models of ED visits is absolutely critical for resource planning in the hospitals. The proposed GA-ANN hybrid model allows daily forecasts with a good quality level. Accurate forecasting of ED visits decreases the overcrowding and optimizes the ED medical staff and space allocation to the actual demand. The current study contributes to ED patient arrival forecasting literature from two sides. From one point of view, a hybrid method outline (GA and ANN) for EDs is proposed. Secondly, it includes meteorological and calendar variables together for patient arrival forecasting. As future work, the forecasts produced by the developed forecasting approach will be compared to the forecasts that will be obtained from other newly adopted algorithms. Additionally, hourly ED visits and multi-site ED visit forecasting contexts will be dealt.

References

[1] Aladeemy, M. Chou, C-A. Shan, X. Khasawneh, M. Srihari, K., & Poranki, S. 2016. Forecasting Daily Patient Arrivals at Emergency Department: A Comparative Study. Proceedings of the 2016 Industrial and Systems Engineering Research Conference, May 2016.

[2] Au-Yeung, S.W.M., Harder, U., McCoy, E.J., & Knottenbelt, W.J. 2009. Predicting patient arrivals to an accident and emergency department. Emergency Medicine Journal, 26(4), 241-244.

[3] Carvalho-Silva, M., Monteiro, M.T.T., de S´a-Soares, F., & D´oria-N´obrega, S. 2017. Assessment of forecasting models for patients arrival at Emergency Department. Operations Research for Health Care, http://dx.doi.org/10.1016/j.orhc.2017.05.001.

[4] Cote, M. J., Smith, M. A., Eitel, D. R., & Akçali, E. 2013. Forecasting emergency department arrivals: A tutorial for emergency department directors. Hospital topics, 91(1), 9-19.

[5] Gul, M., & Guneri, A.F. 2015. Forecasting patient length of stay in an emergency department by artificial neural networks. Journal of Aeronautics and Space Technologies, 8(2), 43-48.

[6] Gul, M., & Guneri, A.F. 2016. Planning the future of emergency departments: Forecasting ED patient arrivals by using regression and neural network models. International Journal of Industrial Engineering, 23(2), 137-154.

[7] Hertzum, M. 2017. Forecasting Hourly Patient Visits in the Emergency Department to Counteract Crowding. The Ergonomics Open Journal, 10, 1-13.

[8] Kadiyala, A., Kaur, D., Kumar, A. 2013. Development of hybrid genetic-algorithm-based neural networks using regression trees for modeling air quality inside a public transportation bus. Journal of the Air & Waste Management Association, 63(2):205-218.

[9] Pekel, E., Soner Kara, S. 2017. Passenger Flow Prediction Based on Newly Adopted Algorithms. Applied Artificial Intelligence, 31(1):64-79.

[10] Wargon, M., Guidet, B., Hoang, T.D., & Hejblum, G. 2009. A systematic review of models for forecasting the number of emergency department visits. Emergency Medicine Journal, 26(6), 395-399.

[11] Xu, M., Wong, T.C., & Chin, K.S. 2013. Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. Decision Support Systems, 54(3), 1488-1498.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 6:
Mathematical Programming and Optimization

PRODUCTION PLANNING IN THE BAKERY VIA DE NOVO PROGRAMMING APPROACH

Zoran Babić
University of Split, Faculty of Economics
Cvite Fiskovića 5, 21000 Split, Croatia
E-mail: babic@efst.hr

Tunjo Perić
University of Zagreb, Faculty of Economics and Business
Trg J. F. Kennedy 6, 10000 Zagreb, Croatia
E-mail: tperic@efzg.hr

Branka Marasović
University of Split, Faculty of Economics
Cvite Fiskovića 5, 21000 Split, Croatia
E-mail: branka.marasovic@efst.hr

Abstract: Optimization of production program is one of the crucial problems of optimization. It may be difficult to apply a linear programming model in real production planning situations because of its assumption of proportionality. A common phenomenon is multiple pricing of raw materials because companies may have more than one source of raw material, which they purchase at different prices or may be offered quantity discounts for bulk purchases. In this paper, the application of De Novo programming in the real production system, which produces various bakery products and in its production uses a set of raw materials with different prices, will be presented.

Keywords: De Novo Programming, bakery, production plan, variable prices

1 INTRODUCTION

De Novo programming, initiated by Zeleny [1], presents a special approach to optimization. Instead of "optimizing a given system", De Novo suggests a way of "designing an optimal system". De Novo approach does not limit the resources, as most of the necessary resource quantities can be obtained at certain prices. Resources are actually limited because their maximum quantity is governed by the budget, which is an important element of De Novo.

Most cases can be handled more effectively using De Novo than by using the standard programming model [2]. Changes in prices, technological coefficients, increasing costs of raw materials, quantity discounts and other similar and real production situations can be easily incorporated into the De Novo model and can provide very satisfactory solutions.

In this paper, two different approaches to optimization with the De Novo concept of optimization: increasing costs and quantity discounts of raw materials will be presented. This paper presents the advantages of De Novo approach in the production plan optimization in a concrete bakery.

2 DE NOVO PROGRAMMING

The traditional resource allocation problem in economics is modeled via standard linear programming formulation of the single-objective product-mix problem.

In De Novo formulation the purpose is to design an optimal system and the following formulation is of interest:

$$\text{Max } z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \qquad (1)$$
$$\text{s.t.} \quad a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = b_1$$

$$a_{21}\,x_1 + a_{22}\,x_2 + \dots + a_{2n}\,x_n = b_2$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \tag{2}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_{m1}\,x_1 + a_{m2}\,x_2 + \dots + a_{mn}\,x_n = b_m$$

$$p_1\,b_1 + p_2 b_2 + \dots + p_m\,b_m \le B \tag{3}$$

$$x_j,\; b_i \ge 0,\; j = 1, 2, \dots, n;\; i = 1, 2, \dots, m \tag{4}$$

where:

$b = (b_1, b_2, \dots, b_m)$ – set of decision variables representing the level of resource i to be purchased,

p_i – unit price of resource i,

B – total available budget for the given system.

Now the problem is to allocate the budget so that the resulting portfolio of resources maximizes the value of the product mix (with given unit prices of m resources, and with given total available budget).

The main difference of the two models lies in the treatment of the resources which become decision variable b_i in the De Novo formulation.

2.1 Varying cost of raw materials

The linear programming model is sometimes difficult to apply in real business situations due to its assumption of proportionality. A frequent phenomenon arising in practice is the varying price of the same resource. Namely, if a company needs additional quantities of raw materials it is possible to buy them from another supplier but at a different (usually higher) price. Let us assume that i raw material can be purchased at the price p_i, but only for the quantity lower (or equal) than Q. To purchase i raw material above that quantity it is necessary to take another supplier whose price is $p_i' > p_i$. Then the relation for the i raw material is transformed into:

$$a_{i1}\,x_1 + a_{i2}\,x_2 + \dots + a_{in}\,x_n = b_i + d_i, \tag{5}$$

with additional constraint $b_i \le Q$, where d_i is the additional quantity of the i raw material with the unit price p_i'.

Let us now consider such production situation when there are quantity discounts granted for bulk orders of raw materials. Therefore, in addition to the increasing cost effect we have to introduce this possibility into the model. Let us assume that for the k resource (b_k) the valid price is p_k as long as the purchased quantity is below Q, and the discounted price p_k' is valid for the entire quantity if the purchased quantity is higher (or equal) than Q. Consequently, the assumption is opposite to the one in the previous model, i.e. $p_k' < p_k$.

The previous formulation is not applicable since the optimization model will prefer using the less expensive material without satisfying the quota (Q). A different model has to be formulated with a slightly more complicated procedure.

Let

b_k, p_k – the amount and price of k raw material if it is purchased at less than the quantity discount volume;

d_k, p_k' – the amount and price of k raw material if it is purchased at the quantity discount.

The new model, in that case, instead of one equation for k raw material has some more relations, and those are:

$$a_{k1}\,x_1 + a_{k2}\,x_2 + \dots + a_{kn}\,x_n = b_k + d_k \tag{6}$$

$$b_k - Q^* \, y_1 \quad \leq 0 \tag{7}$$
$$d_k - Q \, y_2 \quad \geq 0 \tag{8}$$
$$d_k - M \, y_2 \quad \leq 0 \tag{9}$$

and, according to this, the budget constraints is:

$$p_1 \, b_1 + p_2 \, b_2 + \ldots + p_k \, b_k + p_k' \, d_k + \ldots + p_m \, b_m \leq B, \tag{10}$$

where M is a very large positive number ($M \gg 0$), or the upper limit for the procurement of the resource k, and Q^* is a number which is slightly lower than Q. Variables y_1 and y_2 are integer 0 - 1 variables, for which is:

$$y_1 + y_2 = 1 \tag{11}$$

In the above model there are two 0-1 variables y_1 and y_2, where only one of them always equals 1, and the other equals zero. Naturally, if the model comprises a number of resources that can be purchased at a discounted price then there are more 0-1 variables.

Since the same raw material has different price variable, the income from end product unit is not constant anymore. Therefore, maximizing the sum of $c_j \, x_j$, would not be an accurate measure of net income. Net income equation (1) should be recalculated as the difference between sales and total cost of materials, where the objective function will include materials at both prices. Consequently, if s_j is the sales price of j product, the objective function has the following form:

$$\text{Max } z = \sum_{j=1}^{n} s_j x_j - \sum_{i=1}^{m} p_i b_i - \sum_{k \in K} p_k' d_k \tag{12}$$

In that equation set K presents the indices of raw materials that have increasing or discounted prices, and d_k ($k \in K$) stays for those materials which in additional quantities can be bought only at a higher price (p_k'), or the quantities of raw materials if we bought them with quantity discounts.

In the budget equation it is also necessary to introduce costs for additional quantities of raw materials, so that it now takes the following form:

$$\sum_{i=1}^{m} p_i b_i + \sum_{k \in K} p_k' d_k \leq B \tag{13}$$

There is no need to specify that b_i should reach the maximum value of Q first, before allowing d_i greater than zero. The optimization model ensures b_i reaching the maximum value of Q because of the lower penalty, i.e. lower price p_i.

3 PROBLEM SETTING

This paper analyses the production planning problem in one bakery which produces twenty different products. These articles can be seen in Table 1. In this table there are the weights of the articles, lower and upper bounds for the week production and selling prices of all the articles.

Table 2. presents the list of raw materials that are used in production of these articles. There are 27 different raw materials and the purchasing prices for every of them are also presented in the table.

The amount of raw materials in one unit of articles (a_{ij}) are also used in production planning problem but their values are not presented in the paper because of short format of this paper.

Mark	Article name	Weight (in kg)	Weekly amount production		Selling prices s_j
			upper	lower	
A1	Rye mixed round	0,60	310	200	8,48
A2	Corn mixed	0,60	1300	1000	9,05
A3	Bread with sunflower seeds	0,50	320	180	9,43
A4	Wheat mixed semi-white	0,65	2560	1650	6,57
A5	Wheat half-white bread - folk	0,65	1500	1000	6,57
A6	Wheat white sandwich	0,65	7800	7100	8,38
A7	Rye mixed long	0,60	600	400	8,48
A8	Wheat mixed Sun	0,60	380	180	9,05
A9	Swedish bread	0,50	210	100	9,43
A10	Wheat mixed bread - Zagora	0,50	270	130	8,48
A11	White rolls mini salty	0,10	1500	980	3,86
A12	White rolls - milk roll	0,08	360	200	4,71
A13	Stuffed pastry layered cheese	0,09	400	270	5,4
A14	White rolls with salt	0,07	400	220	3,76
A15	White rolls round kaiser	0,06	1400	800	3,38
A16	White mini rolls	0,09	1550	900	4,05
A17	White pastry croissant	0,07	360	200	4,92
A18	Donut	0,07	3200	2000	4,20
A19	White pastry - trace	0,05	330	200	4,33
A20	Stuffed rolls - mixed corn and cheese	0,08	600	300	5,40

Table 2: List of raw materials

Mark	Raw material	Prices (per kg)	Mark	Raw material	Prices (per kg)
S1	Wheat flour T-850	2.255	S15	Margarin Tropic	10.45
S2	Wheat flour T-550	2.2	S16	Eggs (pieces)	0.77
S3	Rye flour	3.3	S17	Marmalade	9.02
S4	Wheat flour T-110	2.255	S18	Rum aroma	41.569
S5	Kitchen salt	1.54	S19	Goldperle plus TBM	33
S6	Yeast	5.775	S20	Vanilla sugar	14.069
S7	Aditiv panifarin	30.25	S21	Butter aroma	126.148
S8	Concentrate with wheat germ	16.346	S22	Rye liquid sourdough	13.057
S9	Grandma mix	21.989	S23	Corn concentrate Aurelia	10.89
S10	Suvita	17.259	S24	Cheese for bakery	15.95
S11	Sugar	5.83	S25	Enhancer	20.79
S12	Pure corn grits	7.7	S26	Wiener note	25.19
S13	Edible oil	7.7	S27	Grainpan Max	12.496
S14	Margarin BV	9.9			

According to these data the production planning problem can be posted as the linear programming model with one or more objective functions. Here we will consider the production problem in a concrete bakery where we have the varying price of the same resource, i.e. increasing costs of the same raw material or quantity discounts for some of raw materials.

Suppose that the available budget for weekly production is equal to the former costs of wanted quantity of resources procurement, i.e. $B = 19000$.

Since raw materials are now of different costs, variable prices of end product are not constant any more. Therefore, maximizing the sum of c_j x_j where c_j is the unit profit for article A_j, would not be an accurate measure of profit. Rather, profit equation should be recalculated as sales income less total cost of materials.

If x_j is the production quantity of i bakery product, the model which will take this increasing costs and quantity discounts into consideration is as follows:
Objective function (total contribution) which has to be maximized is:

$$\text{Max } z = \sum_{j=1}^{20} s_j x_j - \sum_{i=1}^{27} p_i b_i - \sum_{k \in K} p_k' d_k, \quad K = \{1, 2, 6, 23\}$$

In that equation set K presents the indices of raw materials that have increasing or discounted prices. In our case that happens for S_1, S_2, S_6 and S_{23}.

Let us consider such situation for our bakery production model. The first and second raw material (Wheat flour T-850 and Wheat flour T-550) can be purchased at a discounted price if the bought quantity is $Q_1 > 1200$ kg and $Q_2 > 4500$ kg, and this reduced price is valid for the entire quantity supplied, i.e. $p_1' = 2.0295$ and $p_2' = 1.98$.

In addition to this, let us assume increasing costs for yeast (S6) and corn concentrate Aurelia (S23) in this way:

The limit of yeast purchased at a lower price is 145 kg, while this limit in corn concentrate is 140 kg. The purchasing price of the additional quantity of yeast is $p_6' = 6.3525$ (10% more), and of corn concentrate $p_{23}' = 12.197$ (12% more) currency units.

Assuming that the budget level is B = 19000, and selling prices as in the Table 1 the constraints in the production model are:

Raw material constraints: $\sum_{j=1}^{20} a_{ij} x_j - b_i - d_i = 0, \quad i = 1, \cdots, 27$

where $d_i = 0$ except for $i \in K$, and a_{ij} are the amount of raw materials in one unit of specific article.

Constraints for the discounted prices for the first and second raw material:

$$b_1 - 1199\, y_1 \leq 0, \quad d_1 - 1200\, y_2 \geq 0, \quad d_1 - M\, y_2 \leq 0$$
$$b_2 - 4499\, y_3 \leq 0, \quad d_2 - 4500\, y_4 \geq 0, \quad d_2 - M\, y_4 \leq 0$$

where M is a very large positive number ($M \gg 0$), or the upper limit for the procurement of the specific resource k. Variables y_1, y_2, y_3 and y_4 are integer 0 - 1 variables, for which $y_1 + y_2 = 1$ and $y_3 + y_4 = 1$ is valid.

In the above model there are four 0-1 variables, where due to the upper relations only one of them in each pair always equals 1, and the other equals zero. Naturally, if the model comprises a number of resources that can be purchased at a discounted price then there are more 0-1 variables.

Last is the budget constraint: $\sum_{i=1}^{27} p_i b_i + \sum_{k \in K} p_k' d_k \leq 19000$

In addition to that the model has 21 integer variables (20 for units of articles - x_j) and one for the number of eggs (b_{16}).

Of course due to the data from Table 1 all articles have the lower and upper bounds which is 40 more constraints.

There are two more constraints for the raw materials which have increasing costs:

$$b_6 \leq 145, \quad b_{23} \leq 140.$$

It should be remarked that it is the mixed integer programming problem with twenty one integer variables (x_j and b_{16}), thirty continuous variables (26 b_i and 4 d_i), and four binary variables (y_j). Its optimal solution is obtained by MATLAB and is presented in the Table 3.

Optimal value of the objective function is: $z^* = 124076.50$ and for that production the available budget is completely spent.

In this table we can see the required quantities of raw materials. For first and second raw materials (Wheat flour T-850 and Wheat flour T-550) we have the quantity discounts because

we purchase these two types of flour over the limited quantity (Q_1 = 1200 kg and Q_2 = 4500 kg). So we get these two types of flour at the discount prices. In our model, of course, the binary variables y_2 and y_4 are equal 1.

Table 3: Optimal solution

Variables	Optimal solution	Variables	Optimal solution	Variables	Optimal solution		Variables	Optimal solution
X1	300	X15	1400	**d6**	**15.026**		b20	2.336
X2	1000	X16	1550	b7	6.313		b21	0.330
X3	180	X17	360	b8	4.472		b22	4.284
X4	1650	X18	3200	b9	3.420		**b23**	**132.420**
X5	1008	X19	330	b10	10.260		**d23**	**0**
X6	7401	X20	300	b11	15.715		b24	15.300
X7	600	**b1**	**0**	b12	8.957		b25	30.499
X8	380	**d1**	**1200.01**	b13	13.101		b26	1.123
X9	100	**b2**	**0**	b14	14.503		b27	14.301
X10	130	**d2**	**4758.15**	b15	22.698		y1	0
X11	1500	b3	141.593	b16	727		y2	1
X12	360	b4	240.570	b17	32.000		y3	0
X13	270	b5	126.492	b18	2.448		y4	1
X14	400	**b6**	**145.000**	b19	7.264		**z* = 124076.50**	

The sixth raw material (Yeast) have to be purchased over the limited quantity and so the quantity over the limit (Q_6 = 145 kg) is purchased at a higher price. For the S23 (Corn concentrate Aurelia) we didn't cross the limit (Q_{23} = 140 kg) and whole quantity (b23) will be purchased at a lower price.

4 CONCLUSION

In this paper, De Novo programming model in the production plan optimization in a bakery is considered. The efficiency of the proposed model is investigated on the case of a company that produces various bakery products. De Novo presents a special approach to optimization. Instead of "optimizing a given system", it suggests a way of "designing an optimal system". De Novo approach does not limit the resources as most of the necessary resource quantities can be obtained at certain prices. Resources, of course, are actually limited because their maximum quantity is controlled by the budget, which is an important element of De Novo.

The obtained results indicate a high application efficiency of the proposed model by using De Novo programming in solving the production plan optimization problem in various production companies. Using De Novo approach most varied cases can be handled more effectively than by the standard programming models and in this paper increasing costs of raw materials and quantity discounts for some raw materials in bakery production are investigated.

The future work on this issue will investigate the possibilities of introducing new objective functions in bakery production and solving this production problem as the multi-criteria ones [3], [4].

References
[1] Zeleny, M. 1986. *Optimal System Design with Multiple Criteria: De Novo Programming Approach*. Engineering Costs and Production Economics, No. 10: 89-94.
[2] Babić, Z. 2009. *De Novo Programming – A Bluff or not a Bluff*. Proceedings of the 10th International Symposium on Operational Research, Nova Gorica, Slovenia. 3-12.
[3] Babić, Z., I. Pavić. 1996. *Multicriterial Production Planning by De Novo Programming Approach*. International Journal of Production Economics, 43(1): 59-66.
[4] Chakraborty, S., D. Bhattacharya. 2013. *Optimal System Design under Multi-Objective Decision Making Using De Novo Concept: A New Approach*. International Journal of Computer Applications, 63(12): 20-27.

THE EFFECTS OF TRANSFORMATIONS ON THE OPTIMAL SET IN INTERVAL LINEAR PROGRAMMING

Elif Garajová, Milan Hladík

Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics

Malostranské nám. 25, 11800 Prague, Czech Republic

{elif, hladik}@kam.mff.cuni.cz

Miroslav Rada

University of Economics, Prague, Department of Financial Accounting and Auditing,

nám. W. Churchilla 4, 13067 Prague, Czech Republic

miroslav.rada@vse.cz

Abstract: Interval linear programming provides a mathematical tool for handling linear optimization problems affected by uncertainty. Contrarily to classical linear programming, the properties of interval linear programs depend on the form in which the program is given. In this paper, we study how the transformations used in linear programming – imposing non-negativity, changing equations to inequalities and vice versa – affect the optimal set of an interval program.

Since the results are mostly negative in the general case, we also consider interval linear programs with a fixed coefficient matrix. For this special case, we prove that all transformations preserve the optimal solution set.

Keywords: interval linear programming, optimal set, transformations

1   INTRODUCTION

Many real-world optimization problems are affected by uncertainty in the input data due to approximations or measurement errors. There have been various approaches incorporating uncertainty in the optimization models in order to achieve a realistic representation of the problems, such as robust optimization, stochastic programming or parametric programming. In this paper, we consider linear programming with interval-valued uncertainty: we assume that lower and upper bounds enclosing the exact data are given.

Let us first introduce some terminology and notation. Given $\underline{A}, \overline{A} \in \mathbb{R}^{m \times n}$ with $\underline{A} \leq \overline{A}$, an *interval matrix* $\boldsymbol{A}$ is defined as the set of matrices

$$\boldsymbol{A} = [\underline{A}, \overline{A}] = \{A \in \mathbb{R}^{m \times n} : \underline{A} \leq A \leq \overline{A}\},$$

where the inequalities are understood entry-wise. The matrices $\underline{A}, \overline{A}$ are the *lower* and *upper bound* of $\boldsymbol{A}$, respectively. Alternatively, an interval matrix can be defined by its *center* A_c and *radius* A_Δ:

$$A_c = \frac{1}{2}(\overline{A} + \underline{A}), \qquad A_\Delta = \frac{1}{2}(\overline{A} - \underline{A}).$$

Throughout the paper, we denote interval objects by bold lowercase (for one-dimensional intervals and interval vectors) and uppercase (for interval matrices) letters. The symbol $\mathbb{IR}$ is used to denote the set of all (real) intervals.

An *interval linear program* (ILP) is then defined as a family of linear programs

$$\text{minimize } c^T x \text{ subject to } x \in \mathcal{M}(A, b), \tag{1}$$

where $A \in \boldsymbol{A}, b \in \boldsymbol{b}, c \in \boldsymbol{c}$ are the given interval data and $\mathcal{M}(A, b)$ denotes the feasible set. For the sake of brevity, we also write ILP (1) shortly as

$$\text{minimize } \boldsymbol{c}^T x \text{ subject to } x \in \mathcal{M}(\boldsymbol{A}, \boldsymbol{b}). \tag{2}$$

A particular linear program in the family is called a *scenario*. A vector $x \in \mathbb{R}^n$ is a *(weakly) feasible/optimal* solution to the interval linear program, if there exists a scenario $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}, c \in \boldsymbol{c}$ such that x is a feasible/optimal solution for the scenario (see [7] and [3] for overview of the feasibility properties of interval linear systems in different forms).

Usually, we assume that an ILP is given in one of the following three forms:

$$\text{minimize } \boldsymbol{c}^T x \text{ subject to } \boldsymbol{A}x = \boldsymbol{b}, \ x \geq 0, \tag{I}$$

$$\text{minimize } \boldsymbol{c}^T x \text{ subject to } \boldsymbol{A}x \leq \boldsymbol{b}, \tag{II}$$

$$\text{minimize } \boldsymbol{c}^T x \text{ subject to } \boldsymbol{A}x \leq \boldsymbol{b}, \ x \geq 0. \tag{III}$$

In classical linear programming, we can equivalently express a program in any of these forms. However, this is not always possible for interval linear programs due to the so-called dependency problem: even when an interval coefficient appears multiple times in a program, we assume that the real coefficients are chosen from the corresponding intervals independently.

2 TRANSFORMATIONS IN INTERVAL LINEAR PROGRAMMING

When working with linear programs, we often use transformations between equations and inequalities or impose non-negativity on the variables in order to convert a program into the desired form. We will now study how these conversions affect the feasible and optimal solution set, when applied to an interval linear program.

It is easy to see that adding slack variables to an inequality-constrained ILP does not change the set of optimal or feasible solutions. Thus, we can transform an interval inequality system $\boldsymbol{A}x \leq \boldsymbol{b}$ into an equivalent interval system

$$\boldsymbol{A}x + Iy = \boldsymbol{b}, \ y \geq 0.$$

However, applying the other transformations may result in a relaxation of the original problem, since multiple occurrences of the interval coefficients introduced by the transformations are independent.

For example, consider the interval program

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & [0,1]x_1 - x_2 = 0, \\ & x_2 \leq 1, \\ & x_1, x_2 \geq 0. \end{array} \tag{3}$$

The optimal solution set of (3) consists of all values with $x_1 \in [1, \infty)$ and $x_2 = 1$. By splitting the equation $[0,1]x_1 - x_2 = 0$ into two inequalities, we introduce additional scenarios into the problem, such as the scenario

$$1x_1 - x_2 \leq 0, \quad 0x_1 - x_2 \geq 0$$

with the optimal solution $(0,0)$. Therefore, the optimal solution sets of the two problems are not equivalent. Even though the optimal sets are different, the feasible set does not change with this transformation, as proved by Li [5].

The remaining transformation, which introduces the substitution $x = x^+ - x^-$ with non-negative variables x^+, x^- instead of a free variable x, does not preserve the feasible (or optimal) set of an interval linear program, in general. This was already shown in a survey by Hladík [2]. Therefore, the only transformations among the three considered forms of ILPs preserving the optimal solution set are from type (III) to (I) by adding slack variables and the trivial transformation to type (II) (see Figure 1 for an overview).
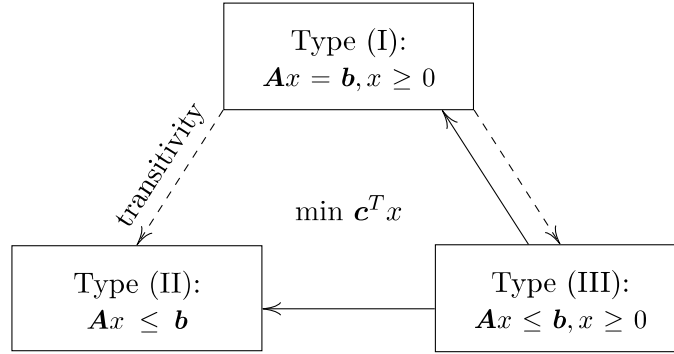
Figure 1: Transformations preserving the feasible (dashed arrows) and optimal (solid arrows) solution set of a general interval linear program.

3   SPECIAL CASE: FIXED COEFFICIENT MATRIX

Let us now consider an important special class of interval programs, in which the uncertainty only affect the objective function and the right-hand-side vector, i.e. problems with a fixed (real) coefficient matrix. This type of interval programs arises in various applications, such as transportation problems [4] or minimum-cost flow problems [1]. We show that unlike general interval linear programs, this type of programs allows for all types of transformations while preserving the optimal solution set.

First, we prove that an ILP of type (I) with a real coefficient matrix can be transformed into a program of type (III) with the same set of optimal solutions, i.e. that splitting an equation into inequalities does not affect the optimal set in the considered special case.

Theorem 3.1 *Let $\mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$ be the optimal solution set of an interval linear program of type* (I) *given by the triplet $(A, \boldsymbol{b}, \boldsymbol{c})$. Then, $\mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$ is equal to the optimal solution set of the problem*

$$
\begin{aligned}
\text{minimize} \quad & \boldsymbol{c}^T x \\
\text{subject to} \quad & Ax \leq \boldsymbol{b}_1, \\
& -Ax \leq -\boldsymbol{b}_2, \\
& x \geq 0.
\end{aligned}
\tag{4}
$$

with $\boldsymbol{b}_1 = \boldsymbol{b}_2 = \boldsymbol{b}$.

Proof. Let $\mathcal{S}(A', \boldsymbol{b}', \boldsymbol{c})$ denote the optimal solution set of program (4). Clearly, the inclusion $\mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c}) \subseteq \mathcal{S}(A', \boldsymbol{b}', \boldsymbol{c})$ holds, since (4) contains all the scenarios of (I).

On the other hand, let x' be an optimal solution of ILP (4) for a scenario determined by the coefficient vectors $c \in \boldsymbol{c}$ and $b_1, b_2 \in \boldsymbol{b}$ satisfying

$$
Ax' \leq b_1, \ -Ax' \leq -b_2.
\tag{5}
$$

Since we have, $b_2 \leq Ax' \leq b_1$, there exists $b_3 \in [b_2, b_1] \subseteq \boldsymbol{b}$ with $Ax' = b_3$. We claim that x' is also an optimal solution of program (I) for the scenario

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & Ax = b_3, \ x \geq 0.
\end{aligned}
$$

Suppose, for the sake of contradiction, that there exists another feasible solution x^* with $c^T x^* < c^T x'$. By the choice of b_3, the vector x^* is also feasible for scenario (5). However, since the objective function is the same for both problems, this contradicts the assumption that x' is optimal in scenario (5). □

Note that even though the optimal sets of both problems are equivalent, the transformation may generate new infeasible linear programs in the ILP due to the dependency problem. This also affects the range of optimal values [6], if we allow infinite bounds for ILPs containing unbounded and infeasible programs. Thus, the two programs are not entirely equivalent with respect to other properties.

We can also prove a similar result for the transformation of an ILP of type (II) into type (III) by splitting a variable into two non-negative variables. Recall that in the general case, this transformation does not even preserve the set of feasible solutions.

Theorem 3.2 *Let $\mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$ denote the optimal solution set of an interval linear program of type* (II) *and let $\mathcal{S}(A', \boldsymbol{b}, \boldsymbol{c}')$ be the optimal solution set of the program*

$$
\begin{array}{ll}
\text{minimize} & \boldsymbol{c}_1^T x^+ - \boldsymbol{c}_2^T x^- \\
\text{subject to} & Ax^+ - Ax^- \le \boldsymbol{b}, \\
& x^+, x^- \ge 0
\end{array}
\tag{6}
$$

with $\boldsymbol{c}_1 = \boldsymbol{c}_2 = \boldsymbol{c}$. Then, the following properties hold:

1. *If $x \in \mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$, then there exists $(x^+, x^-) \in \mathcal{S}(A', \boldsymbol{b}, \boldsymbol{c}')$ with $x = x^+ - x^-$.*

2. *Conversely, if $(x^+, x^-) \in \mathcal{S}(A', \boldsymbol{b}, \boldsymbol{c}')$, then $x^+ - x^- \in \mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$.*

Proof. Let $x \in \mathcal{S}(A, \boldsymbol{b}, \boldsymbol{c})$ be optimal for a scenario $c \in \boldsymbol{c}$, $b \in \boldsymbol{b}$. Decompose x into the non-negative vectors $x^+ = \max(0, x)$ and $x^- = -\min(0, x)$, where the operations max and min are understood entry-wise. The vectors satisfy $x = x^+ - x^-$ and it is easy to see that the pair (x^+, x^-) is optimal in ILP (6) for the scenario determined by the objective vector $(c, -c)$ and the right-hand side b.

Now, let $(x^+, x^-) \in \mathcal{S}(A', \boldsymbol{b}, \boldsymbol{c}')$ be optimal for some $b \in \boldsymbol{b}$ and $c_1, c_2 \in \boldsymbol{c}$. By duality in linear programming, there exists a dual feasible vector y satisfying

$$
\begin{aligned}
& c_1^T x^+ - c_2^T x^- = b^T y, \\
& Ax^+ - Ax^- \le b, \ x^+ \ge 0, \ x^- \ge 0, \\
& A^T y \le c_1, \ -A^T y \le -c_2, \ y \le 0.
\end{aligned}
$$

From the dual feasibility constraints, we have $A^T y = c_3$ for some $c_3 \in [c_2, c_1] \subseteq \boldsymbol{c}$. We will prove that $x = x^+ - x^-$ is optimal for the scenario with objective vector c_3 and right-hand-side vector b in ILP (II), by showing that it satisfies the system

$$
c_3^T x = b^T y, \ Ax \le b, \ A^T y = c_3, \ y \le 0.
$$

Namely, it remains to show that $c_3^T(x^+ - x^-) = b^T y$ holds. Complementary slackness implies that for each $i \in \{1, \dots, n\}$ we have

$$
\begin{aligned}
& (x^+)_i = 0 \ \vee \ (c_1 - A^T y)_i = 0, \ \text{and} \\
& (x^-)_i = 0 \ \vee \ (A^T y - c_2)_i = 0.
\end{aligned}
$$

If $(x^+)_i > 0$ and $(x^-)_i > 0$ for some index i, then we have $(c_1)_i = (c_2)_i = (c_3)_i$ by the choice of c_3. For $(x^+)_i = 0$ and $(x^-)_i > 0$ we have $(c_2)_i = (c_3)_i$, and analogically, $(c_1)_i = (c_3)_i$ for the symmetric case. By substituting into the equation

$$
c_1^T x^+ - c_2^T x^- = b^T y,
$$

we can see that the desired constraint is satisfied in all cases. Therefore, $x^+ - x^-$ is an optimal solution of program (II). $\qquad \square$

By transitivity, Theorems 3.1 and 3.2 imply that all transformations among interval linear programs of the three considered types preserve the set of optimal solutions (see Figure 2).
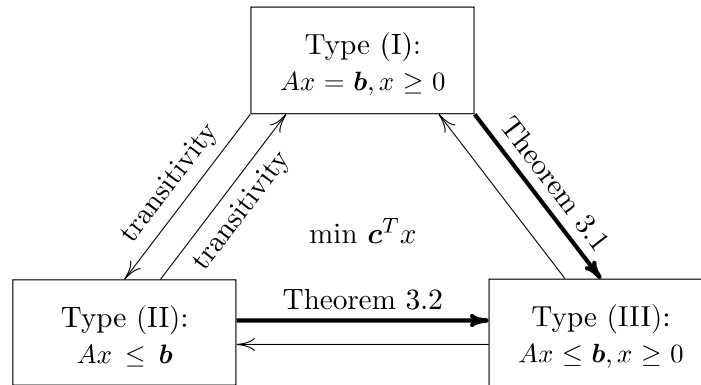
Figure 2: Transformations preserving the optimal solution set of an interval linear program with a fixed coefficient matrix.

4 CONCLUSION

We studied the effects of the basic transformations often used in classical linear programming on the set of all optimal solutions of an interval linear program. We showed by a simple counterexample, that while splitting an equation into two inequalities preserves the feasible set, in general, it may change the optimal solution set. The only transformation that produces an equivalent interval program (with respect to the optimal set) is the conversion of inequalities into equations by adding slack variables.

Stronger results were obtained for a special class of interval programs with a fixed coefficient matrix. We proved that on this class of problems, all considered transformations preserve both the feasible and the optimal solution set. This shows that various results concerning the optimal set are not only relevant to a specific type of programs, but can be generalized to programs in other forms as well. However, even in this case, the programs are not entirely equivalent and may differ in other properties.

5 Acknowledgements

References

[1] Hashemi, S. M., Ghatee, M. and Nasrabadi, E. (2006). Combinatorial algorithms for the minimum interval cost flow problem. Applied Mathematics and Computation, 175(2): 1200–1216.

[2] Hladík, M. (2012). Interval linear programming: A survey. In Mann, Z. A. (Ed.) Linear Programming – New Frontiers in Theory and Applications, chapter 2, (pp. 85–120). New York: Nova Science Publishers.

[3] Hladík, M. (2017). Transformations of interval linear systems of equations and inequalities. Linear Multilinear Algebra, 65(2): 211–223.

[4] Jiménez, F. and Verdegay, J. (1998). Uncertain solid transportation problems. Fuzzy Sets and Systems, 100(1): 45–57.

[5] Li, W. (2015). A note on dependency between interval linear systems. Optimization Letters, 9(4): 795–797.

[6] Rohn, J. (2006). Interval linear programming. In Linear Optimization Problems with Inexact Data, (pp. 79–100). Springer US.

[7] Rohn, J. (2006). Solvability of systems of interval linear equations and inequalities. In Linear Optimization Problems with Inexact Data, (pp. 35–77). Springer US.

GLOBAL OPTIMIZATION FOR SUM-OF-RATIOS PROBLEM USING D.C. PROGRAMMING

Tatiana V. Gruzdeva and Alexander S. Strekalovsky
Matrosov Institute for System Dynamics & Control Theory SB RAS,
Lermontov st., 134, Irkutsk, 664033, Russia

Abstract: This paper addresses a rather general fractional optimization problem. There are two ways to reduce the original problem. The first one is to solve the auxiliary problem with nonlinear nonconvex inequality constraints. The second one is a solution of an equation with the optimal value of an auxiliary nonconvex problem with a vector parameter. Both auxiliary problems turn out to be d.c. optimization problems, which allows to apply Global Optimization Theory and develop two corresponding global search algorithms that have been tested.

Keywords: fractional programming, nonconvex optimization, difference of two convex functions, equation with a vector parameter, global search algorithm.

1 INTRODUCTION

We consider the following problem of the fractional optimization [1, 8]

$$(\mathcal{P}_f) \qquad f(x) := \sum_{i=1}^{m} \frac{\psi_i(x)}{\varphi_i(x)} \downarrow \min_x, \ \ x \in S,$$

where $S \subset I\!\!R^n$ is a convex set and $\psi_i : I\!\!R^n \to I\!\!R, \ \ \varphi_i : I\!\!R^n \to I\!\!R,$

$$(\mathcal{H}_0) \qquad \psi_i(x) > 0, \ \varphi_i(x) > 0 \ \ \forall x \in S, \ \ i = 1, \ldots, m.$$

The fractional programming problems arise in various economic applications and real-life problems [8]. However, it is well-known that the sum-of-ratios program is NP-complete [3]. Surveys on methods for solving this problem can be found in [1, 8], but the development of new efficient methods for a fractional program still remains an important field of research in mathematical optimization.

In the recent two decades, we have succeeded in developing the Global Search Theory, which perfectly fits optimization theory and proved to be rather efficient in terms of computations [4, 9, 10]. Now we are going to apply this theory to solving fractional programming problems.

We propose reduction of the sum-of-ratios problem to the problem of minimizing a linear function on the nonconvex feasible set given by d.c. functions. In this case, we need to solve the following nonconvex problems

$$(\mathcal{DCC}) \quad f_0(x) \downarrow \min_x, \ x \in S, \ \ f_i(x) = g_i(x) - h_i(x) \le 0, \ \ i = 1, \ldots n,$$

where $g_i(\cdot), \ h_i(\cdot) \ i = 1, \ldots n,$ are convex functions, $S \subset I\!\!R^n$, $f_0(\cdot)$ is a continuous function.

Next, generalizing the Dinkelbach's idea [2], we propose to attack such problems with the help of reduction of the fractional program $(\mathcal{P}_f)$ to solving an equation with the optimal value function of a d.c. minimization problem and the vector parameter that satisfies the nonnegativity assumption. To this end, we have to use the solution to the following d.c. minimization problem, treated here as an auxiliary one:

$$(\mathcal{DC}) \qquad f(x) = g(x) - h(x) \downarrow \min, \ \ x \in D,$$

where $g(\cdot), \ h(\cdot)$ are convex functions, and D is a convex set, $D \subset I\!\!R^n$.

Thus, based on the solution of these two classes of d.c. optimization problems we developed two new methods for solving a general fractional program.

The outline of the paper is as follows. In Sect. 2, instead of considering a fractional program directly, we substantiate the reduction of the sum-of-ratios fractional problem to the optimization problem with nonconvex constraints ($\mathcal{DCC}$). In Sect. 3, we propose to combine solving of the corresponding d.c. minimization problem ($\mathcal{DC}$) with a search with respect to the vector parameter. Finally, in Sect. 4, we show some comparative computational testing of two approaches on fractional program instances.

2 REDUCTION TO THE PROBLEM WITH D.C. CONSTRAINTS

In this section we reduce the fractional program to the optimization problem with a nonconvex feasible set.

Proposition 2.1 *[6] Let the pair $(x_*, \alpha_*) \in I\!R^n \times I\!R^m$ be a solution to the following problem:*

$$\sum_{i=1}^{m} \alpha_i \downarrow \min_{(x,\alpha)}, \quad x \in S, \quad \frac{\psi_i(x)}{\varphi_i(x)} \le \alpha_i, \quad i = 1, \dots, m. \qquad (1)$$

Then $\quad \dfrac{\psi_i(x_*)}{\varphi_i(x_*)} = \alpha_{*i}, \quad i = 1, \dots, m.$

Corollary 2.2. *For any solution $(x_*, \alpha_*) \in I\!R^n \times I\!R^m$ to the problem (1), the point x_* will be a solution to Problem $(\mathcal{P}_f)$.*

The inequality constraints in the problem (1) can be replaced by the equivalent constraints $\psi_i(x) - \alpha_i \varphi_i(x) \le 0, \; i = 1, \dots, m$, since $\varphi_i(x) > 0 \; \forall x \in S$. This yields the following problem with m nonconvex constraints:

$$(\mathcal{P}) \qquad f_0 := \sum_{i=1}^{m} \alpha_i \downarrow \min_{(x,\alpha)}, \quad x \in S, \quad f_i := \psi_i(x) - \alpha_i \varphi_i(x) \le 0, \quad i = 1, \dots, m.$$

We intend to solve this problem using the exact penalization approach for d.c. optimization developed in [11]. Therefore, we introduce the penalized problem

$$(\mathcal{P}_\sigma) \qquad \theta_\sigma(x) = f_0(x) + \sigma \max\{0, f_i(x), i \in I\} \downarrow \min, \quad x \in S.$$

It can be readily seen that the penalized function $\theta_\sigma(\cdot)$ is a d.c. function, because the functions $f_i(\cdot) = g_i(\cdot) - h_i(\cdot), \; i \in I \cup \{0\}$, are the same, and we can apply the global search theory to solving this class of nonconvex problems [9, 10].

The theory enables us to construct an algorithm which consists of two principal stages:
a) local search, which provides an approximately critical point [12];
b) procedures of escaping from critical points.

Let the Lagrange multipliers, associated with the constraints and corresponding to the point $z^k, \; k \in \{1, 2, \dots\}$, be denoted by $\lambda := (\lambda_1, \dots, \lambda_m) \in I\!R^m$.

Global search scheme

Step 1. Using the local search method from [12], find a critical point z^k in $(\mathcal{P})$.

Step 2. Set $\sigma_k := \sum_{i=1}^{m} \lambda_i$. Choose a number $\beta : \inf(G_\sigma, S) \le \beta \le \sup(G_\sigma, S)$.

        Choose an initial $\beta_0 = G_\sigma(z^k), \; \zeta_k = \theta_\sigma(z^k)$.

Step 3. Construct a finite approximation

$$R_k(\beta) = \{v^1, \dots, v^{N_k} \mid H_\sigma(v^i) = \beta + \zeta_k, \; i = 1, \dots, N_k, \; N_k = N_k(\beta)\}.$$

Step 4. Find a δ_k-solution $\bar{u}^i$ of the following Linearized Problem:

$$(P_\sigma L_i) \qquad\qquad G_\sigma(x) - \langle \nabla H_\sigma(v^i), x \rangle \downarrow \min_x, \; x \in S,$$

        so that $G_\sigma(\bar{u}^i) - \langle \nabla H_\sigma(v^i), \bar{u}^i \rangle - \delta_k \le \inf_x \{G_\sigma(x) - \langle \nabla H_\sigma(v^i), x \rangle\}.$

Step 5. Starting from the point $\bar{u}^i$, find a KKT-point u^i by the local search method from [12].

Step 6. Choose the point $u^j : \quad f_0(u^j) \le \min\{f_0(u^i), \ i = 1, ..., N\}$.

Step 7. If $f_0(u^j) < f_0(z^k)$, then set $z_{k+1} = u^j, \quad k = k+1$ and go to Step 2.

Step 8. Otherwise, choose a new value of β and go to Step 3.

According to Corollary 2.2, the point z^* resulting from the global search strategy will be a solution to the original fraction program. It should be noted that, in contrast to the approach from next Sect. 3, α_i will be found simultaneously with the solution vector x, because they are variables, although auxiliary ones, of the optimization problem.

3   REDUCTION TO THE D.C. MINIMIZATION PROBLEM

The second method adopts reduction of the fractional minimization problem to the solution of an equation with an optimal value of the d.c. optimization problem with a vector parameter.

Let us consider the following parametric optimization problem

$$(\mathcal{P}_\alpha) \qquad \Phi(x,\alpha) := \sum_{i=1}^m [\psi_i(x) - \alpha_i\varphi_i(x)] \downarrow \min_x, \quad x \in S,$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)^\top \in I\!\!R^m$ is the vector parameter.

Further, let us introduce the function $\mathcal{V}(\alpha)$ of the optimal value to Problem $(\mathcal{P}_\alpha)$:

$$\mathcal{V}(\alpha) := \inf_x \{\Phi(x,\alpha) \mid x \in S\} = \inf_x \left\{ \sum_{i=1}^m [\psi_i(x) - \alpha_i\varphi_i(x)] \ : \ x \in S \right\}.$$

In addition, suppose that the following assumptions are fulfilled:

$$(\mathcal{H}_1) \qquad \begin{array}{l} (a) \ \ \mathcal{V}(\alpha) > -\infty \ \ \forall \alpha \in \mathcal{K}, \text{where } \mathcal{K} \text{ is a convex set from } I\!\!R^m; \\ (b) \ \ \forall \alpha \in \mathcal{K} \subset I\!\!R^m \ \ \text{there exists a solution } z = z(\alpha) \text{ to Problem } (\mathcal{P}_\alpha). \end{array}$$

In what follows, we say that the data of Problem $(\mathcal{P}_f)$ satisfy "the nonnegativity condition", if the following inequalities hold

$$(\mathcal{H}(\alpha)) \qquad \psi_i(x) - \alpha_i\varphi_i(x) \ge 0 \ \ \forall x \in S, \ \ i = 1, \ldots, m.$$

Theorem 3.1 *[5] Suppose that in Problem $(\mathcal{P}_f)$ $\psi_i(x) > 0$, $\varphi_i(x) > 0$ and the assumption $(\mathcal{H}_1)$ is satisfied. In addition, let there exist a vector $\alpha_0 = (\alpha_{01}, \ldots, \alpha_{0m})^\top \in \mathcal{K} \subset I\!\!R^m$ at which "the nonnegativity condition" $(\mathcal{H}(\alpha_0))$ holds. Besides, suppose that in Problem $(\mathcal{P}_{\alpha_0})$ the following equality takes place:*

$$\mathcal{V}(\alpha_0) : = \inf_x \left\{ \sum_{i=1}^m [\psi_i(x) - \alpha_{0i}\varphi_i(x)] \ : \ x \in S \right\} = 0. \qquad (2)$$

Then, any solution $z = z(\alpha_0)$ to Problem $(\mathcal{P}_{\alpha_0})$ is a solution to Problem $(\mathcal{P}_f)$, so that $z \in Sol(\mathcal{P}_{\alpha_0}) \subset Sol(\mathcal{P}_f)$.

According to Theorem 3.1, in order to verify the equality (2), we should be able to find a global solution to Problem $(\mathcal{P}_\alpha)$ for every $\alpha \in I\!\!R^m_+$. Since $\psi_i(\cdot)$, $\varphi_i(\cdot)$, $i = 1, \ldots, m$, are d.c. functions it can be readily seen that Problem $(\mathcal{P}_\alpha)$ belongs to the class of d.c. minimization. As a consequence, in order to solve Problem $(\mathcal{P}_\alpha)$, we can apply the Global Search Theory [9, 10].

Hence, instead of solving Problem $(\mathcal{P}_f)$, we propose to combine solving of Problem $(\mathcal{P}_\alpha)$ with a search with respect to the parameter $\alpha \in I\!\!R_+^m$ in order to find the vector $\alpha_0 \in I\!\!R_+^m$ such that $\mathcal{V}(\alpha_0) = 0$.

Denote $\Phi_i(x) := \psi_i(x) - \alpha_i^k \varphi_i(x)$, $i = 1, \ldots, m$. Let $[0, \alpha^+]$ be a segment for varying α. To choose α^+ we should take into account that due to $(\mathcal{H}(\alpha))$ and $(\mathcal{H}_0)$, we have $\forall i = 1, \ldots, m$:

$$\alpha_i \le f_i(x) \triangleq \frac{\psi_i(x)}{\phi_i(x)} \le \sum_{i=1}^m \frac{\psi_i(x)}{\phi_i(x)} = f(x) \quad \forall x \in S, \text{ so, for example, } \alpha^+ \text{ can be chosen as}$$

$\alpha_i^+ = f_i(x^0)$, $i = 1, \ldots, m$.

Method for solving the equation $\mathcal{V}(\alpha) = 0$

Step 0. (Initialization) $k = 0$, $v^k = 0$, $u^k = \alpha^+$, $\alpha^k = \frac{\alpha^+}{2} \in [v^k, u^k]$.

Step 1. Find a solution $z(\alpha^k)$ to Problem $(\mathcal{P}_{\alpha^k})$ using the global search strategy for d.c. minimization [9, 10].

Step 2. (Stopping criterion) If $\mathcal{V}_k := \mathcal{V}(\alpha^k) = 0$ and $\min_i \Phi_i(z(\alpha^k)) \ge 0$, then STOP: $z(\alpha^k) \in Sol(\mathcal{P}_f)$ (in virtue of Theorem 3.1).

Step 3. If $\mathcal{V}_k > 0$, then set $v^{k+1} = \alpha^k$, $\alpha^{k+1} = \frac{1}{2}(u^k + \alpha^k)$, $k = k + 1$ and go to Step 1.

Step 4. If $\mathcal{V}_k < 0$, then set $u^{k+1} = \alpha^k$, $\alpha^{k+1} = \frac{1}{2}(v^k + \alpha^k)$, $k = k + 1$ and go to Step 1.

Step 5. If $\min_i \Phi_i(z(\alpha^k)) < 0$, then set $\alpha_i^{k+1} = \frac{\psi_i(z(\alpha^k))}{\varphi_i(z(\alpha^k))}$ $\forall i : \Phi_i(z(\alpha^k)) < 0$, $\alpha_i^{k+1} = \alpha_i^k$ $\forall i : \Phi_i(z(\alpha^k)) \ge 0$. In addition, set $v^{k+1} = 0$, $u^{k+1} = t_{k+1}\alpha^{k+1}$, where $t_{k+1} = \dfrac{\alpha^+}{\max\limits_i \alpha_i}$, $k = k + 1$, and return to Step 1.

Let us emphasize the fact that the algorithm for solving Problem $(\mathcal{P}_f)$ of fractional optimization consists of 3 basic stages: the (a) local and (b) global searches in Problem $(\mathcal{P}_\alpha)$ with a fixed vector parameter α and (c) the method for finding the vector parameter α at which the optimal value of Problem $(\mathcal{P}_\alpha)$ is zero, i.e. $\mathcal{V}(\alpha) = 0$.

4   COMPUTATIONAL SIMULATIONS

Two approaches from above for solving the fractional programs $(\mathcal{P}_f)$ via d.c. optimization problems were verified. The algorithm based on the global search scheme from Sect. 2 (F1-algorithm) and the algorithm based on the method for solving the equation $\mathcal{V}(\alpha) = 0$ from Sect. 3 (F2-algorithm) were applied to an extended set of test examples generated randomly. All computational experiments were performed on the Intel Core i7-4790K CPU 4.0 GHz. All convex auxiliary problems (linearized problems) on the steps of F1-, F2-algorithms were solved by the software package IBM ILOG CPLEX 12.6.2.

First, we randomly generated "linear/linear" problems of the form

$$f(x) := \sum_{i=1}^m \frac{\langle a^i, x \rangle + b_i}{\langle c^i, x \rangle + d_i} \downarrow \min_x, \quad Qx \le q, \quad x \ge 0. \tag{3}$$

Table 1 presents results of some comparative computational testing of two approaches (F1-, F2-algorithms) on fractional problems (3) with up to 100 variables and 100 terms in the sum. Tables 1 and 2 below employ the following designations: n is the number of variables (problem's dimension); m is the number of terms in the sum; $f(x^0)$ is the value of the goal function to Problem $(\mathcal{P}_f)$ at the starting point; $f(z)$ is the value of the function at the solution provided by the algorithms; $TimeF1$ and $TimeF2$ stand for the CPU time of computing (seconds) by F1-, F2-algorithms, respectively.

Table 1: Randomly generated problems (3) with affine functions

n	m	$f(x_0)$	$f(z)$	$TimeF1$	$TimeF2$
10	10	10.000000	7.289142	2.19	1.02
10	50	50.000000	48.267163	7.77	5.18
10	100	100.000000	95.932117	66.93	13.66
50	10	10.000000	7.828726	7.16	2.26
50	50	50.000000	45.611627	60.91	19.65
50	100	100.000000	94.173538	420.13	28.58
100	10	10.000000	7.530948	102.15	9.33
100	50	50.000000	44.635263	740.66	60.71
100	100	100.000000	91.738276	1039.28	102.32

Moreover, we tested the proposed algorithms on fractional problems with quadratic functions in the numerators of ratios. We generated the problems from [7] as follows:

$$f(x) := \sum_{i=1}^{m} \frac{\langle x, A_i x \rangle + \langle b^i, x \rangle}{\langle c^i, x \rangle} \downarrow \min_x, \quad Qx \leq q, \quad x \in [1, 5]^n, \tag{4}$$

where $A_i = U_i D_{0i} U_i^T$, $U_i = V_1 V_2 V_3$, $i = 1, \ldots, n$; $V_j = I - 2\frac{w_j w_j^T}{\|w_j\|^2}$, $j = 1, 2, 3$; $w_1 = -i + rand(n, 1)$, $w_2 = -2i + rand(n, 1)$, $w_3 = -3i + rand(n, 1)$; $c_i = i - i \cdot rand(n, 1)$, $b_i = i + i \cdot rand(n, 1)$, $Q = -1 + 2 \cdot rand(5, n)$, $q = 2 + 3 \cdot rand(5, 1)$ [7]. (We denote by $rand(k_1, k_2)$ the random matrix with k_1 rows, k_2 columns and the elements generated randomly on $[0, 1]$.)

Table 2: Randomly generated problems (4) with quadratic functions

n	m	$f(x_0)$	$f(z)$	$TimeF1$	$TimeF2$
10	10	36.738565	36.429457	2.22	1.73
10	50	161.613719	157.926620	18.49	10.98
10	100	303.448935	298.051190	122.03	20.90
50	10	33.151416	31.676804	124.25	4.05
50	50	156.152869	152.686192	918.91	292.17
50	100	304.067383	298.776803	2634.33	179.12
100	10	31.809077	30.521716	1241.75	104.81
100	50	153.990301	149.812955	4986.00	410.82
100	100	302.938895	296.580627	13206.50	492.11

Note that Problems $(\mathcal{P}_\alpha)$ corresponding to the problems (3) appeared to be convex, because we deal with affine functions. In addition, the functions $\psi_i = \langle x, A_i x \rangle + \langle b^i, x \rangle$, $i = 1, \ldots, m$, constructed in the test problems (4), are convex, therefore, corresponding Problems $(\mathcal{P}_\alpha)$ happened to be convex again. That is why the F2-algorithm's run-time turned out to be much less than the run-time of F1-algorithm, in which it is necessary to solve nonconvex problems with m d.c. constraints.

Computational experiments showed that solving of the fraction program should combine the two approaches. For example, we can use the solution to Problem $(\mathcal{P}_\sigma)$ to search for the parameter α that reduces the optimal value function of Problem $(\mathcal{P}_\alpha)$ to zero.

5   CONCLUSIONS

In this paper, we showed how fractional programs can be solved by applying the Global Search Theory of d.c. optimization. The methods developed were justified and tested on an extended

set of problems with linear or quadratic functions in the numerators and denominators of the ratios.

6 Acknowledgments

References

[1] Bugarin, F., Henrion, D. and Lasserre J.-B. (2016) Minimizing the sum of many rational functions. Math. Prog. Comp., 8: 83–111.

[2] Dinkelbach, W.: On Nonlinear Fractional Programming, Manag. Sci. 13, 492–498 (1967)

[3] Freund, R.W. and Jarre, F.(2001) Solving the sum-of-ratios problem by an interior-point method. J. Global Optim., 19(1): 83–102.

[4] Gruzdeva, T.V. (2013) On a continuous approach for the maximum weighted clique problem. J. Global Optim., 56: 971–981.

[5] Gruzdeva, T.V. and Strekalovsky, A.S. (2016) An approach to fractional programming via D.C. optimization. AIP Conf. Proc., 1776: 090010.

[6] Gruzdeva, T.V. and Strekalovsky, A.S. (2016) An Approach to Fractional Programming via D.C. Constraints Problem: Local Search. In: Kochetov, Yu., et all (Eds.) DOOR-2016. LNCS, 9869 (pp. 404–417). Heidelberg: Springer.

[7] Jong, Y.-Ch. (2013) An efficient global optimization algorithm for non-linear sum-of-ratios problems. Optim. Online http://www.optimization-online.org/DB_HTML/2012/08/3586.html.

[8] Schaible, S. and Shi, J. (2003) Fractional programming: the sum-of-ratios case. Optim. Methods Softw., 18: 219–229.

[9] Strekalovsky, A.S. (2014) On solving optimization problems with hidden nonconvex structures. In: Rassias, T.M., Floudas, C.A., Butenko, S. (Eds.). Optimization in Science and Engineering (pp. 465–502). New York: Springer.

[10] Strekalovsky, A.S. (2003) Elements of nonconvex optimization [in Russian]. Novosibirsk: Nauka.

[11] Strekalovsky, A.S. (2016) On the merit and penalty functions for the d.c. optimization. In: Kochetov, Yu., et all (Eds.). DOOR-2016. LNCS, 9869 (pp. 452–465). Heidelberg: Springer.

[12] Strekalovsky, A.S. (2015) On local search in d.c. optimization problems. Applied Math. and Comput., 255: 73–83.

Finiteness of the quadratic primal simplex method when s-monotone index selection rules are applied

Tibor Illés

Coauthors: Adrienn Csizmadia, Zsolt Csizmadia
Budapest University of Technology and Economics, FICO Plc.

`illes@math.bme.hu, adriennagy@gmail.com, zsolt.csizmadia@gmail.com`

Let us consider the *linearly constrained quadratic optimization problem (QP)*

$$\min \frac{1}{2}\, \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x}$$
$$A\mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0},$$

where $Q \in \mathbb{R}^{n \times n}$ and $A \in \mathbb{R}^{m \times n}$ are matrices, $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ are vectors and $\mathbf{x} \in \mathbb{R}^n$ is the vector of unknowns.

We prove the finiteness of the primal quadratic simplex algorithm (PQSA) when applied to the linearly constrained convex quadratic optimization problem, and when ties are resolved using anti-cycling index selection rules. The original quadratic simplex method was developed by Wolfe [8] and Panne and Whinston [4, 5, 6, 7], and was published in several papers in the 1960s. In the original presentation, finiteness was ensured by the means of perturbation.

We show that for the primal quadratic simplex algorithm to cycle, it is necessary that the problem is degenerate (we show more, in fact that there are basis in which all variables in the basis taking part in the ratio test has a primal value of zero), and that in the Karush-Kuhn-Tucker system associated with the problem, all coefficients in the transformed column that correspond to the quadratic objective are zero.

It follows from our proof that the PQSA is finite with the application of all those index selection rules, that only rely on the sign structure of the transformed right-hand-side and the reduced costs, and for which the traditional linear programming primal simplex method is finite [1, 3].

Although our prove of finiteness for PQSA seems to be of theoretical nature, we are convinced that similarly to the linear programming case, the new anti-cycling index selection rules [1, 3] opens up new possibilities in the implementation of PQSA, like for pivot algorithms applied to solve linear programming problems [2].

References

[1] Zs. Csizmadia, T. Illés and A. Nagy, The s-monotone index selection rules for pivot algorithms of linear programming. *European Journal of Operational Research*, 221(3):491–500, 2012.

[2] T. Illés and A. Nagy, Computational aspects of simplex and MBU-simplex algorithms using different anti-cycling pivot rules. *Optimization*, 63(1):49–66, 2014.

[3] A. Nagy, *On the theory and applications of flexible anti-cycling index selection rules for linear optimization problems*. Eötvös Loránd University of Sciences, Budapest, 2014.

[4] C. van de Panne, A. Whinston. Simplicial methods for quadratic programming. *Naval Research Logistics*, 11:273–302, 1964.

[5] C. van de Panne, A. Whinston. The Simplex and the Dual Method for Quadratic Programming. *Operational Research Quarterly*, 15:355–388, 1964.

[6] C. van de Panne, A. Whinston. A Parametric Simplicial Formulation of Houthakker's Capacity Method. *Econometrica*, 34(2):354–380, 1966.

[7] C. van de Panne, A. Whinston. The Symmetric Formulation of the Simplex Method for Quadratic Programming. *Econometrica*, 37(3):507–527, 1969.

[8] P. Wolfe, The simplex method for quadratic programming. *Econometrica*, 27(3):382–398, 1959.

DUALITY GAP IN INTERVAL LINEAR PROGRAMMING

Jana Novotná

Department of Applied Mathematics of the Faculty of Mathematics and Physics at the Charles University, Malostranské náměstí 25, Praha 1, Czech Republic

email: `janca@kam.mff.cuni.cz`

Milan Hladík

Department of Applied Mathematics of the Faculty of Mathematics and Physics at the Charles University, Malostranské náměstí 25, Praha 1, Czech Republic

email: `hladik@kam.mff.cuni.cz`

Tomáš Masařík

Department of Applied Mathematics of the Faculty of Mathematics and Physics at the Charles University, Malostranské náměstí 25, Praha 1, Czech Republic

email: `masarik@kam.mff.cuni.cz`

Abstract:

This paper deals with the problem of linear programming with inexact data represented by real intervals. We extend the concept of duality gap (DG), the difference between the primal and its dual optimal value, into interval linear programming (ILP). We give characterizations of strongly- and weakly-zero DG in ILP and its special case where the matrix of coefficients is real. We show computational complexity of testing weakly- and strongly-zero DG for commonly used types of ILP.

Keywords: Interval Analysis, Linear Programming, Interval Linear Programming, Duality Gap.

1 INTRODUCTION

The area of interval linear programming connects linear programming with interval analysis which is one of approches to deal with inexact data. Optimization problems with interval data are intensively studied for last forty years, see the survey by Hladík for more details [3].

We extend a concept of duality gap (i.e. the difference between the primal and the dual optimal value) from linear programming into the seting of interval linear programming. In our best knowledge we are the first who study this topic in the context of interval linear programming despite the fact that duality gap is one of the basic concepts of linear programming and a very useful property to describe behavior of programs with respect to their duals. We investigate weakly- and strongly-zero duality gap. In particular the second property, strongly-zero DG, is very useful, it allows us to move from a primal to its dual program. Our main interest is in characterizations of weakly- and strongly-zero duality gap and testing their computational complexity.

We focus on interval linear programs and also on their special case — interval linear programs with real constraint matrix (degenerated matrix). Even programs with degenerated matrix are studied [1] in recent years and often used in practical computation (network flows, transportation problems).

1.1 Problem formulation

Given two matrices $\underline{A}$, $\overline{A} \in \mathbf{R}^{m \times n}$ such that $\underline{A} \leq \overline{A}$, we define *interval matrix* $\mathbf{A}$ as the set

$$\mathbf{A} = [\underline{A}, \overline{A}] = \{A \in \mathbf{R}^{m \times n} : \underline{A} \leq A \leq \overline{A}\},$$

where $\leq$ means element-by-element comparison of the coresponding matrices. The set of all interval $m \times n$ matrices is denoted by $\mathbb{IR}^{m \times n}$.

Let $\boldsymbol{A} \in \mathbb{IR}^{m \times n}, \boldsymbol{b} \in \mathbb{IR}^m, \boldsymbol{c} \in \mathbb{IR}^n$ be given. We define an *interval linear program* (ILP[1]) as a family of linear programs minimize $c^\top x$ subj. to $x \in \mathcal{M}(A,b)$ where $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$, $c \in \boldsymbol{c}$ and $\mathcal{M}(\boldsymbol{A}, \boldsymbol{b})$ is a set of linear equations and inequalities. We write it in short as

$$\text{minimize} \quad c^\top x \text{ subj. to } x \in \mathcal{M}(\boldsymbol{A}, \boldsymbol{b}). \tag{1}$$

Duality in ILP straightforwardly extends duality in linear programming.

Duality gap

Duality gap is the difference of optimal values of an primal and its dual program in mathematical programming. Duality gap is zero if and only if at least one of the primal or its dual program is feasible.

Following the concept of mathematical programming we define *duality gap (DG) of an ILP* as the set of duality gaps of all instances of that ILP.

Definition 1.1 *Duality gap of* (1) *is said to be* weakly-zero *if at least one scenario has zero duality gap and* strongly-zero *if all scenarios have zero duality gap.*

Note that if optimal value set of an ILP is not $\{\infty\}$, DG is at least weakly-zero since there exists a scenario with another optimal value (i.e. scenario with zero duality gap).

1.2 Notations, Preliminaries

A concrete realizations of interval values, i.e. a linear program with real values, is called *a scenario*. We say that ILP is *weakly*, resp. *strongly, feasible* if it is feasible for at least one scenario, resp. all scenarios. We denote the *optimal value set* as the set of optimal values of all scenarios. An interval matrix is said to be *degenerated* if $\underline{A} = \overline{A}$. We are using the term *ILP with degenerated matrix* which means an ILP with a real constraint matrix and an interval right hand side and an optimization function. Since different forms of ILP do not have to be equivalent (unlike in linear programming) we have to distinguish between them. Duality in ILP straightforwardly extends duality in linear programming, for some results see [4] and [6]. Basic types and their dual programs are shown in Table 1.

Type	Primal ILP	Dual ILP
(A)	minimize $c^\top x$ subj. to $\boldsymbol{A}x = \boldsymbol{b}$, $x \geq 0$	maximize $\boldsymbol{b}^\top y$ subj. to $\boldsymbol{A}^\top y \leq \boldsymbol{c}$
(B)	minimize $c^\top x$ subj. to $\boldsymbol{A}x \leq \boldsymbol{b}$	maximize $\boldsymbol{b}^\top y$ subj. to $\boldsymbol{A}^\top y = \boldsymbol{c}$, $y \leq 0$
(C)	minimize $c^\top x$ subj. to $\boldsymbol{A}x \leq \boldsymbol{b}$, $x \geq 0$	maximize $\boldsymbol{b}^\top y$ subj. to $\boldsymbol{A}^\top y \leq \boldsymbol{c}$, $y \leq 0$

Table 1: Dual programs.

1.3 Our results

We have introduced the concept of DG to interval linear programming. We give a full characterization of weakly-zero DG in ILP (Prop 2.1) and of strongly-zero DG in ILP with degenerated matrix (Thm 2.6). We give sufficient and necessary conditions of strongly-zero DG in ILP (Prop 2.3 and 2.5). We have determined the complexity of deciding strongly- and weakly-zero DG for types (A), (B) and (C), see Table 2, and their degenerated variants, see Table 3, except strongly-zero DG for type (C) ILP which remains open.

---

[1]Do not confuse it with the same abbreviation ILP for Integer Linear Programming.

Type	Weakly-zero DG	Strongly-zero DG
(A)	NP-hard, Thm 3.3	co-NP-hard, Thm 3.1
(B)	NP-hard, Thm 3.3	co-NP-hard, Thm 3.1
(C)	polynomial, Thm 3.4	unknown

Table 2: Computational complexity of both strongly- and wekly-zero DG.

Type	Weakly-zero DG	Strongly-zero DG
(A)	polynomial, Thm 3.5	co-NP-hard, Thm 3.1
(B)	polynomial, Thm 3.5	co-NP-hard, Thm 3.1
(C)	polynomial, Thm 3.4	polynomial, Thm 3.2

Table 3: Computational complexity of both strongly- and weakly-zero DG with degenerated matrix.

2 DUALITY GAP CHARACTERIZATIONS

In linear programming the duality gap is not zero if and only if the primal and its dual programs are infeasible. We observe that similar conditions hold in interval linear programming. We obtain full characterization for weakly-zero DG, but for strongly-zero DG we have only sufficient and necessary conditions. However, for the degenerate matrix there is a full characterisation of strongly-zero DG as well.

We can easily observe the following full characterization of weakly-zero DG.

Proposition 2.1 *DG is weakly-zero if and only if primal or dual ILP is weakly feasible.*

This observation together with the fact, that type (C) is weakly feasible if and only if the system $\underline{A}x \leq \overline{b}$, $x \geq 0$ is feasible [3], has a nice consequence.

Corollary 2.2 *Type (C)* minimize $c^\top x$ subj. to $Ax \leq b$, $x \geq 0$ *has weakly-zero DG if and only if at least one of the linear systems* $\underline{A}x \leq \overline{b}$, $x \geq 0$ *and* $\underline{A}^\top y \leq \overline{c}$, $y \leq 0$ *is feasible.*

We can easily observe the following propositions for strongly-zero DG.

Proposition 2.3 *DG is strongly-zero if the primal program or its dual counterpart is strongly feasible.*

The proposition above does not hold as an equivalence in general because an ILP can contain just only instances with primal unbounded and dual infeasible and vice versa, see Example 2.4. That ILP has strongly-zero DG, but neither primal nor dual program are strongly-feasible.

The full characterization of strongly-zero DG remains open.

Example 2.4 *Let us have a following primal and its dual program.*

$$\begin{array}{ll} \text{minimize} & -y_2 \\ \text{subj. to } [-1,1]y_1 - y_2 = 1 \\ y_1, y_2 \geq 0. \end{array} \qquad \begin{array}{ll} \text{maximize} & x \\ \text{subj. to } [-1,1]x \leq 0 \\ -x \leq -1 \end{array}$$

Depending on the choise of a value from the interval $[-1,1]$ *either the primal is infeasible and the dual unbounded, for* $[-1,0]$, *or vice versa, for* $(0,1]$.

Strong feasibility appears in sufficient conditions and weak feasibility appears in a necessary condition. We have also a neccessary condition for strongly-zero DG. Our neccessary condition for strongly-zero DG uses weak feasibility of ILP.

Proposition 2.5 *If an ILP has strongly-zero duality gap, then the primal or its dual program is weakly feasible.*

2.1 ILP with degenerate intervals

The next theorem gives us the full characterization of strongly-zero DG. It shows that the converse implication in Proposition 2.3 holds for an ILP with degenerated matrix.

Theorem 2.6 *An ILP with degenerated matrix has strongly-zero DG if and only if primal or dual programs are strongly-feasible.*

Proof. The backward implication holds according to Proposition 2.3.

We suppose for contradiction that primal and dual programs are not strongly-feasible. Thus there exist a selection $b \in \boldsymbol{b}$ such that those scenario of the primal is infeasible. Thus each scenario $(A, b, \boldsymbol{c})$ of primal ILP is infeasible. Analogously there exists $c \in \boldsymbol{c}$ such that for each scenario $(A, \boldsymbol{b}, c)$ the dual is infeasible. These two selections are independent of each other. Let (A, b, c) be a scenario of the ILP then this scenario has both dual and primal programs infeasible, i.e. duality gap of the ILP can not be strongly-zero. □

Theorem 2.6 together with fact, that type (C) is weakly feasible if and only if the system $\underline{A}x \leq \bar{b}$, $x \geq 0$ is feasible [3], give us a nice consecquence for type (C) which uses only two classic linear programs.

Corollary 2.7 *A type (C) ILP with degenerated matrix* minimize $\boldsymbol{c}^{\top}x, Ax \leq \boldsymbol{b}, x \geq 0$ *has strongly-zero DG if and only if at least one of the systems $Ax \leq \underline{b}, x \geq 0$ and $A^{\top}y \leq \underline{c}, y \leq 0$ is feasible.*

3   COMPUTATIONAL COMPLEXITY

We study computational complexity of strongly- and weakly-zero DG. Here, we provide the proofs of the theorems summarized in the introduction, see Table 2 for ILP and Table 3 for degenerated case.

3.1 Strongly-zero DG

Theorem 3.1 *Let* minimize $\boldsymbol{c}^{\top}x$ subj. to $\boldsymbol{A}x = \boldsymbol{b}, x \geq 0$ *be a type (A) ILP. Then decide whether it has the strongly-zero DG is* co-NP-*hard. Moreover, the problem remains* co-NP-*hard even for degenerated matrix $\boldsymbol{A}$.*

Proof. We construct a polynomial-time reduction from the problem of testing strong feasibility for interval linear program of type (A). This problem is known to be co-NP-hard [5] even with degenerated matrix [2].

The ILP $\boldsymbol{A}x = \boldsymbol{b}, x \geq 0$ is strongly-feasible if and only if $\boldsymbol{A}x = \boldsymbol{b}, u - v = d, x \geq 0$ is strongly-feasible. Let

$$\text{minimize } \boldsymbol{c}^{\top}x + u - 2v \text{ subj. to } \boldsymbol{A}x = \boldsymbol{b}, u - v = d, x \geq 0, \tag{2}$$

$$\text{maximize } \boldsymbol{b}^{\top}y + dz \text{ subj. to } \boldsymbol{A}^{\top}y \leq \boldsymbol{c}, z \leq 1, -z \leq -2, \tag{3}$$

be a primal and its dual program. Observe that the Program (3) is feasible for no scenario. It follows that the Program (2) is strongly feasible if and only if it admits strongly-zero DG. □

The same result holds for type (B) ILP as well since type (A) and (B) are mutually dual and then we can use the same reduction to prove co-NP-hardness. To proceed with the reduction it is necessary to observe that the minimization and the maximization ILP has the same complexity with respect to the determination of zero DG.

We remark that we cannot easily extend the hardness result to completeness result because a scenario of ILP cannot be a polynomially large certificate because it could contain an arbitrary real value. Instead, it could be interesting to develop some other characterization that can give

us bounds on the values in the scenario. Sometimes an exponential algorithm could give us that as a byproduct.

According to Corollary 2.7 it suffices to solve two linear programs to decide whether DG is strongly-zero for a type (C) ILP with degenerated matrix. Thus it can be decided in polynomial time.

Theorem 3.2 *Let be a type (C) ILP with degenerated matrix. Then decide whether it has the strongly-zero DG is polynomial.*

3.2 Weakly-zero DG

We show the hardness of the decision of weakly-zero DG, as well. We follow a similar schema of the reduction as in the strong case.

Theorem 3.3 *Let* minimize $c^\top x$ subj. to $\boldsymbol{A}x \le \boldsymbol{b}$ *be a type (B) ILP. Then decide whether it has weakly-zero DG is* NP-*hard.*

Proof. We construct a polynomial-time reduction from the problem of testing strong infeasibility for interval linear program of type (B). This problem is known to be co-NP-hard [5].

The ILP $\boldsymbol{A}x \le \boldsymbol{b}$ is strongly-infeasible if and only if $\boldsymbol{A}x \le \boldsymbol{b}, u + v \le 1$ is strongly-infeasible. Let

$$\text{minimize} \quad c^\top x + u + v \text{ subj. to } \boldsymbol{A}x \le \boldsymbol{b},\ u + v \le 1, \tag{4}$$

$$\text{maximize} \quad \boldsymbol{b}^\top y + z \text{ subj. to } \boldsymbol{A}^\top y = \boldsymbol{c},\ z = 1,\ z = 2,\ y, z \le 0 \tag{5}$$

be a primal and its dual program. Observe that the dual Program (5) is feasible for no scenario. It follows that the original Program (4) is not strongly-infeasible (and therefore it is weakly feasible) if and only if it admits weakly-zero DG. $\square$

Again, The same result holds for ILP of type (A) as well since type (A) and (B) are mutually dual and then we can use the same reduction to prove NP-hardness similarly as in the strongly-zero DG case.

Theorem 3.4 *Let* minimize $c^\top x$ subj. to $\boldsymbol{A}x \ge \boldsymbol{b}, x \ge 0$ *be a type (C) ILP. Then decide whether it has weakly-zero DG problem is polynomial.*

Proof. The proof follows from Corollary 2.2. It suffices to solve two classical linear programs. $\square$

Moreover, testing weakly-zero DG of any type with degenerated matrix is polynomial-time solvable.

Theorem 3.5 *Let be a type (A), type (B) or type (C) ILP with degenerated matrix. Then decide whether it has weakly-zero DG is polynomial.*

Proof. Since testing weak feasibility of types (A), (B) and (C) with degenerated matrix is polynomial [2] the testing weakly-zero DG is according to Proposition 2.1 polynomial too. $\square$

4 CONCLUSIONS

The computational complexity of testing strongly-zero DG of type (C) ILP and the full characterization of strongly-zero DG remain open. Another interesting problem is to find probably super-polynomial algorithms that can decide whether DG is strongly- or weakly-zero for those problems that were proven to be (co)NP-hard.

5   Acknowledgements

References

[1] Gabrel, V., Murat, C. and Remli, N. (2010). Linear programming with interval right hand sides. Int. Trans. Oper. Res., 17(3):397–408.

[2] Garajová, E. and Hladík M. (2017+). On the Properties of Interval Linear Programs with a Fixed Coefficient Matrix. To appear in Springer Proceedings in Mathematics and Statistics, ODS2017 International Conference on Optimization and Decision Science, Sorrento, September 4-7, 2017. Springer.

[3] Hladík M. (2012) Interval linear programming: A survey. In Mann Z. A., editor, Linear Programming – New Frontiers in Theory and Applications, chapter 2, pages 85–120. Nova Science Publishers, New York.

[4] Rohn J. (2006). Interval linear programming. In Fiedler M., Nedoma J., Ramík J., Rohn J. and Zimmermann K., editors, Linear optimization problems with inexact data, chapter 3, pages 79–100. Springer, New York.

[5] Rohn J. (2006), Solvability of systems of interval linear equations and inequalities. In Fiedler M., Nedoma J., Ramík J., Rohn J. and Zimmermann K., editors, Linear optimization problems with inexact data, chapter 2, pages 35–77. Springer, New York.

[6] Serafini P. (2005). Linear programming with variable matrix entries. Oper. Res. Lett., 33(2):165–170.

FINDING THE NASH EQULIBRIA IN RANDOMLY GENERATED HEXAMATRIX GAMES

Andrei Orlov

Matrosov Institute for System Dynamics & Control Theory SB RAS,
Lermontov str., 134, Irkutsk, 664033, Russia

Abstract: This paper addresses numerical solution of a 3-player polymatrix game (hexamatrix game) which is is reduced to a special nonconvex optimization problem with a bilinear structure in the objective function. For the latter problem, the numerical method for finding the Nash equilibrium is built. This method is based on the Global Search Theory developed by A.S. Strekalovsky and was tested on randomly generated hexamatrix games to demonstrate its efficiency and reliability.

Keywords: hexamatrix games, Nash equilibrium, nonconvex optimization problem, Global Search Theory, random generated test problems, numerical solution

1   PRELIMINARIES AND PROBLEM FORMULATION

It is well known that the problem of numerical finding of equilibrium points in game theory is one of the foundational problems of Mathematical Optimization [3, 10]. In this work a new numerical approach to finding the Nash equilibria in hexamatrix games is presented [1, 7, 8, 9, 15]. The approach is based on the equivalence theorem for the game and a special mathematical optimization problem with a bilinear structure in the objective function [15]. This special problem is solved by the Global Search Theory (GST) developed by A.S. Strekalovsky [11, 14].

According to the GST, a global search consists of two main stages: 1) a local search, which takes into account the structure of the problem in question; 2) the procedures based on the Global Optimality Conditions (GOCs) [11, 14], which allow us to improve the point provided by the local search method. This approach turned out to be rather effective and promising when applied to solving several real-life nonconvex problems of Operations Research [5, 6, 12, 13].

Let us recall the formulation of a hexamatrix game with mixed strategies [1, 7, 8, 9, 15]:

$$\left.\begin{aligned} F_1(x,y,z) &\triangleq \langle x, A_1 y + A_2 z \rangle \uparrow \max_x,\ x \in S_m,\quad F_2(x,y,z) \triangleq \langle y, B_1 x + B_2 z \rangle \uparrow \max_y,\ y \in S_n,\\ F_3(x,y,z) &\triangleq \langle z, C_1 x + C_2 y \rangle \uparrow \max_z,\ z \in S_l, \end{aligned}\right\}$$

where $S_p = \{u = (u_1, \ldots, u_p)^T \in \mathbb{R}^p \mid u_i \geq 0,\ \sum_{i=1}^{p} u_i = 1\}$, $p = m, n, l$, and the symbol "$\triangleq$" means "equals by definition".

The goal is to find a Nash equilibrium (approximately) [8, 9, 15] in the game $\Gamma_3 = \Gamma(A, B, C)$ ($A = (A_1, A_2)$, $B = (B_1, B_2)$, $C = (C_1, C_2)$). In such an equilibrium none of the players are profitable to change its optimal strategy. It is well known that due to Nash's Theorem [3, 15] there exists a Nash equilibrium in the game $\Gamma_3 = \Gamma(A, B, C)$ with mixed strategies.

2   REDUCTION THEOREM AND GLOBAL SEARCH METHOD

Consider the following optimization problem ($\sigma \triangleq (x, y, z, \alpha, \beta, \gamma)$):

$$\left.\begin{aligned} \Phi(\sigma) &\triangleq \langle x, A_1 y + A_2 z \rangle + \langle y, B_1 x + B_2 z \rangle + \langle z, C_1 x + C_2 y \rangle - \alpha - \beta - \gamma \uparrow \max_\sigma,\\ \sigma \in D &\triangleq \{(x, y, z, \alpha, \beta, \gamma) \in \mathbb{R}^{m+n+l+3} \mid x \in S_m,\ y \in S_n,\ z \in S_l,\\ &\quad A_1 y + A_2 z \leq \alpha e_m,\quad B_1 x + B_2 z \leq \beta e_n,\quad C_1 x + C_2 y \leq \gamma e_l\}, \end{aligned}\right\} \quad (\mathcal{P})$$

where α, β, γ are additional scalar variables, $e_p = (1, 1, ..., 1) \in \mathbb{R}^p, p = m, n, l$.

Theorem 2.1 [15] *A point* (x^*, y^*, z^*) *is a Nash equilibrium point in the hexamatrix game* $\Gamma(A, B, C) = \Gamma_3$ *if and only if it is a part of a global solution* $\sigma_* \triangleq (x^*, y^*, z^*, \alpha_*, \beta_*, \gamma_*) \in \mathbb{R}^{m+n+l+3}$ *to Problem* $(\mathcal{P})$. *At the same time, the numbers* α_*, β_*, *and* γ_* *are the payoffs of the first, the second, and the third players, respectively, in the game* Γ_3. *In addition, an optimal value* $\mathcal{V}(\mathcal{P})$ *of Problem* $(\mathcal{P})$ *is equal to zero:*

$$\mathcal{V}(\mathcal{P}) = \Phi(x^*, y^*, z^*, \alpha_*, \beta_*, \gamma_*) = 0. \tag{1}$$

Thus, one can conclude that the search for a Nash equilibrium can be carried out by solving Problem $(\mathcal{P})$. It can be also proved that if an approximate solution to Problem $(\mathcal{P})$ is obtained, then we have an approximate Nash equilibrium [15]. However, at present numerical solution of the nonconvex Problem $(\mathcal{P})$ seems to be rather difficult [2, 4], because the classical methods of convex optimization (see e.g. [4]) cannot provide a global solution to nonconvex problems, and they are not capable of escaping a local optimum. In order to solve Problem $(\mathcal{P})$, in this work we employ the Global Search Theory for nonconvex problems with d.c. functions, i.e. the functions that can be represented as a difference of two convex functions [11, 14].

According to the GST, first, one has to build an explicit d.c. representation for the objective function Φ. It can be done, for example, in the following way:

$$\Phi(x, y, z, \alpha, \beta, \gamma) = h(x, y, z) - g(x, y, z, \alpha, \beta, \gamma), \quad \text{where} \tag{2}$$

$$
\begin{aligned}
h(x, y, z) &= \frac{1}{4}\Big(\|x + A_1 y\|^2 + \|x + A_2 z\|^2 + \|B_1 x + y\|^2 + \|y + B_2 z\|^2 + \\
&+ \|C_1 x + z\|^2 + \|C_2 y + z\|^2 \Big), \quad g(\sigma) = \frac{1}{4}\Big(\|x - A_1 y\|^2 + \|x - A_2 z\|^2 + \\
&+ \|B_1 x - y\|^2 + \|y - B_2 z\|^2 + \|C_1 x - z\|^2 + \|C_2 y - z\|^2 \Big) + \alpha + \beta + \gamma.
\end{aligned} \tag{3}
$$

It is easy to see that these functions are convex on (x, y, z) and σ, respectively.

Now, let us consider the GOCs for Problem $(\mathcal{P})$, which constitute the core of the Global Search Algorithm (GSA) [8, 11, 12, 14].

Theorem 2.2 [8, 11, 12, 14] *If a feasible 6-tuple* $\sigma_* = (x^*, y^*, z^*, \alpha_*, \beta_*, \gamma_*)$ *is not a global solution to Problem* $(\mathcal{P})$, *then there exists a triple* $(u, v, w) \in \mathbb{R}^{m+n+l}$, *a vector* $(\bar{x}, \bar{y}, \bar{z}) \in S_m \times S_n \times S_l$, *and a scalar* ξ, *such that*

$$h(u, v, w) - \xi = \zeta \triangleq \Phi(\sigma_*) < 0, \quad g(u, v, w, \alpha(v, w), \beta(u, w), \gamma(u, v)) \le \xi \le \sup(g, D), \tag{4}$$

and the following inequality takes place:

$$g(\bar{x}, \bar{y}, \bar{z}, \bar{\alpha}, \bar{\beta}, \bar{\gamma}) - \xi < \langle \nabla_{xyz} h(u, v, w), (\bar{x}, \bar{y}, \bar{z}) - (u, v, w) \rangle, \tag{5}$$

where $\bar{\alpha} \triangleq \max\limits_{1 \le i \le m} (A_1 \bar{y} + A_2 \bar{z})_i$, $\bar{\beta} \triangleq \max\limits_{1 \le j \le n} (B_1 \bar{x} + B_2 \bar{z})_j$, $\bar{\gamma} \triangleq \max\limits_{1 \le t \le l} (C_1 \bar{x} + C_2 \bar{y})_t$. $\qquad \square$

These GOCs possess a so-called algorithmic (constructive) property [11, 12]. For Problem $(\mathcal{P})$, it means that if one successfully found the 4-tuple $(\bar{u}, \bar{v}, \bar{w}, \bar{\xi})$ from (4), and the point $\bar{\sigma} \triangleq (\bar{x}, \bar{y}, \bar{z}, \bar{\alpha}, \bar{\beta}, \bar{\gamma})$, $(\bar{x}, \bar{y}, \bar{z}) \in S_m \times S_n \times S_l$, such that the inequality (5) holds: $g(\bar{\sigma}) - \bar{\xi} < \langle \nabla_{xyz} h(\bar{u}, \bar{v}, \bar{w}), (\bar{x}, \bar{y}, \bar{z}) - (\bar{u}, \bar{v}, \bar{w}) \rangle$, then, due to the convexity of the function $h(\cdot)$ and the equality in (4), one obtains $\Phi(\bar{\sigma}) > \Phi(\sigma_*)$. It means that the 6-tuple $\bar{\sigma} = (\bar{x}, \bar{y}, \bar{z}, \bar{\alpha}, \bar{\beta}, \bar{\gamma})$ is better than the 6-tuple $\sigma_* = (x^*, y^*, z^*, \alpha_*, \beta_*, \gamma_*)$.

The constructive property forms the basis of the global search algorithms for nonconvex problems [11, 12, 14]. For the hexamatrix games, the GSA can be formulated in the following way.

Let there be given a starting point $(x^0, y^0, z^0, \alpha_0, \beta_0, \gamma_0) \in D$, numerical sequences $\{\tau_k\}$, $\{\delta_k\}$. $(\tau_k, \delta_k > 0, \ k = 0, 1, 2, ..., \tau_k \downarrow 0, \delta_k \downarrow 0, (k \to \infty))$, a set $Dir = \{(\bar{u}^1, \bar{v}^1, \bar{w}^1), ..., (\bar{u}^N, \bar{v}^N, \bar{w}^N) \in$

$I\!\!R^{m+n+l} \mid (\bar{u}^p, \bar{v}^p, \bar{w}^p) \neq 0, p = 1, ..., N\}$, the numbers $\xi_- \triangleq \inf(g, D)$ and $\xi_+ \triangleq \sup(g, D)$, and parameters of the algorithm η and M.

Step 0. Set $k := 0$, $(\bar{x}^k, \bar{y}^k, \bar{z}^k, \bar{\alpha}_k, \bar{\beta}_k, \bar{\gamma}_k) := (x^0, y^0, z^0, \alpha_0, \beta_0, \gamma_0)$, $p := 1$, $\xi := \xi_-$, $\triangle\xi = (\xi_+ - \xi_-)/M$.

Step 1. Start the local search method from the point $(\bar{x}^k, \bar{y}^k, \bar{z}^k, \bar{\alpha}_k, \bar{\beta}_k, \bar{\gamma}_k)$ and construct a τ_k-critical point $\sigma_k := (x^k, y^k, z^k, \alpha_k, \beta_k, \gamma_k) \in D$ to Problem $(\mathcal{P})$. Set $\zeta_k := \Phi(x^k, y^k, z^k, \alpha_k, \beta_k, \gamma_k)$.

Step 2. Using $(\bar{u}^p, \bar{v}^p, \bar{w}^p) \in Dir$, construct a point (u^p, v^p, w^p) of the approximation $\mathcal{A}_k = \{(u^1, v^1, w^1), ..., (u^N, v^N, w^N) \mid h(u^p, v^p, w^p) = \xi + \zeta_k, p = 1, ..., N\}$ of the level surface $\mathcal{U}(\zeta_k) = \{(x, y, z) \mid h(x, y, z) = \xi + \zeta_k\}$ of the convex function $h(x, y, z) : h(u^p, v^p, w^p) = \xi + \zeta_k$.

Step 3. If $g(u^p, v^p, w^p, \alpha(v^p, w^p), \beta(u^p, w^p), \gamma(u^p, v^p)) > \xi + \eta\xi$, then $p := p + 1$ and return to Step 2.

Step 4. Find a δ_k-solution $(\bar{x}^p, \bar{y}^p, \bar{z}^p, \bar{\alpha}_p, \bar{\beta}_p, \bar{\gamma}_p)$ of the following linearized problem:

$$g(x, y, z, \alpha, \beta, \gamma) - \langle \nabla h(u^p, v^p, w^p), (x, y, z) \rangle \downarrow \min_\sigma, \quad \sigma = (x, y, z, \alpha, \beta, \gamma) \in D. \qquad (\mathcal{PL}_p)$$

Step 5. Proceeding from the point $(\bar{x}^p, \bar{y}^p, \bar{z}^p, \bar{\alpha}_p, \bar{\beta}_p, \bar{\gamma}_p)$, build a τ_k-critical point $\hat{\sigma}_p := (\hat{x}^p, \hat{y}^p, \hat{v}^p, \hat{\alpha}_p, \hat{\beta}_p, \hat{\gamma}_p) \in D$ to Problem $(\mathcal{P})$ by means of the local search method.

Step 6. If $\Phi(\hat{\sigma}_p) \leq \Phi(\sigma_k)$, $p < N$, then set $p := p + 1$ and return to Step 2.

Step 7. If $\Phi(\hat{\sigma}_p) \leq \Phi(\sigma_k)$, $p = N$ and $\xi < \xi_+$, then set $\xi := \xi + \triangle\xi$, $p := 1$ and go to Step 2.

Step 8. If $\Phi(\hat{\sigma}_p) > \Phi(\sigma_k)$, then set $\xi := \xi_-$, $(\bar{x}^{k+1}, \bar{y}^{k+1}, \bar{v}^{k+1}, \bar{\alpha}_{k+1}, \bar{\beta}_{k+1}, \bar{\gamma}_{k+1}) := (\hat{x}^p, \hat{y}^p, \hat{v}^p, \hat{\alpha}_p, \hat{\beta}_p, \hat{\gamma}_p)$, $k := k + 1$, $p := 1$ and return to Step 1.

Step 9. If $\Phi(\hat{\sigma}_p) \leq \Phi(\sigma_k)$, $p = N$ and $\xi = \xi_+$, then stop. The point $(x^k, y^k, z^k, \alpha_k, \beta_k, \gamma_k)$ is the obtained solution to the problem.

It can be readily seen that this Algorithm is not an algorithm in the usual sense, because some steps are not specified in it. For example, we do not know how to construct a starting point and the set Dir, how to implement a local search or how to solve the problem $(\mathcal{PL}_p)$ etc. These issues will be considered below.

3   IMPLEMENTATION OF THE GLOBAL SEARCH

First, note that a feasible starting point can be constructed by using the barycenters of standard simplexes: $x_i^0 = \dfrac{1}{m}$, $i = 1, ..., m$; $y_j^0 = \dfrac{1}{n}$, $j = 1, ..., n$; $z_t^0 = \dfrac{1}{l}$, $t = 1, ..., l$;
$\alpha_0 = \max_i(A_1 y^0 + A_2 z^0)_i$; $\beta_0 = \max_j(B_1 x^0 + B_2 z^0)_j$; $\gamma_0 = \max_t(C_1 x^0 + C_2 y^0)_t$.

As for a local search (see steps 1 and 5), it can be based on the consecutive solution of the following LP problems derived from Problem $(\mathcal{P})$:

$$\left.\begin{array}{l} f_1^{(v,w)}(x, \beta) \triangleq \langle x, (A_1 + B_1^T)v + (A_2 + C_1^T)w \rangle - \beta \uparrow \max_{(x,\beta)}, \\[2mm] (x, \beta) \in X(v, w, \bar{\gamma}) \triangleq \{(x, \beta) \mid x \in S_m, \ B_1 x - \beta e_n \leq -B_2 w, \ C_1 x \leq \bar{\gamma}e_l - C_2 v\}; \end{array}\right\} \\ (\mathcal{LP}_x(v, w, \bar{\gamma}))$$

$$\left.\begin{array}{l} f_2^{(u,w)}(y, \gamma) \triangleq \langle y, (B_1 + A_1^T)u + (B_2 + C_2^T)w \rangle - \gamma \uparrow \max_{(y,\gamma)}, \\[2mm] (y, \gamma) \in Y(u, w, \bar{\alpha}) \triangleq \{(y, \gamma) \mid y \in S_n, \ A_1 y \leq \bar{\alpha}e_m - A_2 w, \ C_2 y - \gamma e_l \leq -C_1 u\}; \end{array}\right\} \\ (\mathcal{LP}_y(u, w, \bar{\alpha}))$$

$$\left.\begin{array}{l} f_3^{(u,v)}(z, \alpha) \triangleq \langle z, (C_1 + A_2^T)u + (C_2 + B_2^T)v \rangle - \alpha \uparrow \max_{(z,\alpha)}, \\[2mm] (z, \alpha) \in Z(u, v, \bar{\beta}) \triangleq \{(z, \alpha) \mid z \in S_l, \ A_2 z - \alpha e_m \leq -A_1 v, \ B_2 z \leq \bar{\beta}e_n - B_1 u\}. \end{array}\right\} \\ (\mathcal{LP}_z(u, v, \bar{\beta}))$$

where $(u, v, w, \bar{\alpha}, \bar{\beta}, \bar{\gamma}) \in D$ is a feasible point in Problem $(\mathcal{P})$. This type of the local search is efficient for problems with a bilinear structure [5, 8, 9, 12, 14].

Next, we should add some extra stopping criteria after Steps 1 and 5. If a value of the objective function Φ at the τ_k-critical point is approximately equal to zero, then we have already obtained an approximate Nash equilibrium and further computations should be terminated.

The key moment of the above GSA is the construction of an approximation of the level surface of the convex function $h(\cdot)$, which generates the basic nonconvexity in the problem under consideration. For Problem $(\mathcal{P})$ the approximation $\mathcal{A}_k = \mathcal{A}(\zeta_k)$ has been constructed with the help of special sets of directions. They use information from the problem statement.

$$Dir1 = \{(e^i, e^j, e^t),\ i = 1, ..., m,\ j = 1, ..., n,\ t = 1, ..., l\}.$$

$$Dir2 = \{e^r \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}.$$

$$Dir3 = \{(e^i + x^k, e^j + y^k, e^t + z^k),\ i = 1, ..., m,\ j = 1, ..., n,\ t = 1, ..., l\}.$$

$$Dir4 = \{e^r + (x^k, y^k, z^k) \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}.$$

$$Dir5 = \{([A_1]_j + [A_2]_t + e_m, [B_1]_i + [B_2]_t + e_n, [C_1]_i + [C_2]_j + e_l), i = \overline{1,m}, j = \overline{1,n}, t = \overline{1,l}\}.$$

$$Dir6 = \{[ABC]_r + e_{m+n+l} \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}.$$

$$Dir6' = \{(ABC)_r + e_{m+n+l} \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}.$$

$$Dir7 = \{e^r + (x^k, y^k, z^k) + e_{m+n+l} \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}.$$

$$Dir8 = \{((B_1)_j + (C_1)_t + e_m, (A_1)_i + (C_2)_t + e_n, (A_2)_i + (B_2)_j + e_l), i = \overline{1,m}, j = \overline{1,n}, t = \overline{1,l}\}.$$

$Dir9 = \{p^r \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}$, where vectors p^r are pairwise conjugate relative to the matrix H [4], which generates a basic nonconvexity in problem $(\mathcal{P})$ $(p^1 = (x^k, y^k, z^k))$.

$Dir10 = \{p^r \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}$ is the same but $p^1 = (x^k, y^k, z^k) - (x^{k-1}, y^{k-1}, z^{k-1})$.

$Dir11 = \{u^r \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l\}$, where u_r are eigenvectors of the matrix H.

$$Dir12 = \{e^r - e^s \in I\!\!R^{m+n+l},\ r = 1, ..., m+n+l,\ s = 1, ..., m+n+l\}.$$

Here e^i, e^j, e^t, e^r, e^s are the basic Euclidean vectors, $e_q = (1, 1, ..., 1)$ of an appropriate dimension; (x^k, y^k, z^k) is the part of a current critical point of Problem $(\mathcal{P})$; $[Q]_k$ is a k-th column of the matrix Q; $(Q)_k$ is a k-th row of the matrix Q;

$$H = \frac{1}{2}\begin{pmatrix} 2E_{m\times m} + B_1^T B_1 + C_1^T C_1 & A_1 + B_1^T & A_2 + C_1^T \\ B_1 + A_1^T & 2E_{n\times n} + A_1^T A_1 + C_2^T C_2 & B_2 + C_2^T \\ C_1 + A_2^T & C_2 + B_2^T & 2E_{l\times l} + A_2^T A_2 + B_2^T B_2 \end{pmatrix};\quad ABC = \begin{pmatrix} 0_{m\times m} & A_1 & A_2 \\ B_1 & 0_{n\times n} & B_2 \\ C_1 & C_2 & 0_{l\times l} \end{pmatrix}.$$

These sets were constructed with the help of sets for other problems with a bilinear structure which were solved earlier [5, 8, 12, 14]. In addition, for the sets $Dir1$, $Dir3$, $Dir5$, $Dir8$, $Dir12$, which contain a lot of points when the dimension of the problems grows, we propose special techniques for reducing the number of points in them (see also [12]). We will denote such sets as $Cut(DirX)$. Other stages and parameters of the GSA were realized according to our previous experience [5, 8, 11, 12, 14].

4 NUMERICAL EXPERIMENT

The GSA has been implemented with the help of MATLAB 7.11 R2010b (see http://www.mathworks.com/products/matlab/). During its numerical testing auxiliary LP problems and convex quadratic problems have been solved by IBM ILOG CPLEX 12.6.2 (see http://www-03.ibm.com/software/products/ru/ibmilogcpleoptistud). The computer with Intel Core i5-4690 CPU (3.5 GHz), 16 Gb RAM has been used. The test hexamatrix games were randomly generated by the relevant subroutines of MATLAB.

At the first stage of the computational experiment pseudorandom integer elements of matrices A_1, A_2, B_1, B_2, C_1, and C_2 have a uniform distribution and were chosen from the segment

$[-9, 9]$. The density of these matrices is equal to 0.15. The results of the GSA testing with different approximations on the problem series of dimension $(10 + 10 + 10)$ are presented in Table 1. There were 100 problems in the series.

The following denotations are used in Table 1: Dir is the number of the set of directions for the approximation of the level surface; Loc_{av} is the average number of starts of the local search method for one problem of the series; LP_{av} and QP_{av} are the average numbers of LP problems and convex quadratic problems, respectively, solved in course of the GSA; GIt_{av} is the average number of iterations of the GSA (the average number of escapes from critical points); T_{av} stands for the average CPU time for one problem of the series (in seconds); UnS is the number of problems where an approximate Nash equilibrium has not been obtained with the accuracy $\varepsilon = 10^{-3}$; $S10^{-2}$ and $S10^{-1}$ are the number of unsolved problems which were solved with the accuracies $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-1}$, respectively.

Dir	LP_{av}	Loc_{av}	QP_{av}	GIt_{av}	T_{av}	UnS	$S10^{-2}$	$S10^{-1}$
Dir1	35670.61	3510.45	585.75	1.90	38.89	2	1	1
Dir2	1675.25	134.01	23.05	2.08	1.81	4	0	1
Dir3	43818.04	5325.60	888.30	2.48	48.31	1	1	0
Dir4	14072.80	783.42	131.10	2.99	14.90	11	1	5
Dir5	220974.89	13867.37	2311.79	3.33	235.07	6	2	4
Dir6	23493.38	1601.78	267.49	3.07	24.75	22	4	14
Dir6'	29710.50	2133.64	356.10	3.37	32.00	26	5	15
Dir7	41514.35	3405.73	567.85	3.23	44.30	57	4	18
Dir8	128608.65	14129.27	2355.42	3.53	141.68	5	0	5
Dir9	2122.38	156.33	26.57	2.27	2.30	19	0	5
Dir10	1732.85	152.89	25.98	2.08	1.90	21	1	5
Dir11	8207.63	591.17	99.01	3.05	8.74	18	1	5
Dir12	16067.28	1349.25	225.55	4.66	17.30	1	0	1

Table 1. Comparison of the different approximations

On the basis of these numerical results we can conclude that some approximations can be more efficient with respect to speed and quality (the percentage of solved problems) than others. Therefore, we choose the best of them and build the following chain of approximations for solving games of large dimension: $Dir2 \bigcup \{-z\} \Rightarrow Dir4 \bigcup \{-z\} \Rightarrow Dir10 \bigcup \{-z\} \Rightarrow Cut(Dir3) \bigcup \{-z\} \Rightarrow Cut(Dir12) \bigcup \{-z\} \Rightarrow Dir11 \bigcup \{-z\}$. At the second stage of the computational experiment pseudorandom integer elements of matrices with density 0.1 were chosen from the segment $[-(m+n+l)/10, (m+n+l)/10]$. The results of solving the series of problems of dimension from $(10+10+10)$ up to $(100+100+100)$ are presented in Table 2, where $SolvLoc$ is the number of problems in the series which were solved using a local search only.

$m=n=l$	$Series$	$SolvLoc$	LP_{av}	Loc_{av}	QP_{av}	GIt_{av}	T_{av}	UnS
10	10000	1325	200.06	14.75	3.11	1.19	0.21	0
15	10000	97	806.57	45.84	8.30	2.02	0.94	0
20	1000	1	6057.86	291.23	49.20	2.44	6.98	4
25	1000	0	8243.24	575.84	96.65	2.59	10.08	7
30	1000	0	14038.57	983.67	164.64	2.70	17.81	8
40	1000	0	29920.71	1741.52	290.94	2.93	42.49	11
50	100	0	40650.39	2077.65	346.98	3.51	58.55	1
60	100	0	58387.44	2584.34	431.37	3.58	91.39	1
70	100	0	119430.68	5777.18	963.46	4.07	215.88	2
80	10	0	29301.5	985.0	164.8	3.6	57.6	0
100	10	0	117174.9	3328.4	555.3	4.7	265.0	0

Table 2. GSA testing on the series of problems of different dimension

There were 10000, 1000, 100, and 10 problems in each series depending on the dimension of the problems.

Table 2 shows that the GSA is efficient for these series of problems (more than 99% problems were solved with the given accuracy in a reasonable time). However, for the problems with dense matrices this version of the GSA is rather inefficient because the percentage of unsolved problems and the computation time of the GSA are rather large. The future work implies adjustment of the GSA to problems with dense matrices.

5   Acknowledgements

References

[1] Audet, C., Belhaiza, S. and Hansen, P. (2006). Enumeration of all the extreme equilibria in game theory: bimatrix and polymatrix games. J. Optim. Theory Appl., 129(3): 349–372.

[2] Horst, R. and Tuy, H. (1993). Global Optimization. Deterministic Approaches. Berlin: Springer-Verlag.

[3] Mazalov, V. (2014). Mathematical Game Theory and Applications. New York: John Wiley & Sons.

[4] Nocedal, J. and Wright, S.J. (2000). Numerical Optimization. New York-Berlin-Heidelberg: Springer-Verlag.

[5] Orlov, A.V. and Strekalovsky, A.S. (2005). Numerical search for equilibria in bimatrix games. Comput. Math. Math. Phys., 45(6): 947–960.

[6] Orlov, A.V. (2008). Numerical solution of bilinear programming problems. Comput. Math. Math. Phys., 48(2): 225–241.

[7] Orlov, A.V. (2015). On the optimization approach to polymatrix games. In Proceedings of the 13th International Symposium on Operational Research (SOR'15), Slovenia, Bled, September 23-25, 2015. (pp. 542-547). Ljubljana: Slovenian Society Informatika, Section for Operational Research.

[8] Orlov, A.V., Strekalovsky, A.S. and Batbileg, S. (2016) On computational search for Nash equilibrium in hexamatrix games. Optim. Lett. 10(2): 369–381.

[9] Orlov, A.V. and Strekalovsky, A.S. (2016) On a Local Search for Hexamatrix Games. CEUR Workshop Proceedings. DOOR-SUP 2016, 1623:477–488.

[10] Pang, J.-S. (2010). Three modeling paradigms in mathematical programming. Math. program. Ser.B., 125(2): 297–323.

[11] Strekalovsky, A.S. (2003). Elements of Nonconvex Optimization (in Russian). Novosibirsk: Nauka.

[12] Strekalovsky, A.S. and Orlov, A.V. (2007). Bimatrix Games and Bilinear Programming (in Russian). Moscow: FizMatLit.

[13] Strekalovsky, A.S., Orlov, A.V. and Malyshev, A.V. (2010). On computational search for optimistic solutions in bilevel problems. J. Glob. Optim., 48(1): 159–172.

[14] Strekalovsky, A.S. (2014). On Solving Optimization Problems with Hidden Nonconvex Structures. In Rassias, T.M., Floudas, C.A., Butenko, S. (Eds.). Optimization in Science and Engineering. (pp. 465–502). New-York: Springer.

[15] Strekalovsky, A.S. and Enkhbat, R. (2014). Polymatrix games and optimization problems. Autom. Remote Control, 75(4): 632–645.

Abstract
Infeasible interior-point algorithms for linear optimization problems

Rigó Petra Renáta

Coauthors: Darvay Zsolt, Illés Tibor
Budapest University of Technology and Economics, Babes-Bolyai University
takacsp@math.bme.hu, darvay@cs.ubbcluj.ro, illes@math.bme.hu

The interior point algorithms (IPA) for solving linear optimization problems (LOP) have the best known theoretical complexity. In case of these algorithms the determination of the initial (interior) point can cause difficulties.

The self-dual embedding technique introduced by Ye, Todd and Mizuno [6] and Terlaky [4] can resolve the question of finding initial interior point solution by embedding the original primal-dual problem pair into a problem that has one extra pair of decision variables related to the embedding. The embedded problem has initial, starting interior point, namely the all one vector that lies on the central path of the embedded problem. Solving the embedded problem using IPAs, decides whether the original primal-dual problem pair has optimal solution or it is infeasible.

Because some transformations of the primal-dual problem pair need to be done for producing the embedded problem and the initial, starting interior point of it, and finally some analysis and additional computations are necessary to derive final conclusion about the solvability of the original LOP, in practical implementations of IPAs have not been widely used this theoretically elegant approach. Best known implementations of IPAs uses the so called *infeasible IPAs*. The initial point for infeasible IPAs can be computed by using heuristic procedure, and may not be interior point for the original LOP. One of the most widely used heuristic is given by Mehrotra [2].

The first results that analysed infeasible IPAs were published by Lustig [1] and Tanabe [3]. Zhang [7] was the first who introduced an infeasible IPA which gives an approximate solution in polynomial time. In theory, the infeasible IPAs have two stopping criteria. One is when the method finds ϵ-optimal feasible solution. The other one proves that there is no feasible solution by using the approximate Farkas lemma [5]. But in practice the proof of infeasibility is usually given by the fact that the coordinates of some variables become too large. This practical stopping criterion does not coincide with the theoretical one. Therefore, there are some doubts among some authors about the correctness of infeasible IPAs.

Some authors analyse an infeasible IPA and proves it's polynomial complexity, by assuming that both primal and dual LOPs have interior point solution, just they do not know such an initial starting point.

In our talk, we would like to discuss some pros and cons related to infeasible IPAs from theoretical and practical point of view.

References

[1] I.J. Lustig. Feasibility issues in a primal-dual interior-point method for linear programming. *Math. Program.*, 49(1-3):145–162, 1990.

[2] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. Optim.*, 2(4):575–601, 1992.

[3] K. Tanabe. Centered Newton method for linear programming: Interior and 'exterior' point method. In K. Tone, editor, *New Methods for Linear Programming*, volume 3, pages 98–100. 1990. In Japanese.

[4] T. Terlaky. An easy way to teach interior-point methods. *Eur. J. Oper. Res.*, 130(1):1–19, 2001.

[5] M.J. Todd and Y. Ye. Approximate Farkas lemmas and stopping rules for iterative infeasible-point algorithms for linear programming. *Mathematical Programming*, 81(1):1–21, 1998.

[6] Y. Ye, M.J. Todd, and S. Mizuno. An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm. *Math. Oper. Res.*, 19:53–67, 1994.

[7] Y. Zhang. On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *SIAM J. Optim.*, 4(1):208–227, 1994.

GLOBAL OPTIMALITY CONDITIONS FOR PROBLEM WITH D.C. EQUALITY AND INEQUALITY CONSTRAINTS

Alexander S. Strekalovsky

Matrosov Institute for System Dynamics & Control Theory SB RAS,

Lermontov st., 134, Irkutsk, 664033, Russia

Abstract: This paper addresses the nonconvex optimization problem with the cost function and the inequality and equality constraints given by d.c. functions. The problem is reduced to the one without constraints by the exact penalization. Furthermore, the penalized problem is a d.c. minimization problem. For the latter problem we develop the Global Optimality Conditions (GOCs) which reduce the nonconvex optimization problem to a family of convex (linearized with respect to the basic nonconvexity) problems.

Keywords: d.c. functions, inequality constraints, exact penalty, global optimality conditions, constructive property, Lagrange function, KKT point, linearized problem

1   INTRODUCTION

A little more than 50 years ago the exact penalty method was invented by I.I. Eremin and W.I. Zangvill [1, 9] independently and almost simultaneously. This approach has been widely used and nowadays becomes increasingly popular. Furthermore, almost all real-life optimization problems are explicitly or implicitly nonconvex, with a lot (often a huge number!) of stationary vectors and local pits. For such problems the classical optimization methods provide only the KKT points. On the other hand, the methods of Global Optimization (B&B, cut's method etc.) suffer the so-called "curse of dimension", when the exponential growth of computational efforts corresponds to an increase in dimension of the problem in question.

The current research develops the new GOCs [11, 12] for the problem in question with the help of the exact penalty approach and reduces the original problem with d.c. constraints to a problem without constraints. After that, we represent the cost function of the penalized problem as a d.c. function. The latter helps us develop the necessary GOCs, which turn out to be related to the KKT Theorem in the original problem. Then we identify and explain various features of the GOC developed, in particular, the constructive property.

2   STATEMENT OF THE PROBLEM

Consider the following problem:

$$(\mathcal{P}): \quad \begin{cases} f_0(x) := g_0(x) - h_0(x) \downarrow \min_x, \quad x \in S, \\ f_i(x) := g_i(x) - h_i(x) \le 0, \quad i \in I = \{1, \ldots, m\}, \\ f_i(x) := g_i(x) - h_i(x) = 0, \quad i \in \mathcal{E} = \{m+1, \ldots, l\}; \end{cases}$$

where the functions $g_i(\cdot)$, $h_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, are convex on $\mathbb{R}^n$, so that the functions $f_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, are the d.c. functions [6, 7]. Besides, assume that the set $S \subset \mathbb{R}^n$ is convex and compact. Furthermore, suppose that the set $Sol(\mathcal{P})$ of global solutions to Problem $(\mathcal{P})$, $Sol(\mathcal{P}) := \{x \in \mathcal{F} \mid f_0(z) = \mathcal{V}(\mathcal{P})\}$ and the feasible set $\mathcal{F}$ of Problem $(\mathcal{P})$, $\mathcal{F} := \{x \in S \mid f_i(x) \le 0, i \in I, f_i(x) = 0, i \in \mathcal{E}\}$, are non-empty.

Besides, in what follows the optimal value $\mathcal{V}(\mathcal{P})$ of Problem $(\mathcal{P})$ is supposed to be finite:

$$\mathcal{V}(\mathcal{P}) := \inf(f_0, \mathcal{F}) := \inf_x \{f_0(x) \mid x \in \mathcal{F}\} > -\infty.$$

3 EXACT PENALTY

Introduce the penalty function $W(\cdot)$ for Problem $(\mathcal{P})$ as follows

$$W(x) := \max\{0, f_1(x), \ldots, f_m(x)\} + \sum_{j \in \mathcal{E}} |f_j(x)|. \tag{1}$$

Further, along with Problem $(\mathcal{P})$, consider the penalized problem without the inequality and equality constraints:

$$(\mathcal{P}_\sigma): \qquad \theta_\sigma(x) \downarrow \min_x, \quad x \in S, \tag{2}$$

where $\sigma \geq 0$ is a penalty parameter and the goal function of $(\mathcal{P}_\sigma)$ is of the form

$$\theta_\sigma(x) := f_0(x) + \sigma W(x). \tag{3}$$

As well-known, if $z \in Sol(\mathcal{P}_\sigma)$, and z is feasible in $(\mathcal{P})$, i.e. $z \in \mathcal{F}$, then z turns out to be a global solution to $(\mathcal{P})$: $z \in Sol(\mathcal{P})$ [1, 2, 4, 7, 9]. On the other hand, the inverse assertion does not, in general, hold. Hence, the principal moment of the exact penalization (EP) theory is the existence of a threshold value $\sigma_* \geq 0$ of the penalty parameter $\sigma \geq 0$ for which $Sol(\mathcal{P}_\sigma) \subset Sol(\mathcal{P}) \;\; \forall \sigma \geq \sigma_*$. In other words, for $\sigma \geq \sigma_*$ Problems $(\mathcal{P})$ and $(\mathcal{P}_\sigma)$ turn out to be equivalent in the sense that $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$ (see [7, Chapt. VII, Lemma 1.2.1]). Besides, the existence of exact penalty parameter $\sigma_* \geq 0$ allows us to solve a single unconstrained problem instead of a sequence of unconstrained problems with $\sigma_k \to \infty$ [2, 5].

Recall that under various constraint qualification (CQ) conditions (MFCQ, etc. [2, 8]), the error bound properties [1, 2, 8, 9], the metric sub-regularity conditions, calmness of constraints systems can help to prove the existence of the exact penalty threshold $\sigma_* \geq 0$ even for a global solution [2, 4]. Assume that some regularity condition is fulfilled that ensures the existence of such threshold value $\sigma_* \geq 0$ of penalty parameter.

4 GLOBAL OPTIMALITY CONDITIONS (GOCs)

Before all, let prove that the cost function $\theta_\sigma(\cdot)$ of Problem $(\mathcal{P}_\sigma)$ is a d.c. function. Indeed, it can be readily seen that

$$\theta_\sigma(x) \overset{\triangle}{=} f_0(x) + \sigma \max\{0, f_i(x), \; i \in I\} + \sigma \sum_{i \in \mathcal{E}} |f_i(x)| = G_\sigma(x) - H_\sigma(x), \tag{4}$$

where

$$H_\sigma(x) := h_0(x) + \sigma \left[\sum_{i \in I} h_i(x) + \sum_{j \in \mathcal{E}} (g_j(x) + h_j(x)) \right], \tag{5}$$

$$G_\sigma(x) := \theta_\sigma(x) + H_\sigma(x) =$$
$$g_0(x) + \sigma \max \left\{ \sum_{j \in I} h_j(x); \left[g_j(x) + \sum_{\substack{j \neq i \\ j \in I}} h_j(x) \right], \; i \in I \right\} + 2\sigma \sum_{i \in \mathcal{E}} \max\{g_i(x); h_i(x)\}. \tag{6}$$

Obviously, $G_\sigma(\cdot)$ and $H_\sigma(\cdot)$ are both convex functions [7, 10], so that the function $\theta_\sigma(\cdot)$ is a d.c. function, as claimed. Besides, it is clear, that for a feasible (in $(\mathcal{P})$) point $z \in S$ we have $W(z) \overset{\triangle}{=} \max\{0, f_1(z), \ldots, f_m(z)\} + \sum_{i \in \mathcal{E}} |f_i(z)| = 0$, and therefore, for $\zeta := f_0(z)$, we obtain

$$\theta_\sigma(z) = f_0(z) + \sigma W(z) = f_0(z) = \zeta. \tag{7}$$

Theorem 4.1 *Let a point $z \in \mathcal{F}$ be a solution to Problem $(\mathcal{P})$ and $\sigma \geq \sigma_* > 0$, where $\sigma_* \geq 0$ is a threshold value of penalty parameter σ. Then, $\forall (y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ such that*

$$H_\sigma(y) = \beta - \zeta, \tag{8}$$

the following inequality holds

$$G_\sigma(x) - \beta \geq \langle H'_\sigma(y), x - y \rangle \quad \forall x \in S, \tag{9}$$

with any subgradient $H'_\sigma(y) \in \partial H_\sigma(y)$ of the function $H_\sigma(\cdot)$ at the point y.

Remark 1. In view of Theorem of Moreau-Rockafellar and due to (5) we have

$$H'_\sigma(y) = h'_0(y) + \sigma \Big[\sum_{i \in I} h'_i(y) + \sum_{j \in \mathcal{E}} (g'_j(y) + h'_j(y)) \Big], \tag{10}$$

where $h'_i(y) \in \partial h_i(y)$, $i \in \{0\} \cup I \cup \mathcal{E}$, $g'_j(y) \in \partial g_j(y)$, $j \in \mathcal{E}$, are some subgradients of the corresponding functions at the point y in the classical (convex) sense.

Remark 2. It is not difficult to note that Theorem 4.1 reduces the solution of the nonconvex Problem $(\mathcal{P}_\sigma)$ to an investigation of the family of the convex (linearized) problems

$$(\mathcal{P}_\sigma L(y)): \quad \Phi_{\sigma y}(x) := G_\sigma(x) - \langle H'_\sigma(y), x \rangle \downarrow \min_x, \quad x \in S,$$

depending on the parameters $(y, \beta) \in I\!\!R^n \times I\!\!R$ which fulfill the equation (8).

Remark 3. Suppose, we found a triple (y, β, u) and a subgradient $H'_\sigma(y) \in \partial H_\sigma(y)$, $(y, \beta) \in I\!\!R^n \times I\!\!R$, $H_\sigma(y) = \beta - \zeta$, $u \in S$, $H'_\sigma(y)$ has the form (10), such that the principal inequality (9) is violated, i.e. $0 > G_\sigma(u) - \beta - \langle H'_\sigma(y), u - y \rangle$.

Then, using the equation (8) and the convexity of the function $H_\sigma(\cdot)$, we derive

$$0 > G_\sigma(u) - \beta - H_\sigma(u) + H_\sigma(y) = \theta_\sigma(u) - \zeta = \theta_\sigma(u) - \theta_\sigma(z),$$

or, $\theta_\sigma(z) > \theta_\sigma(u)$, $z \in \mathcal{F}$, $u \in S$. Hence, the point z can not be a solution to $(\mathcal{P}_\sigma)$.

Hence, the conditions (8)–(9) of Theorem 4.1 possess the classical constructive (algorithmic) property (once the conditions are violated, one can find a feasible vector which is better than the point under investigation).

Suppose that a feasible (in Problem $(\mathcal{P})$) point z satisfies the conditions (8)–(9) of Theorem 4.1. Then, we set $y := z$ in (8)–(9). Further we derive that $\beta := H_\sigma(z) + \zeta = G_\sigma(z)$.

Therefore, from (9) it follows the validity of the inequality

$$G_\sigma(x) - G_\sigma(z) \geq \langle H'_\sigma(z), x - z \rangle \quad \forall x \in S,$$

for any subgradient $H'_\sigma(z) \in \partial H_\sigma(z)$ (see (10)) of $H_\sigma(\cdot)$ at z. It implies that the point z (satisfying (8)–(9)) is a solution to the linearized convex problem as follows

$$(\mathcal{P}_\sigma L(z)): \quad G_\sigma(x) - \langle H'_\sigma(z), x \rangle \downarrow \min_x, \quad x \in S,$$

for any subgradient $H'_\sigma(z)$ of $H_\sigma(\cdot)$ at z.

Since $(\mathcal{P}_\sigma L(z))$ is a convex problem, then the following inclusion is, as well-known, the necessary and sufficient optimality condition in order z to be a solution to $(\mathcal{P}_\sigma L(z))$:

$$0_n \in \partial G_\sigma(z) - \partial H_\sigma(z) + N(z \mid S). \tag{11}$$

Thus, the conditions (8)–(9) of Theorem 4.1 entail the well-known optimality conditions [1, 6, 9] for Problem $(\mathcal{P}_\sigma)$.

On the other hand, it is not difficult to see the drawback of (11), where the function $G_\sigma(\cdot)$ is not smooth (see (6)), while the entries of $(\mathcal{P})$ may be differentiable.

In order to overcome this shortcoming, let us apply Lemma 4.1 from [12] (see also [1, 9]), according to which Problem $(\mathcal{P}_\sigma L(z))$ amounts the following problem (see (6) and (10)):

$$\left. \begin{array}{l} g_0(x) - \langle H'_\sigma(z), x \rangle + \sigma\gamma + 2\sigma \sum\limits_{j \in \mathcal{E}} t_j \downarrow \min\limits_{(x, \gamma, t)}, \\[2mm] g_i(x) + \sum\limits_{\substack{j \in I \\ j \neq i}} h_j(x) \leq \gamma, \ \ i \in I, \quad \sum\limits_{j \in I} h_j(x) \leq \gamma, \ \ x \in S, \ \ \gamma \in I\!\!R, \\[2mm] g_k(x) \leq t_k, \ \ h_k(x) \leq t_k, \ \ k \in \mathcal{E}, \quad t = (t_{m+1}, \ldots, t_l)^\top \in I\!\!R^{l-m}. \end{array} \right\} \tag{12}$$

Clearly, the problem (12) is a convex optimization problem with the variables $(x, \gamma, t) \in I\!R^{n+1} \times I\!R^{l-m}$. Secondary, the problem (12) has only inequality constraints that distinguishes it from the original Problem $(\mathcal{P})$.

In addition, in the case of a smooth data of Problem $(\mathcal{P})$ the auxiliary problem (12) stays smooth. Besides, the number $(l - m + 1)$ of suplementary variables $(\gamma, t_{m+1}, \ldots, t_l)$ does not overcomes $(n + 1)$, otherwise Problem $(\mathcal{P})$ may be overdetermined or inconsistent.

Finally, it can be readily seen, that in (12) the Slater's condition holds.

Therefore, in the Lagrange function of the problem (12) we have $\mu_0 = 1$ and, besides, the corresponding KKT-conditions become necessary and sufficient for a tuple (z, γ_*, t_*) being a solution to the convex problem (12). In particular, the point z, satisfying (8)–(9), must be such a solution, since $z \in Sol(\mathcal{P}_\sigma L(z))$. For simplicity of presentation, suppose that $S = I\!R^n$.

Furthermore, suppose that all the functions $g_i(\cdot), \ h_i(\cdot), \ i \in \{0\} \cup I \cup \mathcal{E}$, are differentiable. Then the KKT system for $z \in Sol(\mathcal{P}_\sigma L(z))$, i.e. for $(z, \gamma_*, t_*) \in Sol$(12) with

$$\gamma_* = \max\left\{\sum_{i \in I} h_i(z); \left[g_i(z) + \sum_{\substack{j \in I \\ j \neq i}}^{j \neq i} h_j(z)\right], \ i \in I\right\}, \quad t_{k*} = \max\{g_k(z); h_k(z)]\}, \ k \in \mathcal{E}; \quad (13)$$

can be written as follows. There is a vector $(\mu_1, \ldots, \mu_m, \mu_{m+1}, \eta_{m+1}, \ldots, \eta_l, \nu_{m+1}, \ldots, \nu_l) \in I\!R_+^{m+1} \times I\!R_+^{2(l-m)}$, such that the complementary conditions hold:

$$\left.\begin{array}{l} \mu_i\left[g_i(z) + \sum_{\substack{j \in I \\ j \neq i}}^{j \neq i} h_j(z) - \gamma_*\right] = 0, \ \mu_i \geq 0, \ i \in I; \ \ \mu_{m+1}\left[\sum_{j \in I} h_j(z) - \gamma_*\right] = 0, \ \mu_{m+1} \geq 0, \\ \eta_k[g_k(z) - t_{k*}] = 0 = \nu_k[h_k(z) - t_{k*}], \ \ \eta_k \geq 0, \ \nu_k \geq 0, \ \ k \in \mathcal{E}. \end{array}\right\} \quad (14)$$

In addition, the following equations take place

$$\frac{\partial \mathcal{L}(z, \gamma_*, t_*)}{\partial \gamma} = \sigma - \mu_{m+1} - \sum_{i \in I} \mu_i = 0, \quad \text{i.e.} \quad \mu_{m+1} + \sum_{i \in I} \mu_i = \sigma; \quad (15)$$

$$\frac{\partial \mathcal{L}(z, \gamma_*, t_*)}{\partial t_k} = 2\sigma - \eta_k - \nu_k = 0, \ \ k \in \mathcal{E}, \quad \text{i.e.} \quad \eta_k + \nu_k = 2\sigma \ \forall k \in \mathcal{E}; \quad (16)$$

$$\begin{aligned} \frac{\partial \mathcal{L}(z, \gamma_*, t_*)}{\partial x}(z, \gamma_*) &= g_0'(z) - H_\sigma'(z) + \mu_{m+1}\sum_{i \in I} h_i'(z) + \sum_{i \in I} \mu_i[g_i'(z) + \sum_{\substack{j \in I \\ j \neq i}}^{j \neq i} h_j'(z)] + \\ &+ \sum_{k \in \mathcal{E}} \eta_k g_k'(z) + \sum_{k \in \mathcal{E}} \nu_k h_k'(z) = 0_n \in I\!R^n, \end{aligned} \quad (17)$$

for any collection of subgradients $h_i'(y) \in \partial h_i(y), \ i \in \{0\} \cup I \cup \mathcal{E}, \ g_j'(y) \in \partial g_j(y), \ j \in \mathcal{E}$, of the corresponding functions $h_i(\cdot), \ g_j(\cdot)$.

Further, thanks to (10) with $y = z$ and (16), $\nu_k = 2\sigma_k - \eta_k, \ k \in \mathcal{E}, \ f_i = g_i - h_i, \ i \in \{0\} \cup I \cup \mathcal{E}$, we obtain in the smooth case that

$$0_n = \nabla f_0(z) + \sum_{i \in I} \mu_i \nabla f_i(z) + \left(\sigma - \mu_{m+1} + \sum_{i \in I} \mu_i\right)\sum_{j \in I} \nabla h_j(z) + \sum_{k \in \mathcal{E}} (\eta_k - \sigma)[\nabla g_k(z) - \nabla h_k(z)].$$

The latter equation, in turn, due to (15), leads us to the final relation

$$0_n = \nabla f_0(z) + \sum_{i \in I} \mu_i \nabla f_i(z) + \sum_{k \in \mathcal{E}} (\eta_k - \sigma)\nabla f_k(z). \quad (18)$$

It means that, the principal equation of the KKT system for original Problem $(\mathcal{P})$ holds at the point z with Lagrange multipliers $\lambda_0 = 1, \ \lambda_i, \ i \in I \cup \mathcal{E}$, satisfying the following conditions

$$\lambda_i = \mu_i, \ i \in I, \quad \lambda_k = \eta_k - \sigma, \ k \in \mathcal{E}. \quad (19)$$

Moreover, it can be readily seen that in Problem $(\mathcal{P})$ the complementarity conditions with λ_i, $i \in I$, are also fulfilled:

$$\mu_i f_i(z) = 0, \ i \in I. \tag{20}$$

Hence, we just proved rather unexpected and unpredictable result.

Proposition 4.2 *Let a feasible in Problem $(\mathcal{P})$ point z satisfy the conditions (8)–(9) of Theorem 4.1. Then, z is a KKT point in $(\mathcal{P})$ (fulfills (18) and (20)) with the Lagrange multipliers $\lambda_0 = 1$, λ_i, $i \in I \cup \mathcal{E}$, satisfying (15), (16), (19), where the Lagrange multipliers μ_i, $i \in I$, η_k, ν_k, $k \in \mathcal{E}$, are provided by the auxiliary convex problem (12).#*

In addition, we have now the supplementary information about relations among the vectors $\lambda = (\lambda_1, \ldots, \lambda_l)^T \in \mathbb{R}^l$, $(\mu, \eta, \nu) \in \mathbb{R}^{m+1}_+ \times \mathbb{R}^{2(l-m)}$ and the penalty parameter $\sigma > 0$:

$$\sum_{i \in I} \lambda_i \leq \mu_{m+1} + \sum_{i \in I} \mu_i = \sigma, \quad \eta_k + \nu_k = 2\sigma \ \forall k \in \mathcal{E}. \tag{21}$$

Nevertheless, the natural question arises on whether it is possible to find a triple $(y, \beta, u) \in \mathbb{R}^{2n+1}$, satisfying (8) and which violates the inequality (9).

Theorem 4.3 *Assume, that a feasible in Problem $(\mathcal{P})$ point z is not an ε-solution to $(\mathcal{P})$, i.e.*

$$\inf(f_0, \mathcal{F}) + \varepsilon = \mathcal{V}(\mathcal{P}) + \varepsilon < \zeta := f_0(z). \tag{22}$$

In addition, let a vector $v \in \mathbb{R}^n$ satisfy the following inequality

$$(\mathcal{H}): \qquad \theta_\sigma(v) > \zeta - \varepsilon \tag{23}$$

Then, for any penalty parameter $\sigma > 0$ one can find a tuple (y, β, u), $(y, \beta) = (y(\sigma), \beta(\sigma)) \in \mathbb{R}^{n+1}$, $u \in \mathcal{F}$, such that for any subgradient $H'_\sigma(z) \in \partial H_\sigma(z)$ of the function $H_\sigma(\cdot)$ at y the following conditions take place

$$\left. \begin{array}{ll} (a) \ H_\sigma(y) = \beta - \zeta + \varepsilon; & (b) \ G_\sigma(y) \leq \beta, \\ (c) \ G_\sigma(u) - \beta < \langle H'_\sigma(y), u - y \rangle. \end{array} \right\} \tag{24}$$

5 CONCLUSION

We considered above the optimization Problem $(\mathcal{P})$ with the cost function and inequality and equality constraints given by d.c. functions. Introducing the penalty function we reduced $(\mathcal{P})$ to Problem $(\mathcal{P}_\sigma)$ without constraints and represented the cost function of penalized problem in d.c. form (4)–(6). Furthermore we developed necessary Global Optimality Conditions (GOC) (8)–(9) for $(\mathcal{P})$ and established its relations with the classical Optimality Conditions (the KKT theorem). Finally, thanks the constructive property of GOCs, we presented Theorem 4.3 which asserts the existence of a tuple (y, β, u) violating the GOCs in the case when the point z in question is not an ε-solution to Problem $(\mathcal{P})$. It makes possible constructions of numerical (computational) procedures allowing to escape local pit falls in $(\mathcal{P})$.

6 Acknowledgments

References

[1] Bonnans, J.-F., Gilbert, J.C., Lemaréchal, C. and Sagastizábal, C.A. (2006) Numerical Optimization: Theoretical and Practical Aspects. Berlin, Heidelberg: Springer-Verlag.

[2] Burke, J. (1991) An exact penalization viewpoint of constrained optimization. SIAM J. Control Optim., 29: 968–998.

[3] Byrd, R., Lopez-Calva, G. and Nocedal, J. (2012) A line search exact penalty method using steering rules. Math. Program., Ser. A, 133: 39–73.

[4] Clarke, F.H. (1983) Optimization and Nonsmooth Analysis. New York: Wiley-Interscience.

[5] Di Pillo, G., Lucidi, S. and Rinaldi, F. (2015) A derivative-free algorithm for constrained global optimization based on exact penalty functions. J. Optim. Theory Appl., 164: 862–882.

[6] Floudas, C.A. and Pardalos, P.M. (eds.) (2004) Frontiers in Global Optimization. Dordrecht: Kluwer Academic Publishers.

[7] Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993) Convex Analysis and Minimization Algorithms. Berlin: Springer-Verlag.

[8] Dontchev, A.L. and Rockafellar, R.T. (2014) Implicit Function and Solution Mappings. New York: Springer.

[9] Nocedal, J. and Wright, S.J. (2006) Numerical Optimization. New York: Springer.

[10] Rockafellar, R.T. (1970) Convex Analysis. Princeton: Princeton University Press.

[11] Strekalovsky, A.S. (2014) On solving optimization problems with hidden nonconvex structures. In: Rassias, T.M., Floudas, C.A., Butenko, S. (Eds.). Optimization in Science and Engineering (pp. 465–502). New York: Springer.

[12] Strekalovsky, A.S. (2017) Global Optimality Conditions in Nonconvex Optimization. J. Optim. Theory Appl., 173: 770–792.

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 7:
Multiple Criteria
Decision Making

COMPLEMENTARY USAGE OF MULTI-CRITERIA DECISION MAKING AND SYSTEM DYNAMICS: CASE STUDY OF HUMAN RESOURCE MANAGEMENT

Vesna Čančer
University of Maribor, Faculty of Economics and Business
Razlagova 14, 2000 Maribor, Slovenia
E-mail: vesna.cancer@um.si

Mirjana Pejić Bach, Jovana Zoroja
University of Zagreb, Faculty of Economics and Business
Trg J. F. Kennedy 6, 10000 Zagreb, Croatia
E-mails: mpejic@efzg.hr, jzoroja@efzg.hr

Abstract: System dynamic models can help decision makers in enhancing understanding of system behavior over time. However, previous research has demonstrated that model behavior is in number of cases contradictory. Therefore, the evaluation of the consequences of different policies showed by these models should be supported by multi-criteria decision making. The goal of this paper is to explore the complementary usage of complete system dynamics and multi-criteria decision making. A real-life example of the selection of the most appropriate human resource management policy in one market-research company is developed in order to demonstrate the usage of multi-criteria decision making for the purpose of the evaluation of different policies identified by system dynamics.

Keywords: human resource management, multi-criteria decision making, simulation, strategy, system dynamics

1 INTRODUCTION

System dynamics (SD) models are frequently used for the purpose of explaining the dynamics of complex systems and their behavior in different fields, such as industrial and social systems [7]. Multi-criteria decision making (MCDM) methods that have already turned out to be very applicable in business practice can be used to complement intuition and practical experience in solving complex problems [2, 3]. This paper highlights the capacity of the complementary usage of SD and MCDM methods when dealing with complex problems. The goal of this paper is to explore the possibilities how to employ MCDM as the completion of SD. The paper thus deals with multimethodology [5]. It presents several examples of SD and MCDM in economics and business, showing the mutual assistance of system dynamics and multi-criteria decision tools and approaches.

2 MCDM METHODS COMPLEMENTARY TO SD

SD models are frequently developed and used to represent, analyze and explain the dynamics of complex systems. The dynamics of behavior of a system is defined by its structure and the interactions of its parts. The main goal of SD is to understand how this behavior is produced, and use this understanding to predict the consequences over time of policy changes on the system. Almost four decades ago, Gardiner and Ford [4] emphasized that the point of SD was on developing models that show consequences, not on formally evaluating these consequences. Formal evaluation of the consequences showed by SD models can be supported by MCDM. It has been widely recognized that MCDM methods can help decision makers learn about the problems they face, and consequently make better-informed and justifiable choices. Santos et al. [10] concluded that SD and MCDM can play a major role in detailed analysis of the structure

of the problem under study and the consideration of trade-offs (to understand the causes of poor performance and determine the proper action plan for performance improvement).

Araz [1] pointed out that it is not enough to use merely simulation models when incorporating policy decision makers' preferences into decision-making processes. Araz [1] developed a framework for public health preparedness exercise design that simulates disease spread with selected intervention strategies. The framework integrates an AHP model for MCDM with a simulation model to evaluate policy decision options based on criteria determined by decision makers. SD was used to model different policy interventions. MCDM was used to express preferences to alternatives (interventions strategies, policies) and judgments on criteria's importance.

Santos et al. [10] argue that the integration between SD and multiple criteria analysis can address some issues which require further study if measurement systems are to be supporting the decision making process, and contribute to improve organizational performance. These issues are the identification of key performance factors (or performance drivers), a better understanding of the interrelationships and the consideration of trade-offs between performance measured, the dynamism of organizations and the dynamism of measurement systems. The integration of SD and MCDM can bring new insights to inform and support performance measurement and management. [10]

Pruyt [8] looked at the combination of SD, multiple criteria decision analysis (MCDA) and ethics to support strategy selection in case of dynamically complex multi-dimensional societal issues, with special attention paid to the capacity of the multi-methodology. It was suggested that SD could be used to simulate the multi-dimensional behavior, and MCDA could then be used to describe, evaluate and choose between the strategies simulated with the SD models.

To summarize, MCDM can be used to evaluate policy decision options obtained by SD. We propose the use of the frame procedure of MCDM for the group of methods based on assigning weights [3]: problem definition and structuring, measuring local alternatives' values (by using value functions and pairwise comparisons), criteria weighting (by using the methods based on ordinal, interval and ratio scale), synthesis, ranking and sensitivity analysis.

3 EXAMPLE CASE OF THE USE OF SD AND MCDM

3.1 Description of the case study

Technological advancement on growing markets requires highly educated employees. The firms that are believed to be the drivers of the new industrial revolution, that are expected to increase productivity, shift economics, and foster industrial growth, are also expected to modify the profile of the workforce [9]. The main resource in these firms is highly qualified employees. Firms employ expert professionals; if there are not enough experts, they can employ trainees. The most important variables that influence employee satisfaction are workload (number of projects/employee) and salary. One of the main reasons for decreased productivity of the current employees and thus for employing new staff is stress of the current employees. Stress also decreases the quality of work. When developing this model, demand and business success should also be taken into consideration. Sectors that should be taken into consideration for managing highly qualified employees are therefore employee lifecycle, employee satisfaction, quality of service, demand and business success. The SD model has been developed for identifying efficient human resource management (HRM) policies, which consists of 5 segments: (1) Employee Lifecycle, (2) Employee Satisfaction, (3) Quality of Service, (4) Demand and (5) Business Success. The model was developed with Vensim; software specialized for the system dynamics. The aforementioned system dynamics model is more detailed explained and presented in the article of Pejic-Bach et al. [6].

For the purpose of the demonstration, the first sector is presented in this paper. More in-depth description of the sectors is available in [6].

Sector 1: Employee Lifecycle. Required work in company can be done only by expert professionals who have necessary knowledge and expertise. In a case, when there are not enough experts in the company, they employ trainees. When there is imbalance between the necessary and the current number of employees, new trainees are employed and educated by experts. Therefore, the total number of necessary experts depends on the workload and the number of employed trainees as well as on the number of projects for which the company's manager thinks one expert should be responsible on a monthly basis.

According to company's manager, professional who have to finished 15 projects during one month will have less experts to help him, then other who have to finish 10 projects during one month. The reason for that is that number of experts is decreased depending on the ratio of current and necessary number of experts. In a case of more experts that they are necessary, they are laid of which can be modelled using shorter length of employment. Situation of dissatisfied experts, who leave the company, can also be modelled using shorter length of employment.
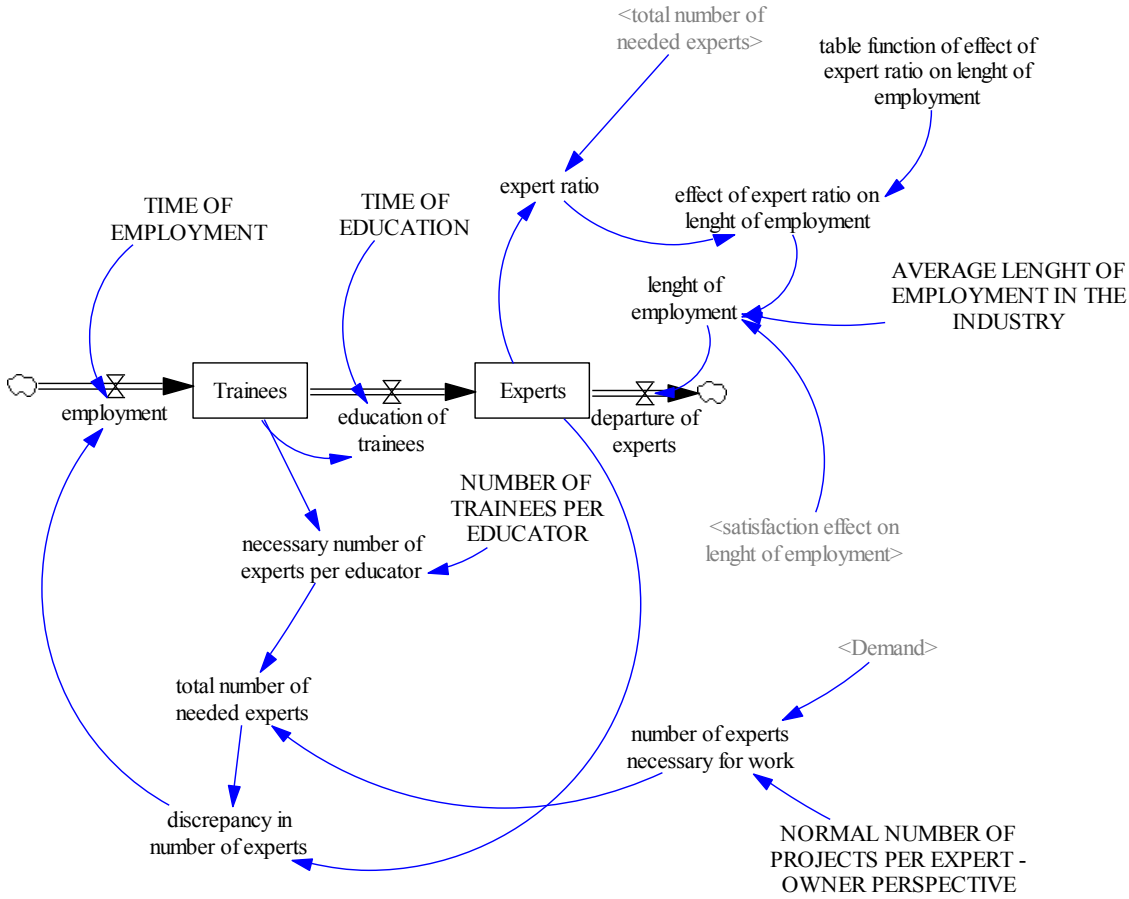
Figure 1: Flow chart of employment

The average length of employment of employees depends on two factors: (i) the normal length of employment and (ii) the effect of experts on length of employment. In addition, the effect of experts on length of employment also depends on two aspects: (i) the ratio of the current number of experts and (ii) the necessary number of experts. It can be concluded that when the current number of professionals is lower or equal to the necessary, the average length of employment of professionals is equal to the normal length of employment. In the case of more

experts than those who are needed, needles experts are fired and the length of employment is shorter than normal. Model for Sector 1 named *Employee Lifecycle* is shown in Figure 1.

3.2 Simulation experiments

We made two simulations: for the one-year period and for the five-year period. The results of the HRM policies' simulation regarding profit, perception of quality and employee satisfaction in different scenarios (normal working conditions, stressful working conditions with an average salary and stressful working conditions with a high salary) are for both simulations (for the one-year and for the five-year period) presented in Table 1.

Table 1: The results of the HRM policies' simulation regarding profit, quality and satisfaction in different scenarios (for the one-year and for the five-year period)

	Profit			Perception of quality			Employee satisfaction		
	Initial	*1 year*	*5 years*	*Initial*	*1 year*	*5 years*	*Initial*	*1 year*	*5 years*
Scenario 1	75600	88870	234104	100	100	100	1	1.01	1.02
Scenario 2	122067	121739	41281	100	22	11	1.5	1.54	1.51
Scenario 3	95667	108702	161657	100	54	47	1.2	1.21	1.21

Note: Scenario 1: Normal working conditions, Scenario 2: Stressful working conditions with an average salary, Scenario 3: Stressful working conditions with a high salary; Measure of employee satisfaction is expressed as dissatisfaction: higher the value, lower the satisfaction

MCDM has been employed to evaluate HRM policies: normal working conditions, stressful working conditions with an average salary, and stressful working conditions with a high salary. The criteria for evaluating these policies are employee satisfaction, quality of service and profit. We followed the frame procedure for MCDM [3]. The goal (selection of the most appropriate HRM policy when managing highly qualified employees), the criteria and the alternatives (the above mentioned HRM policies) were structured in problem hierarchy.

The weights of criteria were determined by using the SWING method that is based on an interval scale (Table 2). Table 2 shows that the change from the worst to the best quality of work is considered the most important, thus 100 points were assigned to quality. With respect to this change importance, 20 points less, i.e. 80 points are given to the change from the lowest to the highest profit, and 50 points less (i.e. 50 points) are given to the change from the highest to the lowest satisfaction.

Table 2: Criteria weighting

Criterion	Points	Weight
Profit	80	0.348
Satisfaction	50	0.217
Quality	100	0.435

The data for entering MCDM process were obtained by SD model by using the SD results of the simulation for the one-year and for the five-year period (Table 1). With respect to profit, the HRM policies were measured by increasing linear value function. To obtain the greatest diversification between the alternatives, the lower bound is equal to the lowest datum, and the upper bound is equal to the highest datum in each simulation period (Table 3).

Criterion	Measuring local alternatives' values						
	One year				Five years		
Profit	Increasing value function, LB: 88870, UB: 121739				Increasing value function, LB: 41281, UB: 234104		
Satisfaction	Pair-wise comparisons:				Pair-wise comparisons:		
		A	B	C		A	B	C
	A	1	5	3	A	1	5	3
	B	0.2	1	0.5	B	0.2	1	0.5
	C	0.33	2	1	C	0.33	2	1
Quality	Pair-wise comparisons:				Pair-wise comparisons:		
		A	B	C		A	B	C
	A	1	5	2	A	1	9	2
	B	0.2	1	0.4	B	0.11	1	0.25
	C	0.5	2.5	1	C	0.5	4	1

The results in Table 4 show that with respect to profit, the 'stressful working conditions with an average salary' achieved the highest value and the 'normal working conditions' policy achieved the lowest value in the one-year period. The results with respect to profit for the five-year period, however, show that the 'normal working conditions' policy is the most appropriate policy and the 'stressful working conditions with an average salary' is the least appropriate policy (Table 4). With respect to employee satisfaction and quality of service, the HRM policies were measured by pairwise comparisons. Preferences to alternatives were expressed by using the AHP verbal scale and then transformed to numerical values (1–the alternatives are equally preffered, 3–the alternative is moderately more preffered than the compared one, 5–the alternative is strongly more preffered than the compared one, 7–the alternative is very strongly more preffered than the compared one, 9–the alternative is extremely more preffered than the compared one). Based on the SD results in Table 1 it can be evaluated that with respect to satisfaction, normal working conditions are strongly more preffered than stressful working conditions with an average salary, and moderately more preferred than stressful working conditions with a high salary (Table 3). The SD results showed that major changes in employee satisfaction occurred in the period from six to nine months of the first year. For this reason, the values of the HRM policies with respect to employee satisfaction after the one-year period are equal to the ones after the five-year period (Table 4).

Table 4: The HRM policies'values obtained by MCDM

	Value with respect to profit		Value with respect to employee satisfaction		Value with respect to quality of service		Aggregate value		Rank	
	One year	Five years	One year	Five years	One year	Five years	One year	Five years	One year	Five years
Scenario 1	0	1	0.648	0.648	0.588	0.626	0.397	0.761	2.	1.
Scenario 2	1	0	0.122	0.122	0.118	0.072	0.426	0.058	1.	3.
Scenario 3	0.603	0.624	0.230	0.230	0.294	0.301	0.388	0.398	3.	2.

Note: Scenario 1: Normal working conditions, Scenario 2: Stressful working conditions with an average salary, Scenario 3: Stressful working conditions with a high salary

With respect to quality of service, the values of the 'normal working conditions' policy and the 'stressful working conditions with a high salary' are slightly higher for the five-year period than for the one-year period (Table 4). With respect to multiple criteria, the 'stressful working conditions with an average salary' policy achieved the highest aggregate value in the one-year

period, and the 'normal working conditions' policy achieved the highest value in the five-year period (Table 4). The gradient sensitivity analysis has been performed to analyze the effects of changes in criteria's weights on alternatives' ranking. The results showed that ranking of alternatives is very sensitive to changes in criteria's weights in the one-year period, but it is not sensitive in the five-year period. The MCDM model results confirmed that – with respect to multiple criteria – normal conditions policy is the best policy in management of highly qualified human resources in the long run.

4 CONCLUSIONS

The case presented in this paper represents an approach to face with risks to humanity's future welfare, including ones that could be created by emerging technologies, e.g. overloading, stress, and unemployment. SD models can help decision makers in enhancing understanding of system behavior over time, and MCDM enables explicit evaluation of this behavior.

References

[1] Araz, O. M. 2013. Integrating complex system dynamics of pandemic influenza with a multi-criteria decision making model for evaluating public health strategies. Journal of System Science and System Engineering, 22(3): 319–339. DOI: 10.1007/s11518-013-5220-y

[2] Čančer, V. 2010. Systemic Thinking in Creative Problem Solving. In Exnarová, A., Pavlíček, A. (Eds.). Systémové přístupy '10 (pp. 22-29). Praha: Vysoká škola ekonomická v Praze, Nakladatelství Oeconomica.

[3] Čančer, V. 2012. Criteria weighting by using the 5Ws & H technique. Business Systems Research, 3(2): 41–48. DOI: 10.2478/V10305-012-0011-3

[4] Gardiner, F. 1980. Which policy run is best, and who says so?. In Legasto, A. A., Forrester, J. W., Lyneis, J. M. (Eds.). System Dynamics: TIMS Studies in the Management Sciences (14: 241–257). Amsterdam: North-Holland.

[5] Mingers, J., Gill, A. 1997. Multimethodology: The Theory and Practice of Combining Management Science Methodologies. Chichester, UK: Wiley.

[6] Pejic Bach, M., Knezevic, B., Strugar, I. 2006. Strategic Decision Making In Human Resource Management Based On System Dynamics Model. Proceedings on WSEAS International Conference on Transactions on Systems, Issue 1, Volume 5, January 2006.

[7] Pejić Bach, M., Zoroja, J., Vrankić, I. 2016. System Dynamics Approach to Teaching Supply and Demand: Preliminary Research. In 18th International Conference on Information Technology, Modeling and Computing (pp. 803–808). Barcelona, Spain: World Academy of Science, Engineering and Technology.

[8] Pruyt, E. 2006. System Dynamics and Decision-Making in the Context of Dynamically Complex Multi-Dimensional Societal Issues. In 24th International Conference of the System Dynamics Society (pp. 2767–2785). Nijmegen, The Netherlands: System Dynamics Society.

[9] Rüßmann, M., Lorenz, M., Gerbert, Ph., Waldner, M., Justus, J., Engel, P., Harnisch, M. 2015. Industry 4.0, The Future of Productivity and Growth in Manufacturing Industries. Boston, MA: BCG, The Boston Consulting Group. http://www.zvw.de/media.media.72e472fb-1698-4a15-8858-344351c8902f.original.pdf [Accessed 06/11/2016].

[10] Santos, S. P., Belton, V., Howick, S. 2001. Integrating System Dynamics and Multicriteria Analysis: Towards Organisational Learning for Performance Improvement. AHR. 1st International Workshop on Performance Measurement, 2001-05-24 - 2001-05-25. (Unpublished)

A COMBINED SOCIAL NETWORK ANALYSIS - ANALYTIC NETWORK PROCESS APPROACH TO EVALUATE SUSTAINABLE TOURIST STRATEGIES

Hannia Gonzalez-Urango
Ingenio (CSIC-UPV), Universitat Politècnica de València
46022 Valencia, Spain
E-mail: hangonur@doctor.upv.es

Mónica García-Melón
Ingenio (CSIC-UPV), Universitat Politècnica de València
46022 Valencia, Spain
E-mail: mgarciam@dpi.upv.es

Abstract: In this paper we present a methodology for the sustainable evaluation of strategic urban development projects in the city of Cartagena de Indias, the most important tourist destination of Colombia. The methodology is based on a combination of Social Network Analysis for stakeholder analysis and the multicriteria technique Analytic Network Process, which allows ranking of the tourist strategies according to the stakeholders. The decision model considers environmental, socio-cultural, sectorial, economic and political aspects. The aim is to provide answers and guide local decision makers towards the prioritization of strategies for the tourist sector with the participation of stakeholders since this prioritization process is key to the strategic planning of the city.

Keywords: stakeholder analysis, social network analysis (SNA), multicriteria decision making, analytic network process (ANP), Cartagena de Indias, sustainable evaluation.

1 INTRODUCTION AND LITERATURE REVIEW

Tourism is a large industry that is currently going through a period of great relevance. For the past decades, it has grown above average and has been characterized by immense innovativeness and great diversity[13]. According to UNWTO, these trend are expected to keep rising especially in emerging economic destinations, such as South America, thanks to the fact that it has some of the main emerging destinations in the sector, such as Colombia[20]. This trend of tourism growth comes with some drawbacks which include an increasing pressure on the territories [4]. The tourism sector can and is firmly committed to playing its part in the 2030 Sustainable Development Agenda. Promoting governments, the private sector, academia and the civil society are expected to work together in order to implement sustainable tourism activities with an emphasis on sustainable land use [20].

Colombia as an emerging destination and Cartagena de Indias as its most representative and important destination, cannot be left behind when it comes to achieve this aim. This city has to prepare and to adapt public policies and managerial strategies to face new challenges and opportunities both for the tourist industry and for the destinations.For several years, the city has been doing long-term planning, in which the city defines its vocation and focuses its efforts on achieving productive transformation and increasing its competitiveness through recognized economic development potentials, such as tourism [7]. The current planning has not yet evolved to deal with upcoming challenges; which is affecting the expansion and placement of new development projects. Environmental perception and attitude of stakeholders generate debates, controversy and contradictions among economic sectors and groups. To solve this problem we propose to evaluate different strategic plans that the city has currently in mind considering sustainable criteria together with integrative and participative approach supported by technical and scientific knowledge [16]. This is a decision making problem that should be approached from the multi-criteria analysis perspective, with the participation of different stakeholders.

Sustainable tourism is regularly linked to the preservation of ecosystems, the promotion of human welfare, inter- and intra-generational equity, and public participation in decision-making [6]. Different sustainable development planning experiences applied to tourism analyze and suggest the implementation of participatory processes. For example, Petrosillo [15] shows that environmental management problems could be more easily solved by applying methodologies based on a participatory approach and Bonzanigo [5] present a participatory decision support process for the analysis of adaptation strategies for local development of an Alpine tourism destination. Additionally, coastal space planning must consider methodologies that take holistic approaches and complexity into account [3]. Complex approaches to linear analysis are preferred, as well as multidisciplinary and multisectoral approaches [19]. Domínguez-Tejo et al. [8] in their work consider the existence of an important knowledge gap when trying to improve the integrated management approach to coastal resources planning. In order to fill this gap, in a previous study conducted by the same authors, an Analytic Network Process (ANP) model was proposed for prioritizing local development strategies and reach a consensus among two sectors [12]. The result of that study suggests that more members have to be involved as key decision makers. Therefore, according to these results, a new approach is proposed here in which the selection of stakeholders was made through a Social Network Analysis (SNA).

The knowledge gained from analyzing the social networks of stakeholders can be used to select some of stakeholders for the participation in planning development initiatives [17]. Likewise ANP is a multicriteria technique (MCDM) thas has been introduced for several authors for sustainability assessment [11]. In this paper we combined both techniques to intend to prove that combined SNA- ANP are an appropriate tool to reach a consensus among different stakeholders on the essential issues of the territorial development. We use this approach to identify the most central stakeholders and evaluate and prioritize sustainable tourism strategies in order to improve the touristic offer of the city.

2 METHODOLOGY
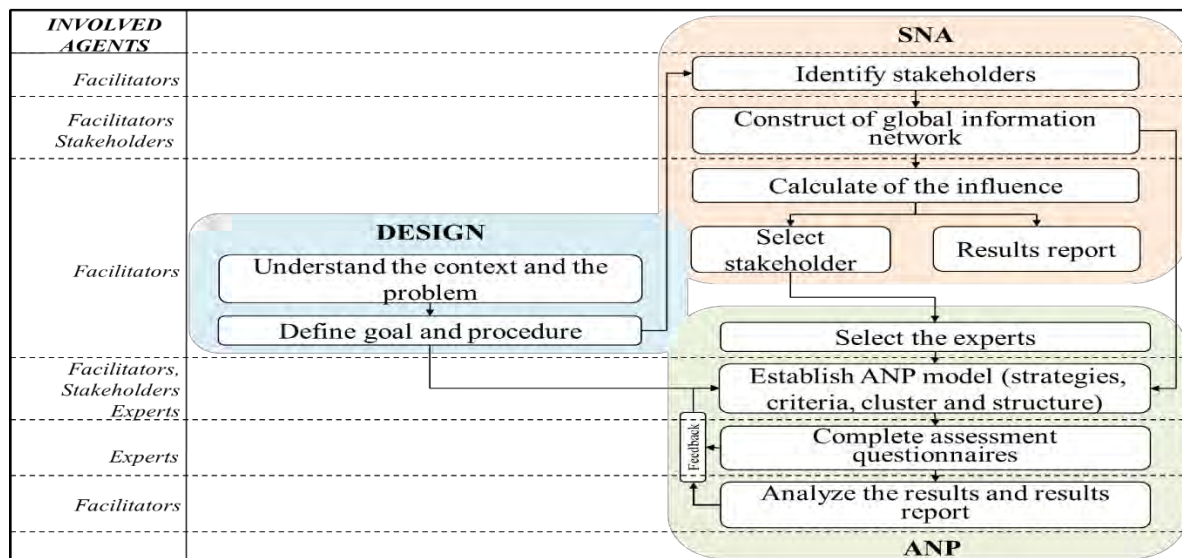
Figure 1 : Methodology proposed

3 APPLYING SNA

According to some Cartagena's Tourist Offices' documents and to the National Colombian Tourism Register, the main groups of the stakeholders were identified. The institutions and

organizations involved in the problem were classified in six groups: G1. Tourist Services Providers, G2. Destination management organizations, G3. Transport concessions, G4. Academy, G5. Civil society and G6. Tourists. They were classified in two categories. According to the groups mentioned above or according to the sector (public, private or public/private mix). To construct the global information network, the information flows were analyzed in both directions. We obtained a non-symmetrical matrix with a directed network of actors (Figure 2).
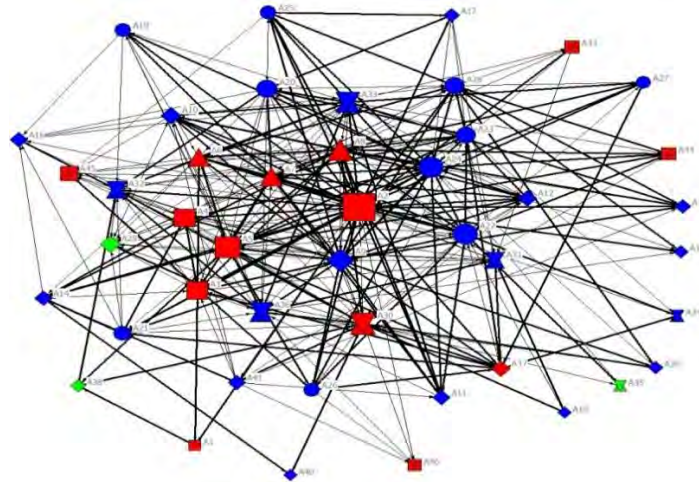
Figure 2 : Graphs showing social network of stakeholders.

We have chosen the nodes´centrality as the most appropriate SNA indicators to assess the influence of the stakeholders [17]. The centrality indices of the actors are calculated in the software program (UCINET©). Table 1 shows the results for the most relevant actors in the information network.

Table 1: Centrality scores for stakeholders.

ID	Institution	Categories		Indicators				
		Sector	Group	Closeness		Degree		Betweeness
				Out	In	Out	In	
A2	Local Tourism office	Public	G2	52	59	38	34	387
A5	Municipality Tourism office	Public	G2	67	71	23	22	105
A8	National Promotion Tourism office	Public	G2	64	72	26	23	141
A9	Local Chamber of Commerce	Mix	G2	61	67	29	26	94
A30	University "A"	Public	G4	76	69	14	24	57

The analysis of the results shows that the actors of groups 2 and 4 do have information and also have power to influence the opinion of the rest. In general the local tourist sector has strong ties, which means that it is a strong sector and able to respond quickly and effectively. Both public and private institutions play an important role in connecting and communicating.

4 THE DECISION PROCESS AND MODELING OF THE PARTICIPATIVE DECISION MAKING PROCESS WITH ANP

This second part is still an ongoing stage. This part aims to support decision makers to evaluate and prioritize sustainable tourism strategies. For the expert selection, the five main influent stakeholders have been considered. In order to define development strategies, a review of local and national plans and programs designed to strengthen the sector was performed. Three proposals (alternatives) were selected, aimed at developing new urban projects in the city. Prioritizing these proposals should allow channeling most of this sector's development and

resources, and should help improving the touristic offer of the city. The selected alternatives are:

- Alternative 1: A1. Tourist complex. Develop of a tourism complex located in insular territory.
- Alternative 2: A2. Tourist boulevard. Develop a coastal protection project to improve the connection and spaces between the most tourist neighbourghoods.
- Alternative 3: A3. Waterborne transport system. Develop a network public transport system using the water resources available around the city.

Criteria which could influence the sustainable evaluation of the proposed alternatives were identified. It was necessary to make sure that these criteria could be grouped, that they were relevant, not redundant and easy to understand for the different actors. The final list of 25 criteria grouped in five evaluation clusters was defined on the basis of a bibliographic review and with the assistance of some of the experts.

After the identification of the model elements, influences among them were determined using a relationship matrix with the help of the experts. The proposed model is illustrated by the network shown in Figure 3.. The bidirectional arrows indicate influences between clusters in both directions.
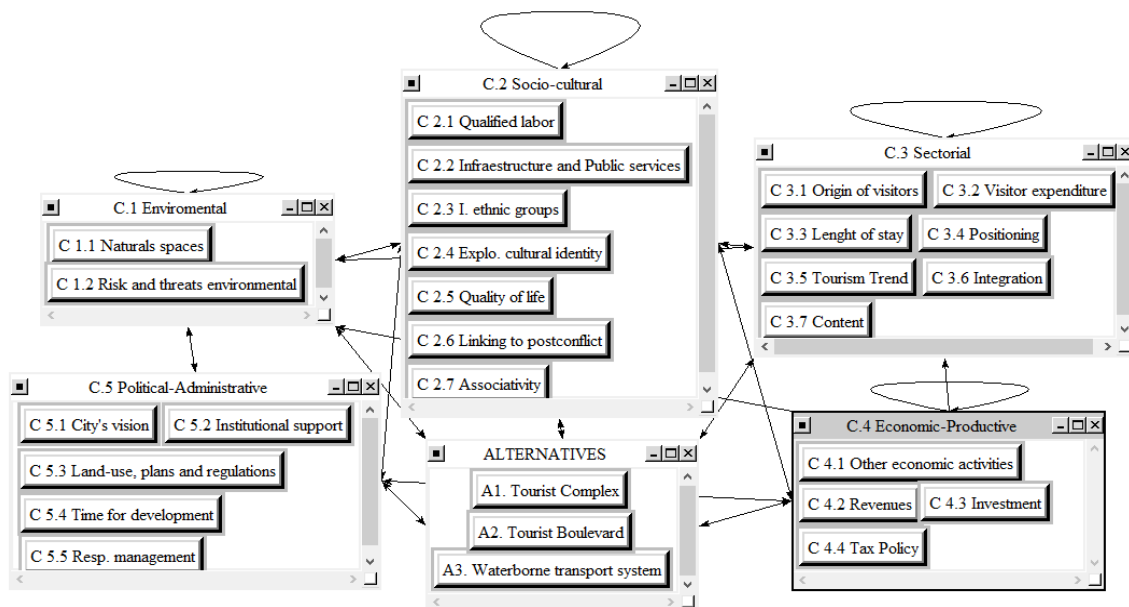
Figure 3: ANP network model of the case study.

A questionnaire was designed with the aim of determining a compliance index for each alternative with regard to all considered criteria. This information was collected among the experts through a questionnaire designed to allow comparisons between pairs of elements. All the calculations were performed using the Superdecision© v.2.0.8. software. Once experts have finished all pairwise comparisons, judgement aggregation was performed using the geometric mean in order to obtain a global judgement [18]. So far, we have one of the five complete questionnaires, the one answered by an Institute of Publics, Regionals and Governments Policies at University "A". We are currently collecting the rest of them. Thus, the results show only one expert's judgements.

The final limit matrix shows the priority obtained for each criterion, a non-dimensional value that can be considered as its relative importance. Preliminary results show (Figure 4) that altogether, the most valued cluster were the environmental (0,25), the political (0,23) and the sociocultural (0,22) ones. The less valued –but not far- are the sectorial (0,16) and the economic (0,15) ones. In concordance with the results by cluster, results for each criteria show that the

most important is C1.1 Use of heritage and naturals spaces (0,15), followed by C1.2 Risk and threats environmental (0,92) and C5.5 Responsible and sustainable management (0,088). The least valued are C2.6 Linking to postconflict (0,003) and C3.1 Origin of visitors (0,003). According to this result, the expert shows a greater interest for the environmental effects and their management.
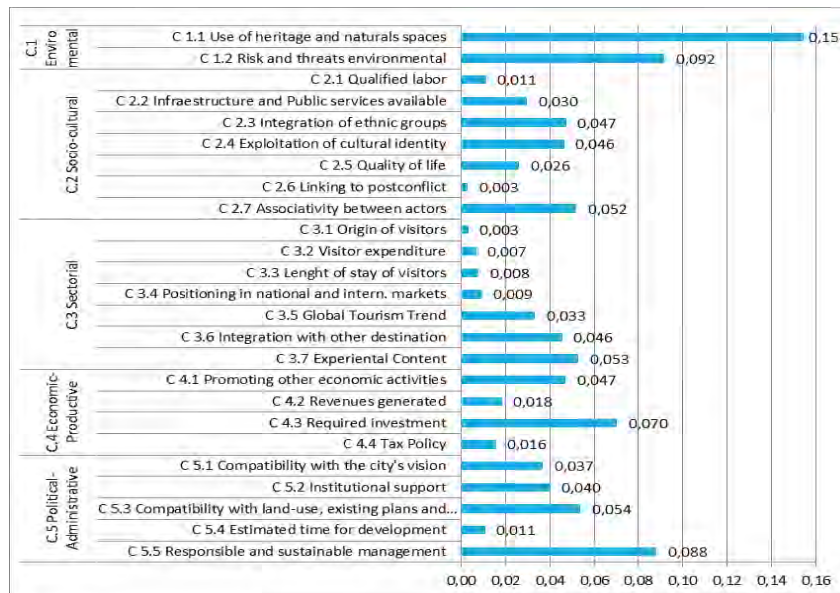
Figure 4: Results for the criteria.

Priorities obtained for the alternatives can be considered as their Preference Index, so the higher this index value, the better the proposal prioritization will be. According to this expert the best strategy to be implemented to improve tourist offer in the city of Cartagena is the *A1Tourist Complex* (44%), followed by *A3 Waterborne transport system* (34%) and *A2 Tourist Boulevard* (22%).

CONCLUSIONS

In this paper we have proposed a methodology that combines both the stakeholder analysis (through SNA) and the participatory sustainable evaluation (through ANP). Regarding the use of these techniques. Social Network Analysis SNA was an appropriate tool to know how the relationship between stakeholders of tourist sector in Cartagena is and who are the most influent and central ones. Concerning the use of ANP as a tool for prioritization, other productive sectors or local development plans of the City of Cartagena could be evaluate using similar techniques. Besides, it is important to remark that procedure could consider sensitivity analysis (by slowly modifying the weight of each criterion in the limit matrix). It would show if the final evaluation and ranking are reliable.

So far, we have not developed the full model, so is not possible to reach conclusions for the complete results only those obtained by our first expert. Thus, our immediate challenges to get the rest of the complete questionnaires, in the next month and analyse them thoroughly. In the same way we hope to compare the answers of some central agents with other less influential ones, following the recommendation of Prell [17].

References

[1] Aminu, M., Matori, A.-N., Yusof, K.W. and Zainol, R.B. 2013. A Framework for sustainable tourism planning in johor Ramsar sites, Malaysia: A geographic information system (GIS) based analytic network process (ANP) approach. Research Journal of Applied Sciences, Engineering and Technology. 6, 3 (2013), 417–422.

[2] Apostolopoulou, E., Drakou, E.G. and Pediaditi, K. 2012. Participation in the management of Greek Natura 2000 sites: Evidence from a cross-level analysis. Journal of Environmental Management. 113, (2012), 308–318.

[3] Berkes, F. 2006. From community-based resource management to complex systems: The scale issue and marine commons. Ecology and Society. 11, 1 (2006).

[4] Berzina, I., Grizane, T. and Jurgelane, I. 2015. The tourism service consumption model for the sustainability of the special protection areas. Procedia Computer Science. 43, C (2015), 62–68.

[5] Bonzanigo, L., Giupponi, C. and Balbi, S. 2016. Sustainable tourism planning and climate change adaptation in the Alps: a case study of winter tourism in mountain communities in the Dolomites. Journal of Sustainable Tourism. 24, 4 (2016), 637–652.

[6] Bramwell, B. 2015. Theoretical activity in sustainable tourism research. Annals of Tourism Research. 54, (2015), 204–218.

[7] Comisión Regional de Competitividad de Cartagena y Bolívar 2010. Plan Regional de Competitividad Cartagena y Bolívar 2008 - 2032.

[8] Domínguez-Tejo, E., Metternicht, G., Johnston, E. and Hedge, L. 2016. Marine Spatial Planning advancing the Ecosystem-Based Approach to coastal zone management: A review. Marine Policy. 72, (2016), 115–130.

[9] Eagles, P., McColl Stephen and Haynes, C. 2002. Sustainable Tourism in Protected Areas: Guidelines for Planning and Management.

[10] García-Melón, M., Gómez-Navarro, T. and Acuña-Dutra, S. 2012. A combined ANP-delphi approach to evaluate sustainable tourism. 34, (2012), 41–50.

[11] Ginevičius, R. and Podvezko, V. 2009. Evaluating the changes in economic and social development of Lithuanian counties by multiple criteria methods. Baltic Journal on Sustainability. 15, 3 (2009), 418–436.

[12] Gonzalez-Urango, H. and García-Melón, M. 2017. A multicriteria model to evaluate strategic plans for the nautical and naval industry in Cartagena de Indias, Colombia. Sustainability (Switzerland). 9, 4 (2017).

[13] Hjalager, A.M. 2010. A review of innovation research in tourism. Tourism Management. 31, 1 (2010), 1–12.

[14] Jeong, J.S., García-Moruno, L., Hernández-Blanco, J. and Jaraíz-Cabanillas, F.J. 2014. An operational method to supporting siting decisions for sustainable rural second home planning in ecotourism sites. Land Use Policy. 41, (Sep. 2014), 550–560.

[15] Petrosillo, I., Valente, D., Zaccarelli, N. and Zurlini, G. 2009. Managing tourist harbors: Are managers aware of the real environmental risks? Marine Pollution Bulletin. 58, 10 (2009), 1454–1461.

[16] Le Pira, M., Ignaccolo, M., Inturri, G., Pluchino, A. and Rapisarda, A. 2016. Modelling stakeholder participation in transport planning. Case Studies on Transport Policy. 4, 3 (2016), 230–238.

[17] Prell, C., Hubacek, K. and Reed, M. 2009. Stakeholder Analysis and Social Network Analysis in Natural Resource Management. Society & Natural Resources. 22, 6 (2009), 501–518.

[18] Saaty, T.L. 2001. The Analytic Network Process: Decision Making with Dependence and Feedback. RWS Publications.

[19] Sierra-Correa, P.C. and Cantera, J.R. 2015. Ecosystem-based adaptation for improving coastal planning for sea-level rise: A systematic review for mangrove coasts. Marine Policy. 51, (2015), 385–393.

[20] World Tourism Organization UNWTO 2017. UNWTO Annual Report 2016. UNWTO.

CONSENSUS MODEL FOR GROUP DECISION PROBLEMS WITH INTERVAL WEIGHTS

Petra Grošelj, Lidija Zadnik Stirn
University of Ljubljana, Biotechnical Faculty
Jamnikarjeva 101, 1000 Ljubljana, Slovenia
petra.groselj@bf.uni-lj.si, lidija.zadnik@bf.uni-lj.si

Abstract: The decision making process often involves multiple decision makers. The aggregation of their opinions into a joint decision is an important topic in the group decision making. The aim of this paper is to provide a new consensus reaching model for deriving group interval weights from individual interval weights. The weights of importance of decision makers, based on the closeness of the individual weights and the widths of the individual intervals, are also taken into account. To demonstrate the effectiveness of the presented model a numerical example is provided. A comparison of the derived results with the results obtained by the method using arithmetic mean is conferred, too.

Keywords: interval weights, consensus model, group decision making, multi-criteria decision making

1 INTRODUCTION

The significance of the research of the group decision making is increasing because today the decisions are very complex and can hardly depend on a single decision maker (DM). Knowledge and experiences of one DM are limited, while a group of DMs can contribute a variety of knowledge, perspectives, views and experiences. Group decision making is a process of aggregating individual judgments, weights or opinions into a common decision. The consensual decision is the most desired one, although it is hard to be achieved. In many multi-criteria decision methods such as SWING, SMART, SMARTER, analytic hierarchy process (AHP), TOPSIS, MAUT, etc. the final result is represented by the weights of criteria and/or alternatives. Due to the vagueness and uncertainty that exist as part of the decision problem and/or the DMs' opinions, the weights could be better expressed by the intervals than by the exact values [1]. Another key issue in group decision making is the importance of DMs. They could be all equally important. Nevertheless, the leader of the decision making process or the participants themselves think that DMs are not equally important. Thus, their weights of importance should also be taken into account. However, there are no widely accepted approaches how to determine the weights of importance of DMs and how to aggregate individual interval weights into group interval weights.

The aim of this paper is to present a consensus model that provides common interval weights from the individual interval weights based on the respect that DMs hold for the other participants of decision making process.

1.1 Consensus reaching models

The consensus group weights can be achieved through negotiations and informal consensus building process or by the employment of the formal mathematical methods. While the informal decision process can be very useful in the problem formulation stage, the structured techniques can be suitable in the stage of assigning and aggregating weights, because they are more transparent and less subjugated to the manipulations of the individuals in the group [6]. Formal consensus reaching models have been widely studied regarding different types of preference relations: multi-criteria group decision making problems [2, 6, 7], fuzzy preference relations [1, 4], linguistic preference relations [8], ordinal preferences [3]. However, the consensus reaching model for aggregating interval weights does not exist.

1.2   Importance of DMs

The importance of DMs is usually implemented through weights of importance that are determined by the leader of decision making process, what can be a prejudiced method [2]. The more appropriate and less biased method requires from each DM_k to assign the weights of importance to all group members (including himself). Since diversity of opinions is always present within a group of DMs, DM_k has similar opinion with several DMs and diversified opinion with the other DMs. Therefore, the weight of importance that DM_k assigns to the DM_j should express the difference between DM_k's weights and DM_j's weights. Consequently, these weights of importance should satisfy several properties [6, 9]. So, each DM provides the highest weight of importance to himself, and gives higher weight of importance to the DMs with similar weights and lower weight of importance to DMs with dissimilar weights.

Furthermore, the accuracy of the interval weights depends also on their width. Wider interval weights are less precise comparing to the narrower intervals. Therefore, wider individual weights should have less impact on the group weights.

2   METHODS

Let $x = [x_l, x_u]$ be an interval number. A few terms can be summarized as follows:

The width of the interval x: $d(x) = x_u - x_l$;

The absolute value of the interval x: $|x| = \max\{|x_l|, |x_u|\}$;

The midpoint of the interval x: $m(x) = \frac{1}{2}(x_l + x_u)$;

The multiplication of the interval x with a scalar a: $ax = [ax_l, ax_u]$.

Let the group be constituted of m DMs. Let

$$w^{(k)} = \left(w_1^{(k)}, w_2^{(k)}, ..., w_n^{(k)}\right) = \left(\left[w_{1l}^{(k)}, w_{1u}^{(k)}\right], \left[w_{2l}^{(k)}, w_{2u}^{(k)}\right], ..., \left[w_{nl}^{(k)}, w_{nu}^{(k)}\right]\right) \tag{1}$$

be the interval weights of DM_k, $k = 1, ..., m$ when comparing n objects (criteria, alternatives). We propose a new model for aggregating individual interval weights into group interval weights, based on Lehrer – Wagner [5] consensus reaching model and its version, adopted for AHP [6]. The consensus building process goes through more iterations until the final group weights are reached. In each iteration the new weights of DM_k are linear combination of weights of all DMs. The coefficients in linear combination express the importance of DMs regarding DM_k and can be called weights of importance. The convergence of the model is guaranteed when coefficients remain unchanged in the iteration process [5].

Based on the interval arithmetic we define the weights of importance that DM_k assigns to DM_j as a vector

$$p^{(kj)} = \left(p_1^{(kj)}, p_2^{(kj)}, ..., p_n^{(kj)}\right), \tag{2}$$

with $p_i^{(kj)}$ denoting the weight of DM_k to DM_j regarding object i, $i = 1, ..., n$. Weights of importance $p_i^{(kj)}$ are composed of two components $a_i^{(kj)}$ and $b_i^{(kj)}$, and finally normalized:

$$p_i^{(kj)} = \frac{\alpha a_i^{(kj)} + \beta b_i^{(kj)}}{\sum\limits_{r=1}^{m} \left(\alpha a_i^{(kr)} + \beta b_i^{(kr)} \right)} \ . \tag{3}$$

The coefficients α and β define the significance of each component, with $\alpha + \beta = 1$. In the paper we will assume that both components are equally important: $\alpha = \beta = 0.5$.

The component $a_i^{(kj)}$ measures the similarity of the weights of DM_k and DM_j of object i. It is based on the distance of the midpoints of their interval weights. Higher values of $a_i^{(kj)}$ express greater similarity. The component $a_i^{(kj)}$ is also normalized to the interval $[0,1]$, assigning 1 to the pair of weights with equal midpoints:

$$a_i^{(kj)} = \frac{\max\limits_{r,s\in\{1,\dots,m\}} \left(\left| m\left(w_i^{(r)}\right) - m\left(w_i^{(s)}\right) \right| \right) - \left| m\left(w_i^{(k)}\right) - m\left(w_i^{(j)}\right) \right|}{\max\limits_{r,s\in\{1,\dots,m\}} \left(\left| m\left(w_i^{(r)}\right) - m\left(w_i^{(s)}\right) \right| \right)} =$$

$$= \frac{\max\limits_{r,s\in\{1,\dots,m\}} \left(\left| w_{iL}^{(r)} + w_{iU}^{(r)} - w_{iL}^{(s)} - w_{iU}^{(s)} \right| \right) - \left| w_{iL}^{(k)} + w_{iU}^{(k)} - w_{iL}^{(j)} - w_{iU}^{(j)} \right|}{\max\limits_{r,s\in\{1,\dots,m\}} \left(\left| w_{iL}^{(r)} + w_{iU}^{(r)} - w_{iL}^{(s)} - w_{iU}^{(s)} \right| \right)} \ . \tag{4}$$

The component $b_i^{(kj)}$ takes into account the widths of the interval weights of DM_k and DM_j of object i. Higher values of $b_i^{(kj)}$ express narrower intervals. The component $b_i^{(kj)}$ is also normalized to the interval $[0,1]$:

$$\mathrm{b}_i^{(kj)} = \frac{2\max\limits_{r\in\{1,\dots,m\}} \left(d\left(w_i^{(r)}\right) \right) - d\left(w_i^{(k)}\right) - d\left(w_i^{(j)}\right)}{2\max\limits_{r\in\{1,\dots,m\}} \left(d\left(w_i^{(r)}\right) \right)} =$$

$$= \frac{2\max\limits_{r\in\{1,\dots,m\}} \left(w_{iU}^{(r)} - w_{iL}^{(r)} \right) - w_{iU}^{(k)} + w_{iL}^{(k)} - w_{iU}^{(j)} + w_{iL}^{(j)}}{2\max\limits_{r\in\{1,\dots,m\}} \left(w_{iU}^{(r)} - w_{iL}^{(r)} \right)} \ . \tag{5}$$

The components of the weights of importance are symmetric, because $a_i^{(kj)} = a_i^{(jk)}$, $b_i^{(kj)} = b_i^{(jk)}$ for all DMs j, $k = 1,\dots,m$ and all objects $i = 1,\dots,n$ but $p_i^{(kj)} \neq p_i^{(jk)}$ because of the normalization.

The iteration process starts with the individual interval weights ${}^0W_i = \left(w_i^{(1)},\dots,w_i^{(m)} \right)$ of object $i=1,\dots,n$ and with the matrices of weights of importance $P_i = \left(p_i^{(kj)} \right)_{m\times m}$, $i = 1,\dots,n$. The updated weights of the object i after the first round of iteration result in

$${}^1W_i = P_i \, {}^0W_i = \left({}^1w_i^{(1)},\dots, {}^1w_i^{(m)} \right) , \ i = 1,\dots,n. \tag{6}$$

The consensus reaching iteration process is repeated with the same weights of importance:

$${}^rW_i = P_i \, {}^{r-1}W_i = \left(P_i \right)^r \, {}^0W_i , \ i = 1,\dots,n. \tag{7}$$

As r approaches infinity, the revised weights of object i converge towards the consensual group interval weights $^c w_i = {}^c w_i^{(1)} = \ldots = {}^c w_i^{(m)}$, $i = 1,\ldots,n$, where c is the number of iterations needed to reach the convergence.

3 NUMERICAL EXAMPLE

In this section, we give a numerical example to validate the effectiveness of the proposed consensus model through a comparative analysis with a simple aggregation approach: arithmetic mean of interval weights.

Consider the following interval weights given by three DMs ranking four objects:

$$DM_1 = \begin{pmatrix} [0.2, 0.3] \\ [0.1, 0.15] \\ [0.4, 0.5] \\ [0.2, 0.3] \end{pmatrix}, \; DM_2 = \begin{pmatrix} [0.2, 0.4] \\ [0.05, 0.2] \\ [0.3, 0.5] \\ [0.4, 0.5] \end{pmatrix}, \; DM_3 = \begin{pmatrix} [0.3, 0.4] \\ [0.15, 0.3] \\ [0.4, 0.6] \\ [0.1, 0.2] \end{pmatrix}. \tag{8}$$

First we calculate the a- and the b-parts of the weights of importance and the normalized weights of importance p for all four objects:

$$A_i = \left(a_i^{(kj)} \right)_{3\times3}, \; B_i = \left(b_i^{(kj)} \right)_{3\times3}, \; P_i = \left(p_i^{(kj)} \right)_{3\times3}, \; i=1,\ldots,4 \tag{9}$$

$$A_1 = \begin{pmatrix} 1 & 0.83 & 0.67 \\ 0.83 & 1 & 0.83 \\ 0.67 & 0.83 & 1 \end{pmatrix}, \; B_1 = \begin{pmatrix} 0.5 & 0.25 & 0.5 \\ 0.25 & 0 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{pmatrix}, \; P_1 = \begin{pmatrix} 0.4 & 0.289 & 0.311 \\ 0.342 & 0.316 & 0.342 \\ 0.311 & 0.289 & 0.4 \end{pmatrix} \tag{10}$$

$$A_2 = \begin{pmatrix} 1 & 1 & 0.67 \\ 1 & 1 & 0.67 \\ 0.67 & 0.67 & 1 \end{pmatrix}, \; B_2 = \begin{pmatrix} 0.75 & 0.5 & 0.5 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{pmatrix}, \; P_2 = \begin{pmatrix} 0.396 & 0.340 & 0.264 \\ 0.409 & 0.341 & 0.25 \\ 0.35 & 0.275 & 0.375 \end{pmatrix} \tag{11}$$

$$A_3 = \begin{pmatrix} 1 & 0.83 & 0.83 \\ 0.83 & 1 & 0.67 \\ 0.83 & 0.67 & 1 \end{pmatrix}, \; B_3 = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.25 & 0 & 0 \\ 0.25 & 0 & 0 \end{pmatrix}, \; P_3 = \begin{pmatrix} 0.409 & 0.295 & 0.295 \\ 0.394 & 0.364 & 0.242 \\ 0.394 & 0.242 & 0.364 \end{pmatrix} \tag{12}$$

$$A_4 = \begin{pmatrix} 1 & 0.33 & 0.67 \\ 0.33 & 1 & 0 \\ 0.67 & 0 & 1 \end{pmatrix}, \; B_4 = \begin{pmatrix} 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}, \; P_4 = \begin{pmatrix} 0.429 & 0.238 & 0.333 \\ 0.294 & 0.529 & 0.176 \\ 0.368 & 0.158 & 0.474 \end{pmatrix} \tag{13}$$

Then the iteration process is started resulting in the final group interval weights, calculated on the three decimal places after seven iteration steps:

$$^c w = \begin{pmatrix} [0.235, 0.365] \\ [0.099, 0.210] \\ [0.370, 0.530] \\ [0.226, 0.326] \end{pmatrix}. \tag{14}$$

Figures 1 - 4 present the interval weights of all three DMs, the group weight derived by the new consensus reaching model, and the group weights derived by the arithmetic mean.
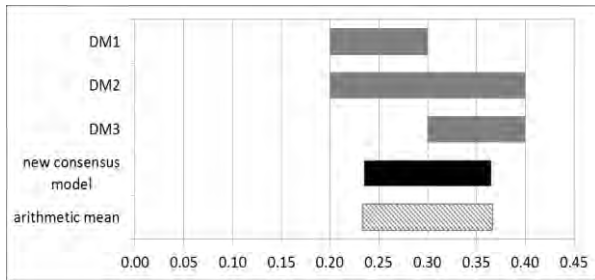
Figure 1: Interval weights for object 1
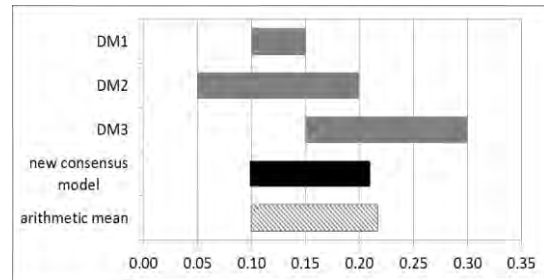
Figure 2: Interval weights for object 2
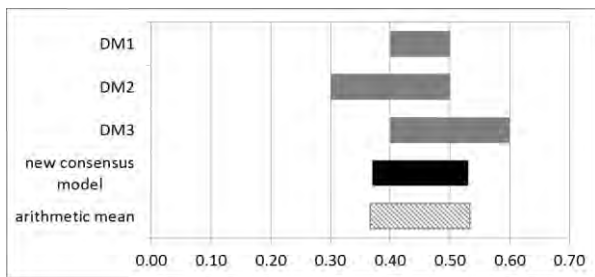
Figure 3: Interval weights for object 3
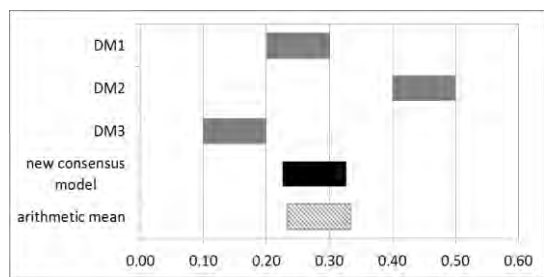
Figure 4: Interval weights for object 4

We can conclude that the interval group weights gained by the proposed model are similar to the weights obtained by the arithmetic mean, what confirms the soundness of the new model. Furthermore, the interval weights derived from the new model are slightly narrower than the arithmetic mean of the interval weights, what also gives reliability to the new model.

4 CONCLUSIONS

In the paper a group decision making problem with the interval weights has been investigated, and a new consensus reaching model has been provided. The presented aggregating algorithm is based on the Lehrer – Wagner model [5], what assures the convergence of the iterative process. The offered numerical example demonstrated the effectiveness of the proposed model. The merits of the new model are the following. It provides a technique for aggregating interval weights that can result from several multi-criteria decision methods. The idea of the presented model is simple and straightforward what assures that the model is easy understandable. The proposed model takes into account two important aspects: the differences between the interval weights, and the widths of the interval weights. The significance of the latter can vary what influences the weights of importance of DMs. In our future research work we plan to extend the proposed model to fuzzy weights.

References

[1] Cabrerizo, F. J., Pérez, I. J. and Herrera-Viedma, E., 2010. Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information. Knowledge-Based Systems, 23(2), pp. 169-181.

[2] Dong, Q. and Cooper, O., 2016. A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making. European Journal of Operational Research, 250(2), pp. 521-530.

[3] Elzinga, C., Wang, H., Lin, Z. and Kumar, Y., 2011. Concordance and consensus. Information Sciences, 181(12), pp. 2529-2549.

[4] Guha, D. and Chakraborty, D., 2011. Fuzzy multi attribute group decision making method to achieve consensus under the consideration of degrees of confidence of experts' opinions. Computers & Industrial Engineering, 60(4), pp. 493-504.

[5] Lehrer, K. and Wagner, C., 1981, Rational Consensus in Science and Society, Reidel, Dordrecht.

[6] Regan, H. M., Colyvan, M. and Markovchick-Nicholls, L., 2006. A formal model for consensus and negotiation in environmental management. Journal of Environmental Management, 80(2), pp. 167–176.

[7] Xu, Y., Li, K. W. and Wang, H., 2013. Distance-based consensus models for fuzzy and multiplicative preference relations. Information Sciences, 253, pp. 56-73.

[8] Xu, Y. and Wang, H., 2013. Optimal Weight Determination and Consensus Formation under Fuzzy Linguistic Environment. Procedia Computer Science, 17, pp. 482-489.

[9] Yaniv, I., 2004. Receiving other people's advice: Influence and benefit. Organizational Behavior and Human Decision Processes, 93(1), pp. 1-13.

MULTILAYER EVALUATION MODEL FOR PROJECT TEAM COMPETENCIES

Domen Ocepek[1] and **Vladislav Rajkovič**[2]
[1] Kopa d. d., Kidričeva 14, Slovenj Gradec[1], domen.ocepek@kopa.si,
[2] Faculty of Organizational Sciences, Kidričeva cesta 55a, Kranj[2],
vladislav.rajkovic@fov.uni-mb.si

Abstract: Transparency of competencies, roles, and conditions have positive impact on project performance. In this paper, a hierarchical multilayer model is presented. The purpose of this model is better understanding and evaluation of each member competencies, roles in a team, impacts, and the scope of the project, presented by the multilayer tree of attributes and proactive communication through the project board, which represents the new way of presentation and visualization. The model was tested using the results of a survey, which was conducted on real IT projects.

Keywords: Evaluation model, competencies, teamwork, organization, project board

1 INTRODUCTION

Process reengineering and digitalization, which are the main goals of IT projects, is a complex and demanding process that requires well-organized project teams, right skills-competencies and understanding the individuals and project team's behaviour. The survey of Rollings [12] shows that technical skills and competencies are not the cause of IT project failure but mainly organizational and other soft skills and competencies. This raises the question what are the competencies that ensure successful implementation of IT and other projects. Although teams are considered to be the building blocks, there is a lack of coherence, integration, and understanding of how team composition effects relate to important team outcomes. Mathieu, Tannenbaum, Donsbach, and Alliger [10] note that the process is going through a team dynamics and that the effectiveness of the team is achieved only when its composition, integration, and understanding is at a high level. Individual competencies are just one layer and understanding the meaning and importance of roles and other conditions is crucial for success of the projects.

Gartner [7] and other surveys show that IT projects success is below expectations. My experiences show that inadequate required and actual competencies lead to inappropriate roles played. Unclear rules on several levels often lead to poorly implemented projects that fail to meet expectations. Digital competences published in Europas [3] cover areas of information processing, content creation, problem solving, communication, and safety. They represent the integral part of the full range of competencies, which are a prerequisite for the successful implementation of IT projects. Digital literacy is becoming important for the successful functioning of an individual, teams, and projects.

Levi [9] notes that the complexity of the project requires a team problem solving. With teamwork problem solving, we can take advantage of the synergy of team competencies. My experiences show that commitment and accepted roles are very important for the success of the projects. For the project manager or team leader, soft competencies are even more important than technical competencies. Teams that provide all the necessary competences, accept roles, responsibility, allow good communication, and respect each other are successful. This article is dedicated to this specific area of project where performance depends on competencies, roles, conditions, and project board, which helps to raise

awareness about what is happening on the project and enables clear communication through the planning and implementation phase of the project.

2 METHODOLOGY

The methodological approach is based on a qualitative MCDM method DEX. The model presented in this paper contains three layers:
Layer 1: Evaluation of individual key competencies.
Layer 2: Evaluation of team key conditions and roles.
Layer 3: Evaluation of project key conditions and roles.

DEX [2] is a *hierarchical* method. Attributes are organized in *hierarchy*. Observed in the top-down direction, hierarchy represents a decomposition of the decision problem into sub-problems. The bottom-up direction denotes dependence, therefore, higher-level attributes depend on the lower-level, more elementary ones. The most elementary attributes, called *basic attributes*, appear as terminal nodes of the hierarchy and represent the basic observable characteristics of alternatives. Higher-level attributes which depend on one or more lower-level ones, are called *aggregated attributes*; they represent evaluations of alternatives.
The survey that was conducted on seventeen IT projects gives the values for attributes in the tree of competencies and confirms adequacy of the model.

The last layer of this model is further divided and presented through the project board with special visualization through matrix of hundred pieces, which are linked to the areas of the project and carry information or message. The scope of these projects board was defined through a guided interview with five experienced project managers.

3 MULTILAYER PERSPECTIVE

The multilayer model is designed from the perspective of the individual and his core competencies highlighted by the Commission of the European Communities (2006). The next layer represents the teams and conditions for their functioning based on Belbin (2017) team roles. The last layer represents a project that can only succeed when those layers are tuned and competencies allow the roles to be played and activities are oriented to provide coordination and communication between all areas and participants of the project.

The company Kopa d. d, whose primarily focus is process digitization, developed dynamic competencies expert system E-DKES [8] for better understanding of actual and required competencies with visualization, which clearly shows the matrix of actual and the desired state. The continuation of this idea was to develop a multi-layer model for wider picture of competencies, roles, and conditions for successful implementation of projects, which is presented in this paper.

3.1 Individual core competencies

The tree of individual competencies is built on the basis of eight competencies (Mother tongue, Foreign language, Mathematical literacy and competence in science and technology, Digital competence, Learning to learn, Interpersonal, intercultural, social and civic competence, Entrepreneurship) that have been identified by the Commission of the European Communities [4]. The scope of competencies in mathematics, science and technology could be analysed deeper. However, in this study, we will stay at this level.

These eight competencies are divided into two main groups: those that are associated with the acquired skills and those that are related to patterns of action and behaviour. The

knowledge is further divided into basic knowledge, communication, functioning, to approach, and features.

1st layer   **2nd layer**   **3rd layer**
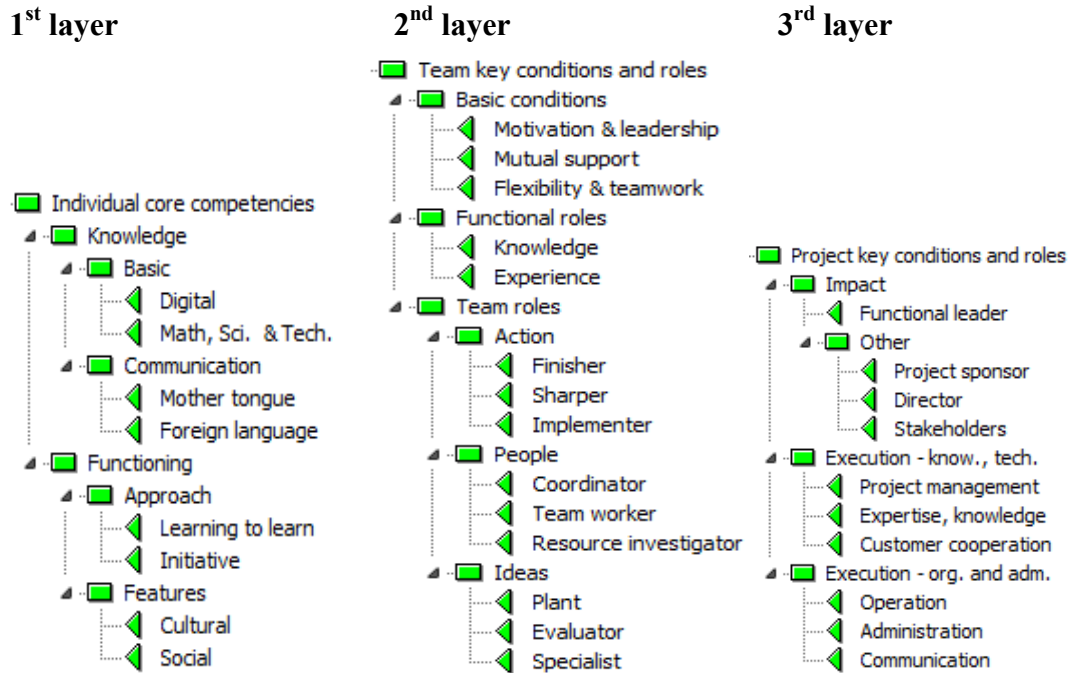
Figure 1: A tree of competencies, conditions, and roles for project performance

The values of attributes (the leaves of the tree) are on the three-step scale (low, medium, and high). These values are aggregated at the level of branches on five-step scale (insufficient, sufficient, good, very good, and excellent).

Middle level means that a person is independent and that he can operate without major problems and can provide knowledge, skills, and experiences that are required. People with medium level of individual competence are independent and occasionally assist those with low level and get help from those with high level. People with low level need assistance of persons with medium and higher level. The person with the highest level is high qualified and can help others on low and middle level.

Medium level for information literacy for example means that the person is able to find a specific Internet content to create documents and spreadsheets, use e-mail, and use tools such as Skype and similar tools for communication. Low level means that the person has difficulty in applying information technology. High level means that the person is efficient user of modern information technology who can find new solutions and help others to use these new solutions and technology.

The utility function is designed on the basis of certain empirical assumptions. For example, if the score on two or more attributes is at a low level, the overall assessment is insufficient. Determining the value of an attribute is carried out based on self-assessment or other appropriate testing such as online tests and interviews.

Digital	Math, Sci. & Tech.	Basic
1 low	1 low	1 low
1 low	2 medium	1 low
1 low	3 high	2 medium
2 medium	1 low	1 low
2 medium	2 medium	2 medium
2 medium	3 high	3 high
3 high	1 low	2 medium
3 high	2 medium	3 high
3 high	3 high	3 high
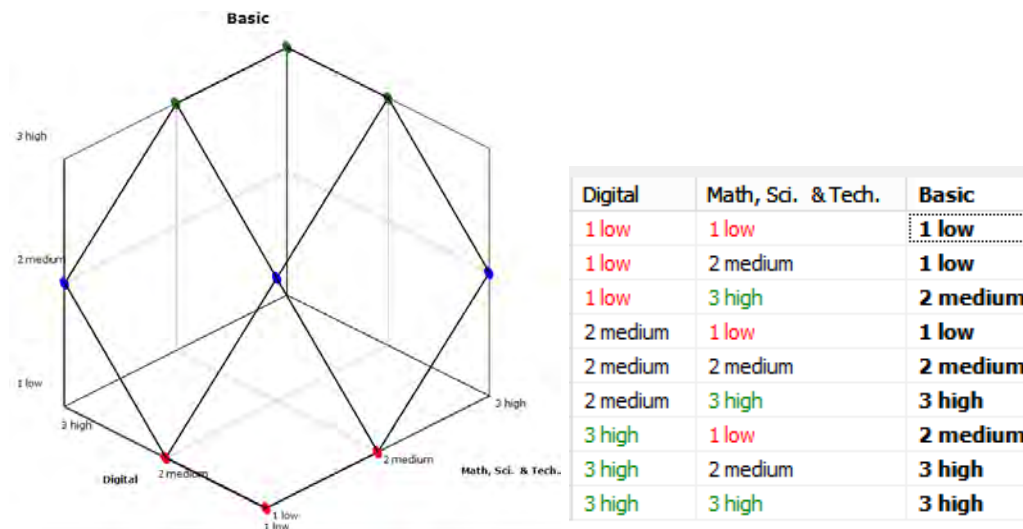
Figure 2: Utility function for basic competencies

3.2 Conditions roles and competencies of the team

The project team requires specific roles, which require appropriate skills-competencies. In addition, to be the operational team, one should meet other conditions, such as leadership and motivation, mutual monitoring of performance, flexibility, and team orientation.

Team conditions and roles are divided to three groups (branches), basic or. fundamental conditions of the team, the functional roles and team roles. Team roles are divided to action, people, and ideas. The second layer in Figure one shows the details of this division.

Utility functions are made on the basis of empirical assumptions.

The rating "good" is the average score, which means that the basic conditions are met at the primary level. As an example for management and motivation, "insufficient" means that the individual is demotivated and demotivate others and disturbing the team. The level of "sufficient" motivation is at a minimum level, which still allows the operation and that person takes a lot of time and energy for motivation and leading. At the middle stage, "good" person is motivated. At the stage of "very good", the person is showing greater enthusiasm and motivation. At the level of "excellent", people are highly motivated and motivate other colleagues and have the characteristics of good leaders and successfully play this role.

The utility function is made on the basis of empirical assumptions.

3.3 Project key conditions and roles

Conditions, roles, and areas of the project, presented with a tree of attributes

This part of the tree represents the third layer, the conditions, roles, and areas of the project that are divided to three branches. The values of attributes (the leaves of the tree) are on the three-step scale (low, medium, and high).

3.4 Project board

The idea of the project board was recognized by five project managers as a useful solution to increase understanding and communication on the project.

The purpose of the project board is to raise the transparency and increase the informational value, which is often lost in complexity of digitization projects, based on the new method, called "meaning content and communications" (hereinafter referred to as

"MCC"). This method assumes that understanding of the whole project can be improved if it is divided into an appropriate number of key shares. This number should not be too high, as this will lose transparency. Forsyth [5] states that Pareto rule, also known as the 80/20 rule, linking the causes and consequences of the ratio is about 80/20 and is found in many business activities. This rule is also reflected in this method. In this case, six areas were identified. Due to the comprehensiveness of percentage, the project panel is divided into one hundred equal pieces. According to the authors, it is not reasonable to deal with less than four building blocks, which represent area of the project. The authors recommend less than eight areas.

The scope of the project - areas	%
Expertise - professional resources	21
Project management	20
Cooperation with the customer - market	20
Team work	18
Communication	15
Administration	6
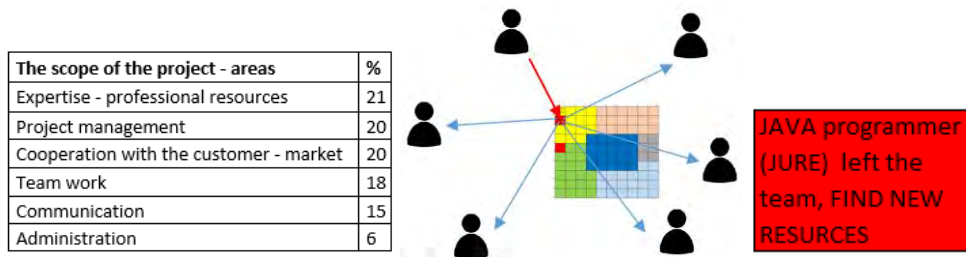
Figure 3: Project board meeting solving problem on the area of professional resources

4 RESULTS

The survey results used for testing, show that the rated value of the project performance with the model matches the rated value from the survey. The figure 4 shows that section of the tree, which has the biggest impact on the result.

	PROJECT 1	PROJECT 2	PROJECT 3
Performance conditions	**5 excelent**	2 sufficient	4 very good
......			
Project key conditions and roles	**5 excelent**	2 sufficient	3 good
Impact	**3 encouraging**	**1 annoying**	2 acceptable
Functional leader	**3 encouraging**	**1 annoying**	2 acceptable
Other	2 acceptable	2 acceptable	2 acceptable
Project sponsor	2 acceptable	**1 annoying**	2 acceptable
Director	2 acceptable	2 acceptable	2 acceptable
Stakeholders	2 acceptable	2 acceptable	2 acceptable
Execution - know., tech.	**3 high**	2 medium	2 medium
Project management	**3 high**	2 medium	2 medium
Expertise, knowledge	**3 high**	2 medium	2 medium
Customer cooperation	**3 high**	2 medium	**3 high**
Execution - org. and adm.	**3 high**	2 medium	2 medium
Operation	2 medium	2 medium	**3 high**
Administration	**3 high**	2 medium	2 medium
Communication	**3 high**	**1 low**	2 medium

Figure 4: Results and detailed explanation for project performance

5 CONCLUSION

Experiences in digitization projects show that success often depends on how competent project teams implement those projects and how successfully they communicate. Technical competences are just a part of a wide range of competences. Roles and conditions in combination with competencies make up a complete multilayer picture of projects performance. The purpose of this hierarchical tree of attributes and project board is to establish a central and comprehensive monitoring of the project and to raise awareness of what is going on with the project. With the introduction of the model for understanding and evaluating competencies, roles, conditions, and using the project board, gives each member of the project an opportunity to highlight the problem on project board, and is there until it is resolved. The model helps to locate and evaluate the area in which the disorder is caused to tackle the real causes and provide the missing competencies, roles and interfere with the disturbances and eliminate disturbing factors occurring to implementation of digitization projects.

Resources

[1] Belbin – Team Roles http://www.belbin.com/about/belbin-team-roles/ [Accessed 05/02/2017].

[2] Bohanec, M., Rajkovič, V., Bratko, I., Zupan, B., & Žnidaršič, M. (2013). DEX methodology: Three decades of qualitative multi-attribute modelling. *Informatica,* 37, 49–54.

[3] Europas. (2017). Digital competencies. https://europass.cedefop.europa.eu/sl/resources/digital-competences/ [Accessed 28/02/2017].

[4] European commission. (2006) Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning (OJ L 394, 30.12.2006, pp. 10-18)

http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Ac11090/

[Accessed 20/02/2016].

[5] Forsyth, P. (2010). *Successful time management* (rev. 2nd ed.). London, Philadelphia, New Delhi: Kogan Page.

[6] Frame, J. D. (1999). *Project management competence: Building key skills for individuals, teams, and organizations*. San Francisco: Jossey-Bass.

[7] Gartner. (2015). 12 Skills Critical to Business Process Management Success. https://www.gartner.com/doc/3107625/-skills-critical-business-process/ [Accessed 28/02/2016].

[8] Kopa d. d. (2014). *E-DKES project documetation – Dinamični kompetenčni ekspertni sistem* (internal documetation). Slovenj Gradec: Kopa d. d.

[9] Levi, D. (2001). *Group dynamics for teams*. Thousand Oaks, CA: Sage Publications.

[10] Mathieu, J. E., Tannenbaum, S. I., Donsbach, J. S., & Alliger, G. M. (2014). A Review and Integration of Team Composition Models: Moving Toward a Dynamic and Temporal Framework. *Journal of Management*, 40(1), 130–160.

[11] Nikoloski, T., Udovič, A., Pavlovič, M., & Rajkovič, V. (2016). Večkriterijski model za oceno primernosti preusmeritve dejavnosti kmetij. *Zbornik 35. mednarodne konference o razvoju organizacijskih znanosti 16.–18. marec 2016* (str. 692–701). Kranj: Moderna organizacija.

[12] Rollings, M. (2013). *Why projects fail? Hint – It's not technical skills.* http://blogs.gartner.com/mike-rollings/2013/03/28/why-projects-fail-hint-its-not-technical-skills// [Accessed 20/02/2016]

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Session 8:
OR Perspectives:
Where we have been,
where we can go

Optimizing enjoyment of mathematics and OR education with introducing psychological concepts flow and grit using simulation-based model of emotional states of learning

Drago Bokal,
Faculty of Natural Sciences and Mathematics, University of Maribor,
Koroška cesta 160, Maribor, Slovenia, d@bokal.net

Abstract: Can studying mathematics and operations research be an enjoying experience? We investigate conditions leading to studying mathematics being a psychologically optimal experience for both professors and students. We combine psychological theory of optimal experience (flow) and mechanisms of academic success (grit) with microeconomic theories of utility maximizing time allocation and practitioners' classification of knowledge discovery progress to develop a discrete time model of learning experience. We find that a combination of existing scientific findings does suggest approaches that lead to enjoying, fulfilling experience with great certainty, and that approaches to studying Operations Research within the mathematics curriculum can play a significant role.

Keywords: higher education in mathematics, operations research education, time allocation, flow, grit, technology readiness levels.

1 INTRODUCTION

How, to what extent could studying and teaching mathematics be an enjoying experience for professors and students? Every individual involved in these studies is bound to answer this question at least implicitly, hence there is no doubt about its relevance. Yet scientific approaches to this problem are scarce, which is understandable due to it personal nature. Nevertheless, there have been some scientific discoveries in psychology that bear significance upon the topic, and it is the aim of this paper to bring them to attention of the relevant audience.

In Section 2, we introduce a model of the processes described in the theory of optimal experience (flow, introduced by M. Csikszentmihalyi [3], [4]) and mechanisms of academic success (grit, studied by A. Duckworth's [6]) starting with the microeconomic theories of utility maximizing time allocation as the starting model. As a very recent review of time use modelling [10] does not feature any discrete time use models and as these are required for our description of the psychological processes we study, we develop such a model. The two psychological topics both feature *purpose* as a significant contributing factor to optimal experience and performance; we posit that *utility* is the corresponding microeconomic concept. We discuss the relationship between both using a recent theory of perception [9] as well as their interpretation in the context of studying and teaching.

In Section 2, we show that simulations using the new model exhibit the behaviours observed by psychologists [6], hence the model allows professors and students to conduct their own simulations of learning and research practices, familiarizing themselves with benefits and pitfalls of different approaches. As the real world data on these topics cannot be easily collected due to difficult experiment design and execution, the simulations may serve as a laboratory to contemplate the topics, and collaboration with psychologists may lead to corroborating real world experiments; in this aspect, our approach is similar to [9].

The paper has two layers of contribution: the informal one, which is rather professional than scientific, steering the discussion about the learning context of mathematics and OR, and the formalized, scientific one, modelling the learning processes described by psychological research. This second layer is developed into a software that can be used by those involved in OR education to simulate effect of various time usage strategies. Due to limited space, we refer
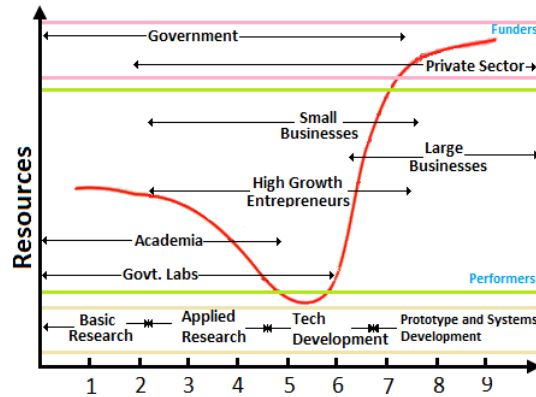
Figure 1: Availability of resources over technology readiness levels and the valley of death at levels 4–6. Adopted from [8].

the readers to relevant bibliography for deeper treatment of the concepts of the context layer, and to either personal communication or a later journal publication for more rigorous treatment in the modelling layer. It is our goal to spread awareness about the concepts presented, as well as to find collaborators who would like to research the usage of these concepts in actual OR curricula.

2   A MODEL OF LEARNING THROUGH GRIT AND FLOW

Hoffman describes in [9] a series of experiments in simulated evolution, demonstrating that veridical perceptions – those true to the structure of the simulated world – are routinely dominated by non-veridical perceptions tuned to fitness. Generalizing and perhaps stating the obvious, when agents are submitted to evolutionary pressure, the behavioural strategies not contributing to fitness of agents are driven to extinction, despite their possible contribution to the more efficient understanding of the environment (which may be irrelevant to fitness). We may extrapolate Hoffman's argument that individual's perception of enjoyment and purpose has evolved to contribute to individual's procreation or proliferating ideas. However, human culture has augmented these primal meanings of purpose, and a plurality of existing cultures allows individuals to even choose their own purpose. But Hoffman's model from mathematical psychology nevertheless allows us to claim that what psychologists understand under *purpose* is the *function that individuals are maximizing in their environment*. In microeconomics, the same function is called *utility*, and as this term is more familiar to mathematicians as widely accepted in game theoretic setting, we will use them interchangeably in our models consciously ignoring the subtle details left over to each respective domain.

Common human cultures build around common definitions of purpose-utility. In higher education, the purpose of professors is producing and proliferating new and existing knowledge, and the purpose of students is to accumulate that knowledge and bring it to businesses where it is employed for societally desired functions resulting in students' careers.

The scale measuring progress of knowledge from basic discovery (where professors do research and where students study) to its applications (where business creates added value and most students seek employment) is called the technology readiness levels (TRL) scale. The concept, popularized by NASA whitepaper of Mankins [13] and adopted by European Commission for evaluating technologies on their path to applicability in H2020 public grants scheme [15], is most tersely illustrated by Figure 1.

550

We discuss TRL in greater detail in a parallel contribution [2]. Here, the key observation is that most of mathematics research falls into basic research (TRL1 – theory, theorems and TRL2 – models, algorithms), whereas Operations Research with implementations and case studies reaches the applied research (TRL3 – prototypes, implementations of models and algorithms, TRL4 – lab validation, TRL5 – relevant environment validation). All mentioned technology readiness levels hence form the context where most of the professors evaluates the utility of their activities, consisting mostly of conducting research, publishing papers, and educating students with knowledge at these TRLs.

The students, on the other hand, aspire to evaluate their activities with utility at higher TRLs, where they will eventually find employment. While seeking employment, they cross the so-called "valley of death" at intermediate TRLs to which the society allocates only few resources. Their ultimate purpose changes from acquiring basic knowledge to earning money offering services and products at higher TRLs. Anticipating this change may be demotivating for students without either intrinsic motivation for studying theoretical content or without trust into the composition of the curriculum to reflect competences required by their career, which may lead to grade inflation [11]. This creates opportunity for OR professors to align their purpose with students to a greater extent, as most student's purpose lies in employment at higher TRLs. Aligning the interests of students and professors yields incentive schemes that contribute to eliminating grade inflation [14]. As OR builds on mastering basic topics such as algebra, analysis, and discrete mathematics, OR curriculum can be a buffer motivating students for more thorough study of these basic courses.

According to a recent survey on time allocation models [10], early models of optimal time allocation [1, 5] feature an agent maximizing utility $U(X_1, \ldots, X_n, T_1, \ldots, T_n)$, where X_i is the i-th commodity consumed and T_i is the time devoted to its consumption. The commodities are subject to budget constraint and the allocated time is subject to total time constraint, some of which is spent for work, i. e. increasing the budget. Activity-based models diverged from commodity-based models in [12], where utility was maximized as $U(T_1, \ldots, T_n, q_1, \ldots, q_n) = \sum_{i=1}^n U(T_i, q_i)$ and T_i is a continuous time variable allocated to activity i with attributes q_i.

From Kitamura's model [12], we adopt a set of activities to which the time is allocated, however, our process requires that time is consumed at a given activity one unit at a time. The agent is to decide which activity to allocate the time on for each unit. Any activity that is performed consumes an unit of time and updates the state of that activity, i. e. $q_i^{t+1} = f(q_i^t)$. The utility of the individual at any time is computed as $U^t = U(q_1^t, \ldots, q_n^t)$. This model mimics the time allocated to studying, practice, training: each practice takes at least some time to be conducted, and during each practice, a certain skill is improved: the utility (i. e. the knowledge of the individual) is progressing and depends on the combination of the skills learned.

How can gathering utility – collecting knowledge – be an enjoying experience? It turns out that this question requires another dimension to be studied: skill level needs to be augmented by the challenge level of the activity in which the skill is being applied.

Csikszentmihalyi has studied emotional states of people conducting various activities, and classified them according to the skill level and the challenge level they felt while performing the activity [3]. The eight emotional states that were identified are depicted in Figure 2. He called the state when both skill level and challenge level are high as the state of *flow*, the optimal state in which *the activity absorbs the person and the person is most efficient conducting the activity*.

From his research [3], we adopt the following process of updating the activities' states: whenever an activity is being practised, the skills required by the activity are increasing according to a *learning* parameter depending on the individual. When the skill increases sufficiently, the challenge also increases in order to avoid boredom. This is repeated, until the individual has reached the area of flow. This will be the basic algorithm of our simulation; the precise equations governing it will be presented in what follows.
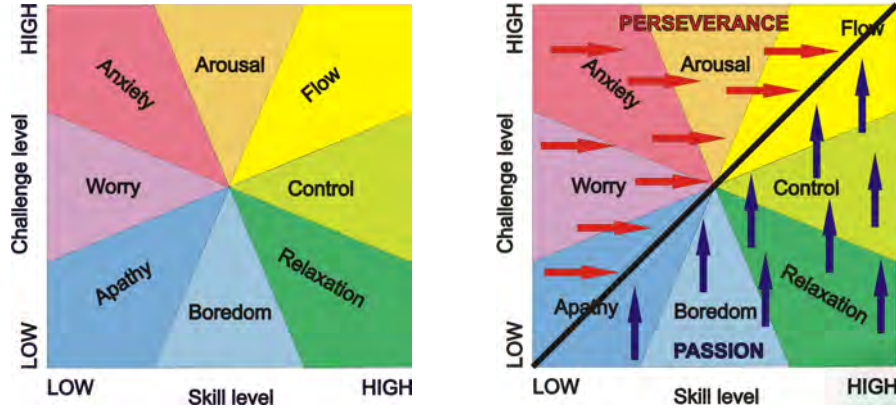
Figure 2: Emotional states depending on skill and challenge levels (left). Grit as the emotional field converging in flow (right).

While flow is a psychological concept relating to an enjoying experience while conducting activity, the process of reaching it is only vaguely described. A much more detailed description of successful learning approaches comes from Duckworth's investigation of *grit* [6]. It has been demonstrated that grit is the best predictor of academic success of students: only the most successful Ivy league students were found slightly less gritty than their less successful peers, but in all other studies described in [6], the most gritty individuals were the most successful.

Duckworth defines *grit* as *a combination of passion for the activity we are performing and perseverance we exhibit when facing obstacles* [6]. In our skill/challenge landscape, passion contributes to increasing challenge when an individual is bored, and perseverance makes the individual keep fighting obstacles when high challenges inhibit immediate success. Grit as a trait in a person hence constitutes an emotional field that helps the person seek challenges when below and improve skills when above the line of balance between skill and challenge, as we illustrate in Figure 2 (right). Such a field directs person's emotional state towards the region of flow, which is both an emotional award as well as a state of frictionless experience of enjoying the activity one is performing well [3].

In her discussion of deliberate practice, Duckworth observes that there are several skills one needs to master when becoming an expert [6]. In our model of learning, we hence posit that time is allocated to developing skills: each skill has two parameters, the ability to perform it (i. e. the skill level) and the challenge level reached in that skill. The individual is allocating time to developing a set of skills required for his (future) job. The utility of a set of skills is a question to be researched. We assume that when performing a job, the least learned skill defines the quality of the result.

From the practical perspective, we add another element to the model: a skill-improving activity can terminate either successfully or with a failure. Failure can occur because an individual either gives-up the particular learning activity, which happens with probability $p_g(1 - a_i)(1 - E(a))(1 - p)$, where p_g is a parameter, a_i is the skill level of the activity i performed, $E(a)$ is the average ability over all skills (with this we model that other skills may compensate for the lack of a specific skill) and p is perseverance of the individual. Failure can also occur if the individual performed the whole activity, but perhaps wrongly: this happens with probability $p_g(1 - E(a))$.

A successful activity results in an increase of the skill level a_i, equal to $\delta_{sa}lc_i(1 - a)$, where δ_{sa} is a factor defining the basic increase (used to control the process speed), l is the learning speed of the simulated individual c_i is the challenge level of the skill i, and the factor $(1 - a_i)$ asserts that the skill level is bounded by 1. Also, at every successful activity, challenge c_i is increased by $\delta_{sc}p(1 - c_i)$, where δ_{bc} controls the speed of challenge increase, p is the passion of
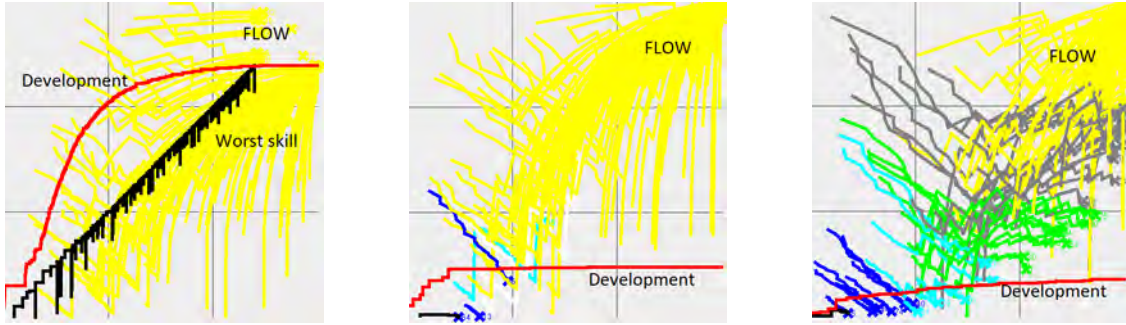
Figure 3: Learning as deliberate practice, stochastic progress, and arrested development.

the simulated individual, and $(1 - c_i)$ bounds the challenge level at 1.

A failed activity results in some learning increase, which is equal to $\delta_{fa} l c_i (1-a) \epsilon$, where δ_{fa} is a factor defining the basic increase at failure, and ϵ is the random element modelling the fact that the failure may have occurred at a random time during the activity. The challenge level is decreased after failure by $\delta_{fc} (1-p) c_i$, where δ_{fc} defines the basic decrease factor, p is the perseverance and $(1-p)$ allows the decrease to be negatively correlated by perseverance, and the challenge level c_i assures the decrease in challenge level is always less than the challenge level itself.

From Duckworth [6], we also adopt three approaches to the learning – simulation process: the first she called *deliberate practice* and consists of identifying weak spots and practising those specific aspects of our knowledge. We simulate this so that at each step, practising the skill that has either smallest skill level or smallest challenge level. The second is the opposite of deliberate practice and we call it stochastic learning. In it, we pick the skill to be trained randomly with equal probability. For the third level, we build upon Duckworth's observation [6] that with some individuals, the passion for an activity decays if it is not stimulated by a trainer or some other mechanism: for these individuals we decrease their passion by a factor slightly less than one after each activity.

We simulate the learning outcomes of the described processes creating a set of 100 skills (corresponding to many aspects required to successfully perform a complex task students are learning for), and randomly choosing the initial skill and challenge levels for these skills. Then we run each of the three variants of the learning process for 10 000 iterations, corresponding to folklore rule of 10 years or 10 000 hours of deliberate practice in a domain to become an expert [6]. All parameters and variables used are from the interval $[0, 1]$, so we can plot the progress of activities as a skill/challenge trail in the unit square. In addition, we also plot the expected skill/challenge level and the progress of expected skill level for each individual. Figure 3 shows typical results.

The leftmost image in Figure 3 exhibits results of deliberate practice: by practising the least skilled or increasing challenge in the most boring aspect of her knowledge, the individual was able to bring all the individual skills into the flow zone (upper right square) and reach the flow zone with the worst skill, as well as reach expert level of knowledge. With stochastic learning (center image), most of individual skills reached the flow zone, yet some weak ones hardly progressed and experienced challenge decay, resulting in inferior overall skill level (red curve). Finally, when passion is decaying, a clear case of arrested development reaching only a suboptimal plateau of competence occurs: some, but not most of the activities reach the flow zone and the overall skill level ("development" curve) does not reach a significant level.

While there is no short answer to the question posed, we summarize the findings claiming that purposeful studying and researching mathematics can be fulfilling, that a combination of existing scientific findings does suggest approaches that lead to this outcome with great

certainty, and that Operations Research as one of the areas of mathematics with highest level of technology readiness could contribute significantly to the optimality of studying and teaching experience and student motivation even for more theoretical topics of mathematics. Parallel contributions to SOR'17 [2, 7] present a particular technique of specific curricular approach that could be directly implemented in both OR education and research, and contribute to grit development and flow experience in students and professors of mathematics.

3 Acknowledgements

References

[1] Becker, G. S. (1965). A Theory of the Allocation of Time. The economic journal, 493–517.

[2] Bokal, D. and Goričan, A. (2017). Operations research as the bridge over technological valley of death. Submitted to SOR 2017.

[3] Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. New York: Harper Perennial modern classics.

[4] Csikszentmihalyi, M. (2004). Good business: Leadership, flow, and the making of meaning. New York: Penguin.

[5] DeSerpa, A. C. (1971). A theory of the economics of time. The Economic Journal, 81(324): 828–846.

[6] Duckworth, A. (2016). Grit: The power of passion and perseverance. New York: Scribner.

[7] Goričan, A., Bratuša, A., and Bokal, D. (2017). Knowledge transfer ontology of photo-voltaic electricity production forecasting. Submitted to SOR 2017.

[8] Hensen, J., Loonen, R., Archontiki, M., and Kanellis, M. (2015). Using building simulation for moving innovations across the "Valley of Death". REHVA Journal, 52(3), 58–62.

[9] Hoffman, D. D., Singh, M., and Prakash, C. (2015). The interface theory of perception. Psychonomic bulletin & review, 22(6), 1480–1506.

[10] Jara-Díaz, S., and Rosales-Salas, J. (2017). Beyond transport time: A review of time use modeling. Transportation Research Part A: Policy and Practice, 97, 209–230.

[11] Johnson, V. E. (2003). Grade Inflation: A Crisis in College Education. New York: Springer-Verlag.

[12] Kitamura, R. (1984). A model of daily time allocation to discretionary out-of-home activities and trips. Transportation Research Part B: Methodological, 18(3), 255–266.

[13] Mankins, J. C. (1995). Technology readiness levels. NASA White Paper.

[14] Smole, A., Kent, D., Jagrič, T., and Bokal, D. (2017). Social dilemma model of grade inflation calls to end focusing on grades in favor of knowledge. Submitted to Educational Evaluation and Policy Analysis.

[15] G. Technology Readiness Levels. (2015). European Commission. https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/ h2020-wp1415-annex-g-trl_en.pdf [Accessed 27/03/2017].

EURO – PER ASPERA AD ASTRA

Jakob Krarup
DIKU, Department of Computer Science, University of Copenhagen
Ydervang 4, DK-3460 Birkerød, Denmark
krarup@di.ku.dk;
https://www.euro-online.org/media_site/PP/KrarupCV.pdf

Abstract: A suitable motto for EURO, The Association of European Operational Research Societies, should be: *"per aspera ad astra",* which translates as "through hardships to the stars". This Latin phrase of uncertain origin (Hesiodos, Seneca, Vergil) is widely used: as an inscription on a multitude of buildings worldwide though notably schools and universities, and the motto of many organisations. It is an apt characterization of EURO's history about which three fragments will be presented here. EURO was launched in a wave of enthusiasm and ideas (1975-1988), then experienced its *"per aspera"* - the turbulent years (1989-1990), and now enjoys it *"ad astra"* years (1991-date) as evidenced by the cornucopia of offers available to today's members.

Keywords: EURO, history.

1 OR AND DATALOGY

Peter Naur (1928-2016) is famed as one of the fathers of the algorithmic language Algol 60 and in 2005 recipient of the prestigious Turing Award. In line with words such as biology and sociology and realizing that *data* is the key notion, he introduced in 1966 the word *datalogy* to characterize the branch of science known elsewhere as computer science. Peter became in 1969 Denmark's first Professor of Datalogy, and DIKU (in Danish: Datalogisk Institut, Københavns Universitet) was founded the following year.

> *"What's in a name? that which we call a rose*
> *By any other name would smell as sweet."*
>   (Shakespeare, Romeo and Juliet, Act II,2).

What's in a name? Though well justified, all attempts to export the term *datalogy* to the outside world appeared to be unsuccessful. Equally unsuccessful was the importation of the term OR. Being a *datalog* (computer scientist) is nowadays a recognized profession in Denmark like e.g. a physician, whereas an *operations analyst* is not - and never was.

Thus, for example, seated at a party next to a nice dinner partner, and being inquired about my profession, the term OR is of no use as it either will ring no bell or give rise to wrong associations. As one of the definitions of OR is "… what OR people are doing", the solution is then to sketch some particularly conspicuous examples of applied OR, or, if that does not suffice, ask her for her email address and send an article on the day after.

Since 2013 my favourite article for this purpose has been the Invited Review [4], co-authored with S. Fores, Manager of EURO, where the abstract reads as follows:

Operational Research (OR) is the science of decision-making. From its military origin on the eve of World War II, OR has over the past seven decades matured to become a discipline that is recognized worldwide for its contributions to managerial planning and complex operations on all levels within both private companies and public institutions. Besides being an indispensable tool as a means of decision support, OR is today a well-established academic discipline and a field with its own institutions. Thus, OR-professionals are joined in national societies worldwide, assembled since 1959 in the global organization *International Federation of OR Societies* (IFORS) which again is subdivided into four *Regional Groupings.* Among those is the *Association of European OR Societies* (EURO) having as members the national societies of 32 countries notably in Europe. Two questions

will be addressed: what is OR all about? How do National OR Societies fit within the frameworks of IFORS and EURO? Partial answers are provided to both.

2 HOMAGE TO A FOUNDING FATHER OF EURO

"Those who cannot remember the past are condemned to repeat it." (George Santayana – Wikiquote https://en.wikiquote.org/wiki/George Santayana). A variation on this theme is that operational researchers ought to be familiar with the origins of OR, the field's subsequent development, and some of the pioneers who propelled OR to prominence. Furthermore, those who teach OR should feel obliged to pass this heritage to their students.

A.A. Assad and S.I. Gass [1] have provided an informal history of OR from 1564 up to 2004. This chronologically ordered series of short sections provides historical notes, anecdotes, photos, and references, and makes truly delightful reading.

The same Assad-Gass team of *American* editors then produced "Profiles in Operations Research" [2] covering the profiles of 43 pioneers and innovators. From a *non-US* point of view, however, one may argue that the choice of profiles is somewhat biased towards the Western side of the Atlantic Ocean.

Further credence to this postulate is provided by Müller-Merbach in his extensive review [7] of "Profiles" where he wrote that: 30 of the 43 profilees are born in the US and have also worked there. Five others are born elsewhere but emigrated to the US where they terminated their professional careers. Six pioneers worked in the UK and four of them were native citizens. These numbers add up to 41. The remaining two comprise one Russian and one French. Müller-Merbach closed his observation with a good question: were no Japanese or German researchers worthy of inclusion?

Without excluding the possibility of potential pioneers outside the US and Europe I shall here focus on the 'missing Europeans' as it is my conviction that a European team of profile editors might have toned down the strong American dominance. To properly remedy this situation, it can be noted that Ron Howard and Bernard Roy, both born in 1934, are the youngest profilees in the Assad-Gass volume. Let, say, 1937 be an upper bound on the birth age of potential new candidates. This condition is satisfied by, amongst others, the following two, Jean-Pierre Brans and Hans-Jürgen Zimmermann, together considered as the founding fathers of EURO.

Hans initiated the birth of EURO [8], was obviously elected EURO's first president (1975-1978), - and has in addition an impressive research record. The person who was instrumental in giving further momentum to EURO, however, was J. Pierre who also has a record of comparable depth and impact.

In their Preface [2] Assad and Gass have noted that "… we were able to find authors who, as close colleagues or friends, were in the best position to relate the professional and personal histories of the person they profiled". When the 'missing Europeans' were considered by Cathal M. Brugha, President of Analytics Society of Ireland, and myself, we decided to give highest priority to J. Pierre as both of us have worked closely with him for decades with particular focus on promoting OR within the framework of EURO. The visible result of our endeavours is a profile [3] concluded in June 2017.

3 THE EARLY YEARS (1975-1988)

Along with the birth of EURO, four so-called *instruments* (though in this context a somewhat unorthodox term in English, coined by J. Pierre and still preserved) were launched: the Association itself, the EURO-k Conferences, the Working Groups (EWGs), and the EURO Bulletin. In addition, the first issue of the European Journal of Operational

Research (EJOR) appeared in January 1977. At the initiative of J. Pierre, three additional instruments were introduced in 1983: the EURO Summer Institutes (ESIs), the EURO Mini Conferences, and the EURO Gold Medal.

There is no need here to repeat what can be found on EURO's extensive and very informative website, https://www.euro-online.org/ covering not only the state-of-affairs (instruments, member societies, people, …) as of today but also the past (who did what and when). Thus, this section on 'the early years' will be reduced to but a short comment on my favourite instrument: the ESIs.

Among the instruments introduced by J. Pierre, possibly EURO's best idea in terms of impact, investment in the future, concern for the next generation targeted carefully selected groups of promising young scientists. Keeping this series alive and preserving its flavour is the best guarantor for the future.

Starting in the summer of 1984 EURO Summer/Winter Institutes (ESWI) have been organised regularly. Each student participant in an ESWI benefits from getting a personal network of academic soulmates that can extend up to 20 countries. How to preserve it and keep it further nurtured? Some do by establishing a EURO Working Group (EWG) on the last day of the ESWI meeting!

Who are heading today's OR departments at several European universities? Who are among today's main contributors to our professional journals? From where might EURO recruit its coming presidents, members of the Executive Committee (EC), other officers, and EWG chairpersons? Answer to all three questions: look at EURO's best idea during the first decade. The famous Madeira-ESI in 1989 on Decision Support Systems generating no less than two later presidents of EURO (Alexis Tsoukiàs and M. Grazia Speranza) is in this respect among the most blatant examples.

According to the rules, nobody can participate in an ESI more than once in her/his lifetime. This rule, however, applies to the "regular" participants only, whereas senior people acting as advisers, guest lecturers, et cetera may enjoy the privilege of several appearances, so far mounting up to eight in my case.

4 THE TURBULENT YEARS (1989-1990)

This tale of different parts of EURO's history is a patchwork of earlier writings mixed with personal recollections. Most of these writings have been generally available as opposed to the 18-pages long President's Report 1989-1990 [6] so far having the EURO Council + EC as its only readership. An abstract is provided on the front page:

The very last sentence of the "Presidential Report 1987-1988" by my predecessor as President of EURO, D. de Werra, reads: "I express my best wishes to Jakob Krarup who will inherit many open problems". Indeed a bold understatement! In retrospect, it can be said that the two-year period 1989-1990 actually marked a state of transition of EURO on many frontiers. Rather than concentrating on the scientific aspects of our Association, the Executive Committee had to devote considerable time to questions of administrative and financial nature as well as to problems of human relations. The present document accounts for the achievements made and sorts of 'set the stage' for EURO on the threshold to the nineties.

Enthusiasm, good will, and the dedication of many are important ingredients to run an association like EURO, but not enough. With the surplus made at the EURO-k conferences as the sole source of income, EURO was in the late 80s financially vulnerable. To enable EURO to carry out a wider range of activities in support of its objectives, different options for raising additional income were permanently on the agenda at several Executive Committee (EC) meetings. It appeared quite unrealistic to increase the fees for conferences or to levy a

557

sufficiently large annual subscription on member societies. A third possibility was then to investigate whether EURO could benefit financially from EJOR.

At EURO XXV (Vilnius, 2012) all past presidents were invited and lined up at the stage to be presented with an award. Three, however, were missing, Bernard Roy, Rolfe Tomlinson, and Maurice Shutler. In 2015 both Rolfe and Maurice sadly passed away. To acknowledge their many and varied contributions to EURO, a Memorial Session was organised at EURO XVIII (Poznan, 2016).

Co-authored with Graham Rand, Lancaster University, an obituary of Maurice [5] appeared in EJOR from which the following paragraphs are excerpted almost *ad verbatim:*

During my EURO presidency (1989-1990) all EC meetings were also attended by Maurice, EURO's President-Elect (1991-1992). It was a great benefit to have him on board: capable of identifying and focusing on the most important issues on long agendas, meticulous with details at times overlooked by others, a true tower of strength during, what in hindsight, were EURO's most turbulent years.

Drawing on his experience of the potential financial benefits of publications, he played a leading role in discussions that resulted in EURO receiving financial benefit from EJOR's publisher Elsevier. The EURO EC commissioned Maurice to obtain competitive quotations from other leading publishers. As a result, at the 1989 EURO conference in Belgrade, Council instructed the EC to transfer publication of EJOR from Elsevier, but expressed the wish that the editors should continue if they were willing. The editors had already become aware of the situation, however, and they backed Elsevier. At, what was described as, an acrimonious meeting of the EC with the editors in Brussels in October 1989, it was agreed that another attempt be made to maintain cooperation with Elsevier, provided that Elsevier would be willing to offer conditions comparable to those of the other publisher. After lengthy negotiations terminating with an ultra-short meeting in Schipol Airport, an agreement between Elsevier and EURO, valid for 10 years, was eventually signed by me on behalf of EURO on 15 March 1990, which allowed the editors to continue and included a payment by Elsevier to EURO from January 1, 1990. At the EURO Council held in Athens in June 1990, at the time of an IFORS conference, this agreement was unanimously approved. It was a great relief to me in [6] to report to the Council that it was believed "that the human relations have been properly repaired and that no frictions henceforth shall disturb the harmonious cooperation between EURO and its foremost instrument. In addition, the deal made with Elsevier has left EURO as a financially sound organization".

In addition to the abrasive discussions with Elsevier, the EURO X conference (Belgrade, 1989) turned out to be a disaster. The 295 participants were almost invisible in the huge Sava Centre capable of accommodating 4,000 persons.The first accounts indicated that EURO might lose a substantial part of its net assets, and some felt that the mere existence of the Association was heavily threatened. Thanks to the admirable efforts of Professor Radivoj Petrovic, however, EURO managed eventually to become financially stable due to the generosity of "Institut Mihajlo Pupin Beograd" which contributed considerably more than planned. End of excerpts from [5].

More than EURO were facing turbulent years during 1989-1990. The building of the 45.1 kilometres long Berlin Wall started on 13 August 1961 and kept East and West Berlin separated from one another until it was reopened in the evening of 9 November 1989.

The six Warsaw Pact countries of Eastern Europe - Bulgaria, Czechoslovakia, Hungary, Poland, Romania, and The German Democratic Republic (GDR) – while nominally independent, were widely recognized by the international community as the Soviet satellite states. All had been occupied by the Soviet Red Army in 1945, had a Soviet style socialist regime imposed on them, and had very restricted freedom of action in either domestic or international affairs. GDR and Poland withdrew from the Pact in 1990. On 25 February

1991, the Pact was declared at an end at a meeting of defence and foreign ministers from the remaining member states. In the same year the Soviet Union itself was dissolved on 26 December.

Already as of 1990, this surprising development prompted several post-communistic countries to "knock on the door" to become admitted to IFORS and, if they succeeded, then to EURO. Accordingly, the report on "new members" in [6] is unusually extensive and encompasses all the former satellite states. Talks were held, letters were exchanged, and obstetric aid was offered, notably to Hungary and Poland. Funnily enough, also the Icelandic OR Society (ICORS) managed to become member of EURO during that period. Apparently ICORS had believed it was a member already, but had forgotten to submit an application for membership. Upon a reminder, this eventually was done on 4 April 1990.

It is not an objective of EURO to take political measures and no such measures were ever taken. It is not forbidden, however, to express concern for those of our colleagues who, in one way or another, are victims of political conflicts. Once a group of such victims, known as *refusnik scientists,* i.e. Jewish scientists who had applied for and refused permission to leave the Soviet Union. Refusniks were normally demoted to low-level jobs or lost their scientific positions altogether. Some of them had their academic degrees removed retroactively; some were in labour camps or exile. In all cases they were cut off from normal interaction with other scientists.

Having participated in the *International Moscow Refusnik Seminar* in December 1988, held under abnormal conditions in private homes, I found it most appropriate at the Closing Session at EURO X (Belgrade, 1989) to choose the refusnik scientists as the subject of my address. Congratulatory notes were afterwards received from several EURO people whereas the Serbian chairman was really upset: "This was not the talk expected from the President of EURO". Three years later, however, at EURO XII (Helsinki, 1992) the situation was reversed. Delegates from Serbia were banned from participating and I found it equally appropriate in my address to let the thoughts go to friends and colleagues in the different parts of former Yugoslavia. The congratulatory note was this time received from the former Serbian chair and we have been best friends ever since.

1989-1990 in hindsight: a full-time job at DIKU together with an almost comparable workload for EURO. It was a great relief on 1 January 1991 to pass the torch to Jaap Spronk, from that day EURO's eighth president.

5 REACHING THE STARS (1991-DATE)

All questions as regards EURO's statutes, its legal status, and domicile occupied in the past the minds of many but have been settled by now. The dream of having a permanent secretariat came true when Philippe van Asbroeck in 1993 became Permanent Secretary until his untimely death in 2012. Many duties have been relieved from the shoulders of various EC-members and organisers of EURO events when Sarah Fores took office as Manager of EURO in 2011. Amongst others, she is also the Editor of the EURO Newsletters circulated regularly via email to all members who wish to subscribe via the EURO website.

'Reaching the stars': once again, let EURO's website, now accompanied by the series of Newsletters, speak for themselves. Keep *challenges* as the motive power but do also note the set of new keywords: *sustainable development, ethics*, and *history*.

6 EPILOGUE

I regret my absence when EURO was founded in January 1975 at the Hotel Sheraton, Brussels, in conjunction with the EURO I conference but 'joined the party' on 21st June

when the first EJOR Editorial Board Meeting was held in Paris. Thus, 42 years so far have been spent with our Association. Viewed as a whole: lots of work when called for but, by and large, a true *con amore* activity throughout.

The interest in devoting time and energy to EURO matters did in no way fade away upon my retirement in 2006. Most precious, however, is to preserve and further nurture the lasting friendships that have evolved with many EURO-*nists* along the way. Besides, I have appreciated the privilege of writing joint papers like [3,4,5], being member of various committees, attending EURO-k conferences with their by now flavour of family gatherings, meetings organized by the national OR societies, EURO Summer Schools or Working Groups. Coming events, looked forward to with great anticipation: EURO XXIX (Valencia, 2018) and EURO XXX (Dublin, 2019).

In closing, let me express my sincere thanks to Lidija Zadnik Stirn for the honourable invitation to become among the contributors to the SOR 2017 Proceedings.

References

[1] A.A. Assad, S.I. Gass, **An Annotated Timeline of Operations Research.** Kluwer Academic Publishers, 2005.

[2] A.A. Assad, S.I. Gass, **Profiles in Operations Research.** Springer Science + Business Media, 2011.

[3] C.M. Brugha, J. Krarup, "Jean-Pierre Brans – Portrait of a Fiery Soul" (2017). Submitted for publication.

[4] S. Fores, J. Krarup "On the origins of OR and its institutions", Invited Review, **Central European J. OR 21.2** (2013) 265-275.

[5] G. K. Rand, J. Krarup, "Maurice F. Shutler (1931-2015)", **EJOR 252** (2016) 699-700.

[6] J. Krarup, "EURO on the threshold to the nineties, President's Report 1989-1990". EURO, Fribourg, Switzerland (1992). Presented to the EURO Council at EURO XII (Helsinki, 1992).

[7] H. Müller-Merbach, "Publikationen: Profiles in Operations Research", **OR News 45** (2011) 56-58.

[8] H.-J Zimmermann, "The Founding of EURO, The Association of European Operational Research Societies within IFORS" **EJOR 87** (1995) 404-407.

IMPORTANCE OF HIGHER EDUCATION AND INVESTMENT IN HIGHER EDUCATION IN CESEE COUNTRIES

Blanka Škrabić Perić
Faculty of Economics, Department of Quantitative Methods
Cvite Fiskovića 5, 21 000 Split, Croatia
E-mail: bskrabic@efst.hr

Zdravka Aljinović
Faculty of Economics, Department of Quantitative Methods
Cvite Fiskovića 5, 21 000 Split, Croatia
E-mail: lil@efst.hr

Hrvoje Mamić
OTP - Splitska banka, Risk managament
Domovinskog rata 61, 21 000 Split, Croatia
E-mail: hmamic.44@gmail.com

Abstract: This paper investigates the importance of enrolment in higher education and investment in higher education in 11 CESEE countries, EU members, during the period from 1994 to 2015. The results of panel data analysis confirm that higher education plays significant role in growth process of CESSE countries. More precisely, the results of the augmented Solow model indicate a statistically significant influence of gross enrolment ratio in the tertiary education and expenditure on tertiary education on economic growth. Additionally, the results confirm positive influence of investment while the influence of population growth is not clear as in the similar research.

Keywords: higher education, CESEE countries, panel data estimators

1 INTRODUCTION

In the recent time the higher education is recognized as a significant contributor to economic growth, prosperity and social cohesion in European countries. On the other hand, in the empirical studies, higher education is rarely considered. Previous empirical studies tried to prove relationship between overall education and economic growth. They used indicators of quality education or average years of schooling in the countries. Additionally, most empirical studies included developed and developing countries around the world. The obtained results are sensitive regard to observed countries, education indicators, methods and used time periods. Some studies [4,10,11], in growth model included the indicator of secondary education as variable of human capital. The importance of secondary education is almost always confirmed.

Among rare studies which consider the importance of higher education, we can find contradictory results. While studies [10] and [11] did not show a significant influence of higher education, studies [5] and [12] confirmed a significant influence of higher education. Moreover, Chatterji [5] indicated stronger influence of higher education than secondary education. Additionally, he noted that it may be the case that higher education is more important determinant for developing countries, but less important for developed economies where tertiary enrolment reached its maximum.

His conclusion motivates us to investigate influence of higher education on economic growth in Central, Eastern and South-Eastern European (CESSE) countries. Since nineties these countries passed through transformations and reforms in all sectors and they also covered higher education. Hegerty [9] indicated that these countries improved living standard in the last decades, while their economies did not reach EU averages. Additionally,

561

Machková [13] indicated that the public expenditure on education is lower than in most successful European countries.

Therefore, it can be interesting to investigate influence of higher education on economic growth in these countries. In the other words, this paper investigate is it possible that higher education is one of the main factors in achieving economic growth as in developed countries. Results of [6], give additional support to our research idea. Grdinic [6] found that expenditure on education, tertiary educated workforce and number of researchers have positive impact of GDP growth in the simple linear regression. This paper goes one step further, it investigates influence of higher education enrolment and government expenditure on higher education by estimating endogenous growth model precisely augmented Solow model.

After Introduction, second section provides data and methodology. Results and economic interpretation are presented in section Result and discussion. Finally, last section summarizes the results and gives policy implications.

2 DATA AND METHODOLOGY

Data set consists of 11 CESEE EU members[1] over the period 1994-2015. Data are obtained from World Bank Database. Recent literature, which deals with education and economic growth, uses different versions of growth model. This paper almost follows model in [4]. They used augmented Solow model which beside investment rate and population growth includes variable of human capital. They used secondary education enrolment ratio as measure of human capital while this paper uses indicators of higher education.

Investment rate is measured by gross capital formation as percentage of gross domestic product. Considering the fact that population growth for some countries in some periods is negative, instead regular population growth, we use ratio of total population for country i in the period t to total population for country i in year 2010 multiplied by 100. For human capital variable, two indicators of higher education are used. The first indicator is Gross enrolment ratio in tertiary education. The second indicator Expenditure on tertiary education, defined as government expenditure per tertiary student as percentage of GDP per capita. Therefore, growth model can be written by equation:

$$\triangle GDPpc_{it} = \gamma GDPpc_{i,t-1} + \beta_1 INV_{it} + \beta_2 POP_{it} + \beta_3 EDU_{it} + \alpha_i + \varepsilon_{it},$$
$$i = 1,\ldots,N \text{ and } t = 1,\ldots,T, \tag{1}$$

where $\triangle GDPpc_{it}$ is the first difference of the logarithm of GDP per capita for country i in the period t, $GDPpc_{i,t-1}$ is the logarithm of GDP per capita for country i in the period t-1, INV_{it} is the logarithm of investment rate for country i in the period t, POP_{it} is the logarithm of population growth for country i in the period t and EDU_{it} is logarithm value of higher education indicator for country i and period t. α_i is the fixed or random effect while remainder disturbance ε_{it} is IID$(0, \sigma_\varepsilon^2$). N is number of countries while T is number of time periods. γ, β_1, β_2 and β_3 are coefficients.

To estimate equation (1) it is necessary to choose adequate estimator (Pooled Model (PM), Fixed Effects Model (FE) or Random Effects Model (RE)) by using diagnostic tests. F test is performed to choose is FE more adequate than PM. The second test is LM test and it tests is the RE more appropriate than PM. If results indicate that FE and RE are more appropriate than PM, Hausman test is performed. Result of Hausman test gives answer is FE or RE the

---

[1] Bulgaria, the Czech Republic, Croatia, Estonia, Latvia, Lithuania, Hungary, Poland, Romania, Slovakia, Slovenia

most adequate estimator. After choosing the adequate estimator, Wooldridge test is performed to detect existence of first order of autocorrelations between residuals.

If Wooldridge test indicates problem of autocorrelation, then following equation is estimated to capture this problem:

$$\triangle GDPpc_{it} = \gamma GDPpc_{i,t-1} + \beta_1 INV_{it} + \beta_2 POP_{it} + \beta_3 EDU_{it} + \alpha_i + \varepsilon_{it},$$
$$i = 1,\ldots,N \text{ and } t = 1,\ldots,T \tag{2}$$

where ε_{it} following a stationary AR(1) process

$$\varepsilon_{it} = \rho\varepsilon_{i,t-1} + u_{it}, \tag{3}$$

where u_{it} is IID($0, \sigma_\varepsilon^2$). These estimators are called Estimators with Autocorrelated Disturbances (AR(1)) and they can contain fixed or random effects [2].

However, in the recent empirical papers [4], [7] and [14], growth regression is very often rewritten as:

$$GDPpc_{it} = (\gamma + 1) GDPpc_{i,t-1} + \beta_1 INV_{it} + \beta_2 POP_{it} + \beta_3 EDU_{it} + \alpha_i + \varepsilon_{it},$$
$$i = 1,\ldots,N \text{ and } t = 1,\ldots,T. \tag{4}$$

This form of growth regression is adjusted to use the most popular dynamic panel data estimators difference GMM (AB) [1] or system GMM (BB) [3]. They allow endogenous regressors in the model and they enable test of endogeneity of model (Sargan test). But, these two estimators are suitable for data set with large N and small number of time periods T what is not case in our research. Our data set consists of 11 countries and 22 time periods. Therefore, these two estimators are not adequate for our sample. One of these estimators will serve as a robustness check of our results and to investigate possible endogeneity in our model.

3 RESULTS AND DISCUSION

The first version of our model for the variable of higher education uses Gross enrolment ratio in tertiary education (GET_{it}). The results of diagnostic tests and results of estimations are shown in the Table 1. From diagnostic tests is evident that FE is the most adequate estimator. Result of F test (p value 0.001) indicates that FE is more adequate than PM while LM test (p value 0.4717) indicates that PM is more adequate than RE. Therefore, Hausman test is not conducted. Results of FE are shown in the column one. However, results of this model are not relevant because of existence of autocorrelation between residuals (results of Wooldridge test p value 0.0001). Therefore, Fixed effect model with AR(1) disturbances is estimated in column (2) an it presents our main results while all other columns serve as additional robustness checks.

The results from column (2) indicate that GDP$pc_{i,t-1}$ has negative sign and it is statistically significant. Investment has positive and statistically significant influence while population has positive sign, but it is statistically significant at 10 percent. For robustness checks equations (3) and (4) are estimated by AB estimator. In column (3), all independent variables are treated as exogenous. Results of Sargan test (p value 0.9127) did not indicate problem of endogeneity. However, in economic theory exists possibility of endogeneity of independent variables. Therefore, in column (4), all independent variables are treated as endogenous. Results are almost in the line with results from column (2). Results indicate robust influence of all independent variables except population growth.

Table 1: Results of growth model with Gross enrolment ratio in tertiary education as variable of human capital

	(1)	*(2)*	*(3)*	*(4)*
	$\triangle GDPpc_{it}$	$\triangle GDPpc_{it}$	b $\triangle GDPpc_{it}$	$\triangle GDPpc_{it}$
$GDPpc_{i,t-1}$	-0.166***	-0.118***	-0.3428***	-0.389***
	a(0.0264)	(0.0240)	(0.0650)	(0.0720)
INV_{it}	0.0377***	0.0338***	0.0511**	0.0715***
	(0.00774)	(0.00820)	(0.0233)	(0.0256)
POP_{it}	-0.310***	0.0956*	-1.341**	-1.286***
	(0.111)	(0.0489)	(0.601)	(0.438)
GET_{it}	0.0625***	0.0647***	0.123***	0.153***
	(0.0148)	(0.0213)	(0.0378)	(0.0427)
const	2.644***	0.323**	8.773***	8.767***
	(0.678)	(0.161)	(3.104)	(2.263)
N	212	201	200	200
Method	FE	FE with AR(1)	AB	AB with endogenous variables
R^2	0.267	0.174		
F test(p value)	0.0001	0.0045		
LM test(p value)	0.4717			
Hausman test(p value)				
Wooldrige test(p value)	0.0001			
Sargan test(p value)			0.9127	1
AR(2) test(p value)			0.1107	0.2460

Note:*, **, *** indicates statistical significance at the level of 10%, 5% and 1%;
avalues in parentheses are standard errors
bAfter estimation using AB estimator, equation (4) is rewritten as equation (1) and coefficient γ of $GDPpc_{i,t-1}$ in columns (3) and (4) is calculated by the Delta method.
Source: Authors' calculations.

In Table 2, Expenditure on tertiary education (ET_{it}) is considered as variable of human capital. The results of diagnostic tests and results of estimations are shown in the Table 2. It is evident that FE is the most adequate estimator[2]. Wooldridge test indicates existence of autocorrelation between residuals. Fixed effects model with AR(1) disturbance is used as the most adequate estimator. The main results are in column (2) of Table 2 while other columns serve as robustness checks.

The results from column (2) indicate that GDP$pc_{i,t-1}$ is negative and statistically significant. Investment has positive and statistically significant influence while population has positive sign, but is statistically significant at 10 percent. For robustness check we estimate equations (3) and (4) by AB estimator[3]. Results are almost in the line with results from column (2).

---

[2] Results are in the line with results of diagnostic tests from column (1) of Table 1.
[3] Same logic as for Table 1, columns (3) and (4)

Table 2: Results of growth model with Expenditure on tertiary education as variable of human capital

	(1)	*(2)*	*(3)*	*(4)*
	$\triangle GDPpc_{it}$	$\triangle GDPpc_{it}$	[b] $\triangle GDPpc_{it}$	$\triangle GDPpc_{it}$
$GDPpc_{i,t-1}$	-0.118***	-0.138***	-0.3499***	-0.0456***
	[a](0.0223)	(0.0233)	(0.0789)	(0.10981)
INV_{it}	0.146***	0.190***	0.187***	0.222***
	(0.0186)	(0.0213)	(0.0236)	(0.0295)
POP_{it}	-0.00983	0.0839*	-0.240	-0.295
	(0.143)	(0.0434)	(0.281)	(0.411)
ET_{it}	0.100***	0.0811**	0.0780	0.179***
	(0.0312)	(0.0329)	(0.0504)	(0.0512)
const	0.412	0.0806*	1.144	0.797
	(0.815)	(0.0430)	(1.475)	(2.233)
N	142	131	124	124
R^2	0.463	0.508		
Method	FE	FE with AR(1)	AB	AB with endogenous variables
F test(p value)	0.0001	0.0005		
LM test(p value)	1.000			
Hausman test(pvalue)				
Wooldrige test(p value)	0.0000			
Sargan test(p value)			1	0.7729
AR(2) test(p value)			0.0741	0.0847

Note:*, **, *** indicates statistical significance at the level of 10%, 5% and 1%;

[a]values in parentheses are standard errors

[b]After estimation using AB estimator , equation (4) is rewritten as equation (1) and coefficient γ of $GDPpc_{i,t-1}$ in columns (3) and (4) is calculated by the Delta method

Source: Authors' calculations.

The results from Table 1 and Table 2 are almost uniform. Negative sign of lagged value GDP per capita suggests conditional convergence. Countries with larger GDP per capita can expect smaller growth. Investment has positive and statistically significant influence as it is expected and confirmed in existing studies. The results of the variable population change direction and significance in different model specifications. However, absence of robust effect of total population growth on economic growth is well known fact in the recent times [8]. Finally, results from Table 1 and Table 2 confirm positive effect of gross enrolment ratio in tertiary education and expenditure for tertiary education. Our results confirm Chatterji's conclusion [5] that developing countries can use higher education to increase economic growth. They have to increase investment in higher education especially because their current investment in the higher education is significantly lower than investment of the most successful European countries [13].

4 CONCLUSION

This paper investigates role of higher education enrolment and government expenditure on higher education on economic growth in CESEE countries during the period from 1994 to 2015. The results of research indicate negative influence of lagged GDP per capita, positive influence of investment and unclear influence of population growth.

However, results of research confirm importance of both indicators of higher education for economic growth in CESEE countries. Therefore, our results confirm view that higher education is important contributor of economic growth in these countries.

Our results provide several policy implications. Governments of CESEE countries have to use potential of higher education to increase economic growth. However, this potential is not indefinitely, but it exists regard to higher education enrolment and investment in the higher education of world leaders. Second, governments of these countries have to promote policy which will attract young people to universities. They have to increase expenditure per student in higher education to ensure better student standards and quality of higher education. That will produce higher quality labour force with higher productivity and consequently higher economic growth.

References

[1] Arellano, M., Bond, S. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. The Review of Economic Studies, 58(2):277–297.

[2] Baltagi, B. H., Li, Q. 1991. A transformation that will circumvent the problem of autocorrelation in an error-component model. Journal of Econometrics, 48(3): 385–393.

[3] Blundell, R., Bond, S. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1): 115–143.

[4] Bond, S. R., Hoeffler, A., Temple, J. R. W. 2001. GMM Estimation of Empirical Growth Models Centre for Economic Policy Research Discussion Paper No. 3048 London

[5] Chatterji, M. 1998. Tertiary Education and Economic Growth. Regional Studies, 32(4), 349–354.

[6] Grdinic. 2014. Higher Education as a Means of Achieving Economic Growth and Development - A Comparative Analysis of Selected EU and Former Soviet Union Countries. Medunarodna Revija za Javno Upravo, 12(4), 93–110.

[7] Hauk, W., Wacziarg, R. 2009. A Monte Carlo study of growth regressions. Journal of Economic Growth, 14(2), 103–147.

[8] Headey, D. D., Hodge, A. 2009. The Effect of Population Growth on Economic Growth: A Meta-Regression Analysis of the Macroeconomic Literature. Population and Development Review, 35(2), 221–248.

[9] Hegerty S.W. 2016. Regional Convergence and Growth Clusters in Central and Eastern Europe: An Examination of Sectoral-Level Data. Eastern European Business and Economics Journal, 2(2), 95–110.

[10] Holmes, C. 2013. Has the Expansion of Higher Education Led to Greater Economic Growth? National Institute Economic Review, 224(1): 29–47.

[11] Keller, K. R. I. 2006. Investment in Primary, Secondary, and Higher Education and the Effects on Economic Growth. Contemporary Economic Policy, 24(1): 18–34.

[12] Kubík, R. 2015. What is the Real Effect of Schooling on Economic Growth? Prague Economic Papers, 24(2): 125–135.

[13] Machková, H. 2016. Higher Education in Central Europe and its Impact on Countries' Competitiveness. Central European Business Review, 2016(1): 62–68.

[14] Soto, M. 2009. System GMM estimation with a small sample. Barcelona Graduate School of Economics Working Paper No. 395. Barcelona

The 14th International Symposium on
Operational Research in Slovenia

SOR '17

Bled, SLOVENIA
September 27 - 29, 2017

Appendix
Authors' addresses

Addresses of SOR'17 Authors

(The 14th International Symposium on OR in Slovenia, Bled, SLOVENIA, September 27 – 29, 2017)

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
1.	Ana	Aleksić	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedy 6	10000	Zagreb	Croatia	aaleksic@efzg.hr
2.	Zdravka	Aljinović	University of Split, Faculty of Economics	Cvite Fiskovića 5	21000	Split	Croatia	lil@efst.hr
3.	Željana	Aljinović Barać	University of Split, Faculty of Economics	Cvita Fiskovića 5	21000	Split	Croatia	zbarac@efst.hr
4.	Hemmak	Allaoua	Department of Computer Science, University of M'sila		28000	M'Sila	Algeria	hem_all@yahoo.fr
5.	Josip	Arnerić	University of Zagreb, Faculty of Economics and Business Zagreb	Trg J.F.Kennedyja 6	10000	Zagreb	Croatia	jarneric@efzg.hr
6.	Zoran	Babić	University of Split, Faculty of Economics	Cvite Fiskovića 5	21000	Split	Croatia	babic@efst.hr
7.	Daria	Battini	Department of Management and Engineering, University of Padova	Stradella San Nicola 3	36100	Vicenza	Italy	daria.battini@unipd.it
8.	Nina	Begičević Ređep	Faculty of Organisation and Informatics	Pavlinska 2	42000	Varaždin	Croatia	nbegicev@foi.hr

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
9.	Jani	Bekő	Faculty of Economics and Business, University of Maribor, Department of Political Economy	Razlagova 14	2000	Maribor	Slovenia	jani.beko@um.si
10.	Serge	Bogaerts	PRACE	Rue du Trône 98	1050	Ixelles, Brussels	Belgium	S.Bogaerts@staff.prace-ri.eu
11.	David	Bogataj	Department of Management and Engineering, University of Padova	Stradella San Nicola 3	36100	Vicenza	Italy	david.bogataj@unipd.it
12.	Marija	Bogataj	CERRISK - Zavod INRISK	Vrtača 9	1000	Ljubljana	Slovenia	marija.bogataj@ guest.arnes.si
13.	Marko	Bohanec	Jožef Stefan Institute, Department of Knowledge Technologies	Jamova cesta 39	1000	Ljubljana	Slovenia	marko.bohanec@ijs.si
14.	Marko	Bohanec	Salvirt Ltd.	Dunajska cesta 136	1000	Ljubljana	Slovenia	marko.bohanec@ salvirt.com
15.	Drago	Bokal	University of Maribor, Faculty of Natural Sciences and Mathematics	Koroška cesta 160	2000	Maribor	Slovenia	d@bokal.net
16.	Amadeja	Bratuša	University of Maribor, Faculty of Natural Sciences and Mathematics	Koroška cesta 160	2000	Maribor	Slovenia	bratusa.amadeja@ gmail.com
17.	Alenka	Brezavšček	University of Maribor, Faculty of Organizational Sciences	Kidričeva cesta 55a	4000	Kranj	Slovenia	alenka.brezavscek@ fov.uni-mb.si
18.	Andrej	Brodnik	University of Ljubljana, Faculty of Computer and Information Science	Večna pot 113	1000	Ljubljana	Slovenia	andrej.brodnik@ fri.uni-lj.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
19.	Ufuk	Bolukbas	Yildiz Technical University, Vocational School	Gaziosmanpaşa		Istanbul	Turkey	bolukbas@yildiz.edu.tr
20.	Luboš	Buzna	University of Žilina			Žilina	Slovak Republic	lubos.buzna@ fri.uniza.sk
21.	Martina	Calzavara	Department of Management and Engineering, University of Padova	Stradella San Nicola 3	36100	Vicenza	Italy	
22.	Erkan	Celik	Munzur University, Department of Industrial Engineering		62000	Tunceli	Turkey	erkancelik@ munzur.edu.tr
23.	Zeynel Abidin	Çil	Department of Manufacturing Engineering, University of Batman		72060	Batman	Turkey	cilzeynelabidin@ gmail.com
24.	Adrienn	Csizmadia	Budapest University of Technology and Economics, FICO Plc.				Hungary	adriennagy@gmail.com
25.	Zsolt	Csizmadia	Budapest University of Technology and Economics, FICO Plc.				Hungary	zsolt.csizmadia@ gmail.com
26.	Valmir Ferreira	da Cruz	Universidade Nove de Julho/Industrial Engineering Graduate Program	Francisco Matarazzo, Av. 612		São Paulo	Brazil	valmir.vfc@gmail.com
27.	Peter	Czimmermann	University of Žilina			Žilina	Slovak Republic	peter.czimmermann@ fri.uniza.sk
28.	Borut	Čampelj	University of Maribor, Faculty of Organisational Sciences	Kidričeva 55	4000	Kranj	Slovenia	borut.campelj@gov.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
29.	Vesna	Čančer	University of Maribor, Faculty of Economics and Business	Razlagova 14	2000	Maribor	Slovenia	vesna.cancer@um.si
30.	Anita	Čeh Časni	University of Zagreb, Faculty of Economics and Business, Department of Statistics	Trg J.F. Kennedy 6	10000	Zagreb	Croatia	aceh@efzg.hr
31.	Draženka	Čizmić	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	dcizmic@efzg.hr
32.	Mirjana	Čižmešija	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	mcizmesija@efzg.hr
33.	Mirko	Čorić	University of Split/Faculty of Maritime Studies	Ruđera Boškovića 37	21 000	Split	Croatia	mcoric@pfst.hr
34.	Zsolt	Darvay	Budapest University of Technology and Economics, Babes-Bolyai University			Budapest	Hungary	darvay@cs.ubbcluj.ro
35.	Leyla	Demir	Pamukkale University, Department of Information Processing Center	Kinikli Campus		Denizli	Turkey	ldemir@pau.edu.tr
36.	Blaženka	Divjak	Faculty of Organisation and Informatics	Pavlinska 2	42000	Varaždin	Croatia	bdivjak@foi.hr
37.	Samo	Drobne	University of Ljubljana, Faculty of Civil and Geodetic Engineering	Jamova cesta 2	1000	Ljubljana	Slovenia	samo.drobne@ fgg.uni-lj.si
38.	Ksenija	Dumičić	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedy 6	10000	Zagreb	Croatia	kdumicic@efzg.hr
39.	Igor	Đukanović	University of Maribor	Razlagova 14	2000	Maribor	Slovenia	igor.djukanovic@um.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
40.	Gencer	Erdogan	SINTEF Digital	P.O. Box 124 Blindern	0314	Oslo	Norway	gencer.erdogan@ sintef.no
41.	Nataša	Erjavec	Faculty of Economics and Business, University of Zagreb, Croatia, Department of Statistics	Trg J. F. Kennedyja 6	10000	Zagreb	Croatia	nerjavec@efzg.hr
42.	Sarah	Fores		3 Woodlea Court, Meanwood	LS6 4SL	Leeds	UK	sarahfores@gmail.com
43.	Lýdia	Gábrišová	University of Žilina, Faculty of Management and Informatics	Univerzitná 1	010 26	Žilina	Slovak Republic	lydia.gabrisova@ fri.uniza.sk
44.	Elif	Garajová	Charles University, Faculty of Mathematics And Physics, Department of Applied Mathematics	Malostranské nám. 25	11800	Prague	Czech Republic	elif@kam.mff.cuni.cz
45.	Mónica	García-Melón	Ingenio (CSIC-UPV), Universitat Politècnica de València		46022	Valencia	Spain	mgarciam@dpi.upv.es
46.	Helena	Gaspars-Wieloch	Poznan Universtity of Economics and Business, Department of Operations Research	Al. Niepodleglosci 10	61-875	Poznan	Poland	helena.gaspars@ ue.poznan.pl
47.	Martin	Golasowski	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 33	Ostrava	Czech Republic	martin.golasowski@vsb. cz
48.	Tomás	Gómez-Navarro	Energy Engineering Institute (IIE), Universitat Politècnica de València		46022	Valencia	Spain	tgomez@dpi.upv.es
49.	Hannia	Gonzalez-Urango	Ingenio (CSIC-UPV), Universitat Politècnica de València		46022	Valencia	Spain	hangonur@ doctor.upv.es

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
50.	Dorota	Górecka	Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management, Department of Econometrics and Statistics	Ul. Gagarina 13A	87-100	Toruń,	Poland	dgorecka@umk.pl
51.	Anja	Goričan	University of Maribor, Faculty of Natural Sciences and Mathematics	Koroška cesta 160	2000	Maribor	Slovenia	gorican.a@gmail.com
52.	Ekaterina	Grakova	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 00	Ostrava	Czech Republic	ekaterina.grakova@vsb.cz
53.	Petra	Grošelj	University of Ljubljana, Biotechnical Faculty	Jamnikarjeva 101	1000	Ljubljana	Slovenia	petra.groselj@bf.uni-lj.si
54.	Tatiana V.	Gruzdeva	Matrosov Institute for System Dynamics & Control Theory SB RAS	Lermontov st. 134	664033	Irkutsk	Russia	gruzdeva@icc.ru
55.	Anita	Gudelj	University of Split/Faculty of Maritime Studies	Ruđera Boškovića 37	21 000	Split	Croatia	anita@pfst.hr
56.	Muhammet	Gul	Munzur University, Department of Industrial Engineering		62000	Tunceli	Turkey	muhammetgul@munzur.edu.tr
57.	Alev Taskin	Gumus	Yildiz Technical University, Department of Industrial Engineering	34349 Besiktas		İstanbul	Turkey	ataskin@yildiz.edu.tr
58.	Roman	Gumzej	Faculty of Logistics, University of Maribor	Mariborska cesta 7	3000	Celje	Slovenia	roman.gumzej@um.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
59.	Ali Fuat	Guneri	Yildiz Technical University, Department of Industrial Engineering	Yıldız	34349	Istanbul	Turkey	guneri@yildiz.edu.tr
60.	Gregory	Gurevich	Department of Industrial Engineering and Management, SCE - Shamoon College of Engineering	Bialik Sts. 56, P.O. Box 950	84100	Beer Sheva	Israel	gregoryg@sce.ac.il
61.	Yossi	Hadad	Industrial Engineering and Management Department, SCE - Shamoon College of Engineering	Bialik Sts. 56, P.O. Box 950	84100	Beer Sheva	Israel	yossi@sce.ac.il
62.	Adela	Has	University of Josip Juraj Strossmayer in Osijek, Faculty of Economics	Gajev trg 7	31000	Osijek	Croatia	adela.has@efos.hr
63.	Milan	Hladík	Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics	Malostranské nám. 25	11800	Prague	Czech Republic	hladik@kam.mff.cuni.cz
64.	Eloy	Hontoria	Technical University of Cartagena, Business Management Department		30202	Cartagena	Spain	eloy.hontoria@upct.es
65.	Domen	Hudoklin	University of Ljubljana, Faculty of Electrical Engineering	Tržaška c. 25	1000	Ljubljana	Slovenia	domen.hudoklin@ fe.uni-lj.si
66.	Tibor	Illés	Budapest University of Technology and Economics, FICO Plc.			Budapest	Hungary	illes@math.bme.hu

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
67.	Saša	Jakšić	Faculty of Economics and Business, University of Zagreb, Croatia, Department of Statistics	Trg J. F. Kennedyja 6	10000	Zagreb	Croatia	sjaksic@efzg.hr
68.	Jaroslav	Janáček	University of Žilina, Faculty of Management and Informatics	Univerzitná 1	010 26	Žilina	Slovak Republic	jaroslav.janacek@ fri.uniza.sk
69.	Marta	Janáčková	University of Žilina, Department of Applied Mathematics	Univerzitna 8215/1	01026	Žilina	Slovak Republic	marta.janackova@ fstroj.uniza.sk
70.	Peter	Jankovič	University of Žilina, Faculty of Management Science and Informatics	Univerzitná 1	010 26	Žilina	Slovak Republic	peter.jankovič@ fri.uniza.sk
71.	Ľudmila	Jánošíková	University of Žilina, Faculty of Management Science and Informatics	Univerzitná 1	010 26	Žilina	Slovak Republic	ludmila.janosikova@ fri.uniza.sk
72.	Eva	Jereb	University of Maribor, Faculty of Organisational Sciences	Kidričeva 55	4000	Kranj	Slovenia	eva.jereb@fov.uni-mb.si
73.	Elza	Jurun	University of Split, Faculty of Economics Split, Department of Quantitative Methods	Cvite Fiskovića 5	21000	Split	Croatia	elza@efst.hr
74.	Nikola	Kadoić	Faculty of Organisation and Informatics	Pavlinska 2	42000	Varaždin	Croatia	nkadoic@foi.hr
75.	Igor	Karnet	University of Maribor, Faculty of Organisational Sciences	Kidričeva 55	4000	Kranj	Slovenia	igor.karnet@gmail.com
76.	Alenka	Kavkler	Faculty of Economics and Business, University of Maribor, Department of Quantitative Economic Analysis	Razlagova 14	2000	Maribor	Slovenia	alenka.kavkler@ uni-mb.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
77.	Marta	Kavšek	Fakulteta za organizacijske študije,	Ulica talcev 3	8000	Novo mesto	Slovenia	marta.kavsek@ gmail.com
78.	Klemen	Kenda	Jožef Stefan Institute	Jamova ulica 39	1000	Ljubljana	Slovenia	klemen.kenda@ijs.si
79.	Baruch	Keren	Industrial Engineering and Management Department, SCE - Shamoon College of Engineering	Bialik Sts. 56, P.O. Box 950	84100	Beer Sheva	Israel	baruchke@sce.ac.il
80.	Mirjana	Kljajić Borštnar	University of Maribor, Faculty of Organizational Sciences	Kidričeva 55a	4000	Kranj	Slovenia	mirjana.kljajic@ fov.uni-mb.si
81.	Davorin	Kofjač	University of Maribor, Faculty of Organizational Sciences	Kidričeva cesta 55a	4000	Kranj	Slovenia	davorin.kofjac@ fov.uni-mb.si
82.	Michal	Koháni	University of Žilina, Faculty of Management Science and Informatics	Univerzitna 8215/1	01026	Žilina	Slovak Republic	michal.kohani@ fri.uniza.sk
83.	Vedran	Kojić	University of Zagreb, Department of Mathematics and Department of Managerial Economics	Trg J. F. Kennedyja 6	10000	Zagreb	Croatia	vkojic@efzg.hr
84.	Tomaž	Kokolj					Slovenia	tomaz.kokolj@ gmail.com
85.	Lana	Kordić	University of Split, Faculty of Economics	Cvite Fiskovića 5	21000	Split	Croatia	lana.kordic@efst.hr
86.	Danijel	Kovačić	MEDIFAS	Mednarodni prehod 6	5290	Šempeter pri Gorici	Slovenia	kovacic.danijel@ gmail.com

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
87.	Renata	Kožul Blaževski	University of Split / University Department of Professional Studies	Kopilica 5	21000	Split	Croatia	rkozulb@oss.unist.hr
88.	Sasho	Kjosev	University "Ss. Cyril and Methodius", Faculty of Economics	Blvd. Goce Delchev 9V	1000	Skopje	Republic of Macedonia	skosev@ eccf.ukim.edu.mk
89.	Mehmet Ulaş	Koyuncuoğlu	Pamukkale University, Department of Information Processing Center	Kinikli Campus		Denizli	Turkey	ulas@pau.edu.tr
90.	Blaženka	Knežević	University of Zagreb, Faculty Economics and Business, Department of Trade	Trg J.F. Kennedya 6	10000	Zagreb	Croatia	bknezevic@efzg.hr
91.	Jakob	Krarup	DIKU, Department of Computer Science, University of Copenhagen	Ydervang 4	DK-3460	Birkerød	Denmark	krarup@di.ku.dk
92.	Maja	Krčum	University of Split/Faculty of Maritime Studies	Ruđera Boškovića 37	21 000	Split	Croatia	mkrcum@pfst.hr
93.	I. P.	Krommyda	Independent researcher		45333	Ioannina	Greece	ikrommyd@gmail.com
94.	Nataša	Kurnoga	University of Zagreb, Faculty Economics and Business, Department of Statistics	Trg J.F. Kennedya 6	10000	Zagreb	Croatia	nkurnoga@efzg.hr
95.	Marek	Kvet	University of Žilina, Faculty of Management Science and Informatics	Univerzitná 1	010 26	Žilina	Slovak Republic	marek.kvet@ fri.uniza.sk
96.	A. G.	Lagodimos	University of Piraeus, Department of Business Administration		18534	Piraeus	Greece	alagod@unipi.gr

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
97.	Mitja	Lakner	University of Ljubljana, Faculty of Civil and Geodetic Engineering	Jamova cesta 2	1000	Ljubljana	Slovenia	mitja.lakner@ fgg.uni-lj.si
98.	Ulrike	Leopold-Wildburger	University of Graz	Universitaetsstrasse 15	8010	Graz	Austria	ulrike.leopold@ uni-graz.at
99.	Mario	Lešnik	University of Maribor, Faculty of agriculture and life science	Pivola 11	2311	Hoče	Slovenia	mario.lesnik@um.si
100.	Iván	Ligardo-Herrera	Energy Engineering Institute (IIE), Universitat Politècnica de València		46022	Valencia	Spain	ivliher@doctor.upv.es
101.	Zrinka	Lukać	University of Zagreb, Department of Mathematics	Trg J.F. Kennedya 6	10000	Zagreb	Croatia	zlukac@efzg.hr
102.	Hrvoje	Mamić	OTP - Splitska banka, Risk managament	Domovinskog rata 61	21 000	Split	Croatia	hmamic.44@gmail.com
103.	Jasmina	Mangafić	University of Sarajevo, School of Economics and Business Sarajevo	Trg oslobođenja - Alija Izetbegović 1	71000	Sarajevo	Bosnia and Herzegovina	jasmina.mangafic@ efsa.unsa.ba
104.	Branka	Marasović	University of Split, Faculty of Economics	Cvita Fiskovića 5	21000	Split	Croatia	branka.marasovic@ efst.hr
105.	Vladislav	Maraš	University of Belgrade, Faculty of Transport and Traffic Engineering	Vojvode Stepe 305	11000	Belgrade	Serbia	v.maras@sf.bg.ac.rs
106.	Jan	Martinovič	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 00	Ostrava	Czech Republic	jan.martinovic@vsb.cz

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
107.	Tomáš	Masařík	Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics	Malostranské nám. 25	11800	Prague	Czech Republic	masarik@ kam.mff.cuni.cz
108.	Ivana	Matić	Imex Bank Split	Tolstojeva 6	21000	Split	Croatia	ivana.matic@ imexbanka.hr
109.	Marina	Matošec	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	mmatosec@efzg.hr
110.	Süleyman	Mete	Department of Industrial Engineering, Munzur University		62000	Tunceli	Turkey	suleyman489@ gmail.com
111.	Lorena	Mihelač	School center Novo mesto, IT department,		8000	Novo mesto	Slovenia	lorena.mihelac@ sc-nm.si
112.	Dunja	Mladenić	Jožef Stefan International Postgraduate School	Jamova ulica 39	1000	Ljubljana	Slovenia	dunja.mladenic@ijs.si
113.	Helena	Nikolic	University of Zagreb, Faculty of Economics & Business, Department of Trade	J.F. Kennedy 6	10 000	Zagreb	Croatia	hmiloloza@efzg.hr
114.	Blagica	Novkovska	University of Tourism and Management, Faculty of Economics	Blvd. Partizanski Odredi No. 99	1000	Skopje	Republic of Macedonia	blagica@novkovski.com
115.	Jana	Novotná	Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics	Malostranské nám. 25	11800	Prague	Czech Republic	janca@kam.mff.cuni.cz
116.	Bjørn	Nygård	Statoil ASA	Forusbeen 50	4035	Stavanger	Norway	bjnyg@statoil.com

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
117.	Phillips Edomwonyi	Obasohan	Department of Liberal Studies, College of Administrative and Business Studies, CABS, Niger State, Polytechnic, Bida Campus				Nigeria	philiobas@yahoo.com
118.	Domen	Ocepek	Kopa d. d.	Kidričeva 14	2380	Slovenj Gradec	Slovenia	domen.ocepek@kopa.si
119.	Furkan	Ömerustaoğlu	Yildiz Technical University, Department of Industrial Engineering	34349 Besiktas		İstanbul	Turkey	furkanomerustaoglu@ gmail.com
120.	Andrei	Orlov	Matrosov Institute for System Dynamics & Control Theory SB RAS	Lermontov str., 134	664033	Irkutsk	Russia	anor@icc.ru
121.	Nidžara	Osmanagić Bedenik	University of Zagreb, Department of Mathematics and Department of Managerial Economics	Trg J. F. Kennedyja 6	10000	Zagreb	Croatia	nosmanagic@efzg.hr
122.	Eren	Özceylan	Department of Industrial Engineering, Gaziantep University		27300	Gaziantep	Turkey	erenozceylan@ gmail.com
123.	Irena	Palić	University of Zagreb, Faculty of Economics and Business, Department of Statistics	Trg J.F. Kennedy 6	10000	Zagreb	Croatia	ipalic@efzg.hr
124.	Polona	Pavlovčič Prešeren	University of Ljubljana, Faculty of Civil and Geodetic Engineering	Jamova cesta 2	1000	Ljubljana	Slovenia	polona.pavlovcic@ fgg.uni-lj.si
125.	Karmen	Pažek	University of Maribor, Faculty of agriculture and life science	Pivola 11	2311	Hoče	Slovenia	karmen.pazek@um.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
126.	Mirjana	Pejić-Bach	University of Zagreb, Faculty Economics and Business, Department of Informatics	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	mpejic@efzg.hr
127.	Engin	Pekel	Yildiz Technical University, Faculty of Machine, Department of Industrial Engineering	A-622	34300	İstanbul	Turkey	pekelc@hotmail.com
128.	Matjaž	Perc	University of Maribor, Faculty of Natural Sciences and Mathematics	Koroška cesta 160	2000	Maribor	Slovenia	matjaz.perc@uni-mb.si
129.	Tunjo	Perić	University of Zagreb, Faculty Economics and Business, Department of Informatics	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	tperic@efzg.hr
130.	Fabio Henrique	Pereira	Universidade Nove de Julho/Informatics and knowledge management Graduate Program	Francisco Matarazzo, Av. 612		São Paulo	Brazil	fabiohp@uni9.pro.br
131.	Alessandro	Persona	Department of Management and Engineering, University of Padova	Stradella San Nicola 3	36100	Vicenza	Italy	alessandro.persona@ unipd.it
132.	Snježana	Pivac	University of Split, Faculty of Economics	Cvita Fiskovića 5	21000	Split	Croatia	snjezana.pivac@efst.hr
133.	Janez	Povh	University of Ljubljana, Faculty of mechanical engineering; Institute of mathematics, physics and mechanics	Aškerčeva cesta 6; Jadranska 19	1000	Ljubljana	Slovenia	janez.povh@ fs.uni-lj.si
134.	Vít	Ptošek	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 33	Ostrava	Czech Republic	vid.ptosek@vsb.cz

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
135.	Miroslav	Rada	University of Economics, Prague, Department of Financal Accounting and Auditing	Nám. W. Churchilla 4	13067	Prague	Czech Republic	miroslav.rada@vse.cz
136.	Uroš	Rajkovič	University of Maribor, Faculty of Organizational Sciences	Kidričeva cesta 55 a	4000	Kranj	Slovenia	uros.rajkovic@ fov.uni-mb.si
137.	Vladislav	Rajkovič	University of Maribor, Faculty of Organizational Sciences	Kidričeva cesta 55 a	4000	Kranj	Slovenia	vladislav.rajkovic@ fov.uni-mb.si
138.	Bernt Kvam	Randeberg	Oilfield Technology Group	Vassbotnen 1	4313	Sandnes	Norway	bernt.kvam.randeberg@ otg.no
139.	Nada	Ratković	University of Split, Faculty of Economics Split, Department of Quantitative Methods	Cvite Fiskovića 5	21000	Split	Croatia	nada.ratkovic@efst.hr
140.	Atle	Refsdal	SINTEF Digital	P.O. Box 124 Blindern	0314	Oslo	Norway	atle.refsdal@sintef.no
141.	Emina	Resić	University of Sarajevo, School of Economics and Business Sarajevo	Trg oslobođenja - Alija Izetbegović 1	71000	Sarajevo	Bosnia and Herzegovina	emina.resic@ efsa.unsa.ba
142.	Petra Renáta	Rigó	Budapest University of Technology and Economics, Babes-Bolyai University			Budapest	Hungary	takacsp@math.bme.hu
143.	Marko	Robnik-Šikonja	University of Ljubljana, Faculty of Computer and Information Science	Večna pot 113	1000	Ljubljana	Slovenia	marko.robnik@ fri.uni-lj.si
144.	Valerija	Rogelj	Fakulteta za organizacijske študije,	Ulica talcev 3	8000	Novo mesto	Slovenia	valerijarogelj@ gmail.com
145.	Bojan	Rosi	Faculty of Logistics, University of Maribor	Mariborska cesta 7	3000	Celje	Slovenia	bojan.rosij@um.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
146.	Ole Petter	Rosland	Statoil ASA	Forusbeen 50	4035	Stavanger	Norway	olpr@statoil.com
147.	Črtomir	Rozman	University of Maribor, Faculty of agriculture and life science	Pivola 11	2311	Hoče	Slovenia	crt.rozman@um.si
148.	Robert	Rupnik	SIJ Acroni d.o.o.	Cesta Borisa Kidriča 44	4270	Jesenice	Slovenia	rupnik642@gmail.com
149.	Darja	Rupnik Poklukar	University of Ljubljana, Faculty of Mechanical Engineering	Aškerčeva ulica 6	1000	Ljubljana	Slovenia	darja.rupnik@fs.uni.lj.si
150.	Sukran	Seker	Yildiz Technical University, Department of Industrial Engineering	Barbaros Street	34349	Istanbul	Turkey	seker.sukran@ gmail.com
151.	Mircea	Simionica	UniCredit Business Integrated Solutions, Financial Risks Factory			Milan	Italy	mircea.simionica@ gmail.com
152.	Fabio	Sgarbossa	Department of Management and Engineering, University of Padova	Stradella San Nicola 3	36100	Vicenza	Italy	fabio.sgarbossa@ unipd.it
153.	K.	Skouri	University of Ioannina, Department of Mathematics		45110	Ioannina	Greece	kskouri@uoi.gr
154.	Katerina	Slaninová	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 00	Ostrava	Czech Republic	katerina.slaninova@ vsb.cz
155.	Radim	Sojka	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 00	Ostrava	Czech Republic	radim.sojka@ vsb.cz

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
156.	Selin	Soner Kara	Yildiz Technical University, Faculty of Machine, Department of Industrial Engineering	A-631	34300	İstanbul	Turkey	ssoner@yildiz.edu.tr
157.	Petar	Sorić	University of Zagreb, Faculty of Economics and Business	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	psoric@efzg.hr
158.	Alexander S.	Strekalovskiy	Matrosov Institute for System Dynamics and Control Theory of SB RAS	Lermontov St., 134	664033	Irkutsk	Russia	strekal@icc.ru
159.	Morapitiye	Sunil				Budapest	Hungary	musz.sunil@gmail.com
160.	Alžbeta	Szendreyová	University of Žilina, Department of Mathematical Methods and Operations Research	Univerzitna 8215/1	01026	Žilina	Slovak Republic	alzbeta.szendreyova@fria.uniza.sk
161.	Simona	Šarotar Žižek	University of Maribor, Faculty of Economics and Business	Razlagova 14	2000	Maribor	Slovenia	simona.sarotar-zizek@um.si
162.	Jiří	Ševčík	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 33	Ostrava	Czech Republic	jiri.sevcik@vsb.cz
163.	Vanja	Šimićević	Croatian Studies	Kampus Borongaj, Borongajska cesta 83d	10000	Zagreb	Croatia	vanja.simicevic@hrstud.hr
164.	Nika	Šimurina	University of Zagreb, Faculty Economics and Business, Department of Finance	Trg J.F. Kennedya 6	10000	Zagreb	Croatia	nsimurina@efzg.hr
165.	Blanka	Škrabić Perić	University of Split, Faculty of Economics	Cvite Fiskovića 5	21000	Split	Croatia	bskrabic@efst.hr

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
166.	Tihana	Škrinjarić	University of Zagreb, Department of Mathematics and Department of Managerial Economics	Trg J. F. Kennedyja 6	10000	Zagreb	Croatia	tskrinjaric@efzg.hr
167.	Maja	Škurić	University of Montenegro, Maritime Faculty Kotor	Dobrota 36	85330	Kotor	Montenegro	majaskuric@gmail.com
168.	Ivana	Tadić	University of Split, Faculty of Economics	Cvita Fiskovića 5	21000	Split	Croatia	itadic@efst.hr
169.	Stanislav	Tojnko	University of Maribor, Faculty of agriculture and life science	Pivola 11	2311	Hoče	Slovenia	stanislav.tojnko@um.si
170.	Andrea	Trgo	KentBank d.d., Corporate banking department	Poljička cesta 26	21000	Split	Croatia	andreatrgo@gmail.com
171.	Tatjana	Unuk	University of Maribor, Faculty of agriculture and life science	Pivola 11	2311	Hoče	Slovenia	tatjana.unuk@um.si
172.	Wim	Van Grembergen	University of Antwerp/ Information Systems Management		2000	Antwerpen	Belgium	wim.vangrembergen@ uantwerpen.be
173.	Luk N.	Van Wassenhove	INSEAD	Boulevard de Constance	77305	Fontainebleau Cedex	France	luk.van-wassenhove@ insead.edu
174.	Jan	Vargovský	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 00	Ostrava	Czech Republic	jan.vargovsky@vsb.cz
175.	Živa	Veingerl Čič	University of Maribor, Faculty of Economics and Business	Razlagova 14	2000	Maribor	Slovenia	zivana.veingerl1@um.si

ID	First name	Surname	Institution	Street and Number	Post code	Town	Country	E-mail
176.	Aleksander	Vesel	University of Maribor; Faculty of Natural Sciences and Mathematics	Koroška cesta 160	2000	Maribor	Slovenia	vesel@uni-mb.si
177.	Jelena	Vidović	University of Split, University Department of Professional Studies	Kopilica 5	21000	Split	Croatia	jvidovic@oss.unist.hr
178.	Tea	Vizinger	Faculty of Logistics, University of Maribor	Mariborska cesta 7	3000	Celje	Slovenia	tea.vizinger@um.si
179.	David	Vojtek	IT4Innovations National Supercomputing Centre, VŠB-Technical University of Ostrava	17. listopadu 15/2172	708 33	Ostrava	Czech Republic	david.vojtek@vsb.cz
180.	Lidija	Zadnik Stirn	University of Ljubljana, Biotechnical Faculty	Jamnikarjeva 101	1000	Ljubljana	Slovenia	lidija.zadnik@bf.uni-lj.si
181.	Marijana	Zekić-Sušac	University of Josip Juraj Strossmayer in Osijek, Faculty of Economics	Trg Lj. Gaja 7	31000	Osijek	Croatia	marijana@efos.hr
182.	Jovana	Zoroja	University of Zagreb, Faculty Economics and Business, Department of Informatics	Trg J.F. Kennedyja 6	10000	Zagreb	Croatia	jzoroja@efzg.hr
183.	Janez	Žerovnik	University of Ljubljana, Faculty of Mechanical Engineering	Aškerčeva 4	1000	Ljubljana	Slovenia	janez.zerovnik@ fs.uni-lj.si
184.	Berislav	Žmuk	University of Zagreb, Faculty of Economics and Business	Trg J.F.Kennedyja 6	10000	Zagreb	Croatia	bzmuk@efzg.hr