Proceedings of the NSAIS-ROW'19 Workshop on Adaptive and Intelligent Systems and Real Options 2019

22.-23.8.2019 Lappeenranta, Finland

In association with Manufacturing 4.0, Finnish Strategic Research Council Project No. 313396

NSAIS
North-European Society for Adaptive and Intelligent Systems

ROW
Real Option Workshop

MFG 4.0

Edited by Mikael Collan

Proceedings of the NSAIS-ROW'19 Workshop on Adaptive and Intelligent Systems and Real Options 2019

Organized by the North-European Society of Adaptive and Intelligent Systems and Finnish Real Options Society together with Manufacturing 4.0 - Finnish Strategic Research Council project No. 313396

22.-23.8.2019 Lappeenranta, Finland

Editor: Mikael Collan

LUT Scientific and Expertise Publications
Research Reports, 97
ISSN 2243-3376

ISBN 978-952-335-400-5 (PDF)

9 789523 354005

FOREWORD

This workshop NSAIS-ROW´19 is the third one in the series of workshops organized by the North-European Society for Adaptive and Intelligent Systems (NSAIS). The first workshop took place more than ten years ago, it was followed by a workshop in 2016, and this is the third international event organized by NSAIS. The idea of this workshop is to continue the series of workshops around the areas of interest to the NSAIS in Northern Europe.

This time the workshop is organized in association with the Finnish Real Options Society (FROS) and includes real options as one key topical area. The workshop is supported by the Manufacturing 4.0 (MFG40) Finnish Strategic Research Council project No. 313396. The conference is organized at and supported by Lappeenranta-Lahti University of Technology (LUT) Lappeenranta campus – the organizers thank the aforementioned organizations for their support.

This proceedings includes altogether nineteen abstracts of the submissions presented at the NSAIS-ROW´19 workshop. They include authors representing eleven nationalities. The papers are a mix along the topical theme of the conference with many contributions to the fields of economic and financial decision-making and analytics in general. All abstracts have undergone an editorial peer review. The organizers thank the international scientific program committee and the reviewers for their support.

Pasi Luukka, Chairman, NSAIS-ROW´19 program committee, President of NSAIS

Mikael Collan, Co-Chair, NSAIS-ROW´19 program committee, President of FROS

Mariia Kozlova, Co-Chair, NSAIS-ROW´19 local organizing committee

NSAIS-ROW'19 International Scientific Program Committee

Pasi Luukka, LUT, Finland (Chair)
Mikael Collan, LUT, Finland, (Co-chair)
Mariia Kozlova, LUT, Finland
Jyrki Savolainen, LUT, Finland
Sheraz Ahmed, LUT, Finland
Eero Pätäri, LUT, Finland
Azzurra Morreale, LUT, Finland
Jan Stoklasa, LUT, Finland and Palacky U., Czech Republic
Luigi Mittone, U. Trento, Italy
Matteo Brunelli, U. Tre
Jana Krejčí, U. Trento, Italy
Mario Fedrizzi, U. Trento, Italy
Yuri Lawryshyn, U. Toronto, Canada
Christer Carlsson, Åbo Akademi, Finland
Jozsef Mezei, Åbo Akademi, Finland
Xiaolu Want, Åbo Akademi, Finland
Anna Sell, Åbo Akademi, Finland
Markku Heikkilä, Åbo Akademi, Finland
Julian Yeomans, U. York, Canada
Matt Davison, U. Western Ontario, Canada
Irina Georgescu, CSIE, Romania
Roberto Montemanni, IDSIA, Switzerland
Onesfole Kurama, LUT, Finland
Andri Riid, Tallinn UT., Estonia
Iulian Nastac, Polytechnic U. Bucharest, Romania

Local organizing committee

Pasi Luukka, LUT, Finland (Chair)
Mikael Collan, LUT, Finland (Co-Chair)
Mariia Kozlova, LUT, Finland (Co-Chair)

TABLE OF CONTENTS

Abstracts

NSAIS-ROW19 Workshop Schedule

Thursday – August 22nd

Venue: Auditorium - Student Association building

09.00 Opening ceremony – welcome words

09.15 Session I –Behavioral and cognitive considerations (Chair: Jan Stoklasa)

1. Mariia Kozlova and Jan Stoklasa - Satisfaction and Extremity Effects under the Baseline Approach in Multiple-Criteria Decision-Making Problems

2. Jana Stoklasová, Jan Stoklasa, Tomáš Talášek and Azzurra Morreale- Risk attitude in the context of preference extraction through questionnaires

3. Anssi Tarkiainen, Pasi Luukka, Ari Jantunen, Jukka-Pekka Bergman, Olli Kuivalainen and Jan Stoklasa - Managerial cognition and firm performance

4. Olga Bogdanova and Mariia Kozlova - First steps of Data Collection Planning for Development of Artificial Intelligence Technology for the European Crowdfunding Aggregator Platform

11.15 Lunch break (Lunch at the Student Union building)

12.15 Paper Session II – Methodological advances (Chair: Pasi Luukka)

1. Christoph Lohrmann and Pasi Luukka - Fuzzy similarity and entropy (FSAE) feature selection revisited by using intra-class entropy and a normalized scaling factor

2. Jan Stoklasa and Jana Siebert - On the use of pairwise comparisons for absolute-type evaluation

3. Christoph Lohrmann and Mariia Kozlova - Developing a Decision Tree Block for the Exercise Boundary Fitting Method

13.45 Break

14.00 Paper Session III - Business and industrial applications (Chair: Mikael Collan)

1. Stein-Erik Fleten, Benjamin Fram, Magne Ledsaak, Sigurd Mehl, Ola Røssum and Carl Ullrich - Analysing Peaking Generator Availability in PJM: The Effect of the Reliability Pricing Model

2. Aleksandr Popkov and Mikhail Zaboev - Extended MobileNet convolutional neural network with Bayesian approach for economic applications

3. Jyrki Savolainen and Michele Urbani - Optimizing maintenance scheduling of multi-machine industrial systems with simulation – example from metals mining

4. Yuri Lawryshyn - Using Boundary Fitting and Machine Learning to Value Multi-Stage Real Option Investments

16.00 End of the first workshop day

In the evening: Evening program – Dinner

Friday – August 23rd

Venue: Auditorium - Student Association building

09.30 Paper Session IV – Analytics with applications (Chair: Pasi Luukka)

1. Christoph Lohrmann and Alena Lohrmann - Missing value imputation of the cooling technology for the global thermal power plant fleet using a decision tree classifier

2. Zhukov Dmitry and Tatiana Khvatova - A stochastic dynamics model for shaping stock indexes considering self-organization processes, memory and oscillations

3. Lyudmila Vyunenko - Features of the Calibration Procedure for AK-type Macroeconomic Models

4. Mahinda Mailagaha Kumbure, Anssi Tarkiainen, Pasi Luukka, Jan Stoklasa and Ari Jantunen - Examining Inter-Causal Relationships in Shared Cognitive Maps

11.30 Lunch break

12.30 Paper Session V– Finance and Business (Chair: Mikael Collan)

1. Lyudmila Gadasina and Kirill Tashlykov - Analysis of a Digital Transformation Factors' Impact On Company Economic Performance

2. Irina Georgescu and Jani Kinnunen – Optimal Saving by Expected Utility Operators

3. Sheraz Ahmed and Muhammad Naeem - Adaptive market hypothesis: A comparative analysis of seven major digital currencies

4. Mika Jaakkola, Andreas Mikkelsen, Sheraz Ahmed, and Mikael Collan – Using chine learning in optimizing pairs trading strategies in the Nordic context

14.40 Closing ceremony – closing words

Abstracts

Satisfaction and Extremity Effects under the Baseline Approach in Multiple-Criteria Decision-Making Problems

Mariia Kozlova[1] and Jan Stoklasa[1,2,*]
[1]LUT University, School of Business and Management,
Lappeenranta, Finland
[2]Palacký University Olomouc, Faculty of Arts,
Department of Applied Economics, Olomouc, Czech Republic
*corresponding author, e-mail: jan.stoklasa@lut.fi

***Abstract*—This extended abstract presents further exploration of the behavioral biases of a decision-maker under the baseline approach in multiple-criteria decision making (MCDM). More specifically it presents a simple formal model of these biases based on a specific mechanism of criteria weights generation. While classical MCDM defines the decision problem as a choice between multiple alternatives, the baseline approach recently proposed by us highlights a special role of one of the alternatives, the baseline. The baseline alternative is the one that the decision-maker is currently in possession of or a set of the characteristics of the absence of a solution in case no current solution is available. We have recently proposed two algorithms of weights creation – "craving the unavailable" and "appreciating the possessed" and their combination in MCDM with a baseline. In this work we explore how different level of satisfaction with the baseline and the extremity of its evaluation with respect to the criteria affect the decision-maker's perception of other alternatives.**

Keywords—multi-criteria decision making, baseline approach, behaviorism, operations research, customer choice.

I. Introduction

Various decision making biases in multiple choice problems have been detected experimentally [1, 2]. A recent study aims at explaining the underlying mathematical mechanism that governs irrational evaluation of alternatives [3]. The baseline approach challenges the assumption of standard multiple-criteria decision-making (MCDM) problem formulation of the equal nature of all alternatives [4-6]. Instead, a special role is assigned to the baseline alternative, the one that occupies the current solution slot or, in other words, the one that is currently possessed (or the "empty solution slot", i.e. the characteristics of the absence of a solution are used to represent the baseline). A simple combination of rules for weight formation that depends on the evaluation of the baseline leads to arising of multiple behavioral effects observable in reality. Such as the willingness to stay with the baseline when an alternative is the same or even slightly better as long as the satisfaction with the baseline is high; constant switching between opposite alternatives (in terms of their evaluation with respect to the criteria) when satisfaction with both is relatively low [3]. Satisfaction here represents the overall evaluation of the alternative or its equivalent. In this paper, we use (5) to calculate the value of satisfaction with an alternative X, i.e. $S_X \in [-1,1]$.

In this paper we further explore the baseline approach by systematically examining the results of the valuations under different satisfaction levels with the baseline alternative and different extremity of the evaluation of the alternatives with respect to the criteria. We illustrate the conclusions with an artificial example.

II. The Baseline Approach

The baseline approach postulates that the baseline is perceived the better the higher the criteria values it has (the appreciating possessed rule), while the alternatives are perceived the better the higher they fulfill the criteria that are saturated lower by the baseline (the craving unavailable effect). Let us assume $n+1$ alternatives $A_1, \dots, A_n$ and a *baseline alternative* B, and m criteria $C_1, \dots, C_m$. Without any loss of generality, let us assume a benefit type scale represented by the interval $[e_{min}, e_{max}] \subset \mathbb{R}$ for all the criteria.

Assuming linear relationships the *appreciating the possessed weights* are defined as

$$w_j^{AP} = \frac{e_B^{C_j} - e_{min}}{e_{max} - e_{min}}, \tag{1}$$

where j is criterion index, $j = 1, \dots, m$, and

e_B^{Cj} is the estimate of the j criterion in the baseline,

e_{min} and e_{max} are the corresponding minimum and maximum of the evaluation scale. These weights reflect the decision-maker's need to preserve saturation of criteria that he/she is already used to under the given baseline alternative.

The *craving the unavailable weights* are defined in the opposite way, representing the craving for improvement of saturation in those criteria, that are currently not sufficiently saturated by the baseline alternative. These weights are computed in the following way:

$$w_j^{CU} = \frac{e_{max} - e_B^{C_j}}{e_{max} - e_{min}}. \tag{2}$$

Acknowledgements: The authors would like to acknowledge the financial support received from the Finnish Strategic Research Council, grant number 313396 / MFG40 - Manufacturing 4.0 and by the Specific University Research Grant IGA FF 2019 002, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2019.

The whole set of the alternatives is evaluated based on the combination of these weights

$$w_j^{Comb} = \begin{cases} w_j^{AP} & \text{for the baseline } B, \\ w_j^{CU} & \text{for the other alternatives } A_1, \dots A_n. \end{cases} \quad (3)$$

using the weighted sum

$$e_x = \sum_{j=1}^{m} (w_j^{Comb} \cdot e_x^{c_j}), \quad (4)$$

where e_x is the overall evaluation of an alternative A_x, $x \in \{1, \dots, n\}$.

Finally, the satisfaction with the baseline $S_B \in [-1,1]$ is defined by how much its evaluations with respect to the criteria are different from the mean value of the evaluation scale $[e_{min}, e_{max}]$:

$$S_B = \frac{\sum_{j=1}^{m} \left[e_B^{c_j} - \left(\frac{e_{max} - e_{min}}{2} \right) \right]}{m \left(\frac{e_{max} - e_{min}}{2} \right)} \quad (5)$$

$S > 0$ implies that the baseline estimates on average are higher than the mean of the evaluation scale.

III. Numerical experiments and Conclusion

The results of the numerical experiments are presented in the Appendix 1. The chosen evaluation scale is [0,10] with benefit-type interpretation. In the rows the extremity of estimates is changing starting from the case where the baseline is evaluated (close to) the best under one of the criteria and (close to) worst under the other criterion (this is denoted as *the minmax case*) to the *close to average case* (all evaluations of the baseline with respect to all the criteria are close to the middle scale value). The columns represent different satisfaction levels. The set of alternatives in every case includes:

- the baseline alternative,
- an alternative evaluated the same with respect to all the criteria (denoted as "*same*"),
- an alternative for which the evaluations with respect to the two criteria are switched (denoted as "*opposite*"),
- an alternative average in all the criteria (denoted "50/50"),
- an alternative with the evaluation with respect to one criterion maximized and the evaluation with respect to the other one having the average value (denoted "100/50"), and
- an alternative with both evaluations with respect to the criteria maximized (denoted "100/100").

The main conclusion from the experiments is that the positive satisfaction leads to high cost of switching where the opposite alternative is not preferred over the baseline, and in the minmax case even the "star" alternative "100/100" is not preferred over the baseline. Contrary, negative satisfaction with the baseline leads to the opposite and the "star" alternative is always preferable, and even the same alternative becomes attractive in the "close to average" case.

As elaborated in the original introduction of the framework in [3], these effects can be interpreted by different perception of the uncertainty associated with switching to another alternative. When the satisfaction with the current solution is high, such uncertainty could be perceived as possible loss, while with low satisfaction the perception shifts towards possible potential or gain.

References

[1] T. Kogut, "Choosing what I want or keeping what I should: The effect of decision strategy on choice consistency," Organizational Behavior and Human Decision Processes, vol. 116, pp. 129-139, September 2011. 2011.

[2] J.S. Trueblood and J.C. Pettibone, "The phantom decoy effect in perceptual decision making," J.Behav.Decis.Making, vol. 30, pp. 157-167, 2017.

[3] J. Stoklasa and M. Kozlova, "Baseline approach – a new framework for multiple-criteria decision making," European Journal of Operational Research, pp. 26, Under review.

[4] P. Yu, Multiple Criteria Decision Making: Concepts, Techniques and Extensions, New York: Plenum Press, 1985.

[5] G. Mitra, H.J. Greenberg, F.A. Lootsma, M.J. Rickaert and H.-. Zimmermann, Mathematical Models for Decision Support, Berlin Heidelberg New York London Paris Tokyo: Springer-Verlag, 1988.

[6] T.L. Saaty, Fundamentals of decision making and priority theory with the analytic hierarchy process, Pittsburgh: RWS publications, 2000.

Appendix 1.
Perceived value of alternatives under different cases of satisfaction and extremity of estimates of the baseline.

ZERO SATISFACTION

MINMAX

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	0 %	**10**	**0**	**10,0**	
Same		10	0	0,0	-10,0
Opposite		0	10	10,0	0,0
50/50		5	5	5,0	-5,0
100/50		10	5	5,0	-5,0
100/100		10	10	10,0	0,0
Perceived weight (baseline)		1,00	0,00		
Perceived weight (alternative)		0,00	1,00		

MODERATE

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	0 %	**8**	**2**	**6,8**	
Same		8	2	3,2	-3,6
Opposite		2	8	6,8	0,0
50/50		5	5	5,0	-1,8
100/50		10	5	6,0	-0,8
100/100		10	10	10,0	3,2
Perceived weight (baseline)		0,80	0,20		
Perceived weight (alternative)		0,20	0,80		

CLOSE TO AVERAGE

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	0 %	**6**	**4**	**5,2**	
Same		6	4	4,8	-0,4
Opposite		4	6	5,2	0,0
50/50		5	5	5,0	-0,2
100/50		10	5	7,0	1,8
100/100		10	10	10,0	4,8
Perceived weight (baseline)		0,60	0,40		
Perceived weight (alternative)		0,40	0,60		

POSITIVE SATISFACTION

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	10 %	**10**	**1**	**10,1**	
Same		10	1	0,9	-9,2
Opposite		1	10	9,0	-1,1
50/50		5	5	4,5	-5,6
100/50		10	5	4,5	-5,6
100/100		10	10	9,0	-1,1
Perceived weight (baseline)		1,00	0,10		
Perceived weight (alternative)		0,00	0,90		

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	10 %	**8**	**3**	**7,3**	
Same		8	3	3,7	-3,6
Opposite		3	8	6,2	-1,1
50/50		5	5	4,5	-2,8
100/50		10	5	5,5	-1,8
100/100		10	10	9,0	1,7
Perceived weight (baseline)		0,80	0,30		
Perceived weight (alternative)		0,20	0,70		

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	10 %	**6**	**5**	**6,1**	
Same		6	5	4,9	-1,2
Opposite		5	6	5,0	-1,1
50/50		5	5	4,5	-1,6
100/50		10	5	6,5	0,4
100/100		10	10	9,0	2,9
Perceived weight (baseline)		0,60	0,50		
Perceived weight (alternative)		0,40	0,50		

NEGATIVE SATISFACTION

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	-10 %	**9**	**0**	**8,1**	
Same		9	0	0,9	-7,2
Opposite		0	9	9,0	0,9
50/50		5	5	5,5	-2,6
100/50		10	5	6,0	-2,1
100/100		10	10	11,0	2,9
Perceived weight (baseline)		0,90	0,00		
Perceived weight (alternative)		0,10	1,00		

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	-10 %	**7**	**2**	**5,3**	
Same		7	2	3,7	-1,6
Opposite		2	7	6,2	0,9
50/50		5	5	5,5	0,2
100/50		10	5	7,0	1,7
100/100		10	10	11,0	5,7
Perceived weight (baseline)		0,70	0,20		
Perceived weight (alternative)		0,30	0,80		

	Satisfaction level	Criterion 1	Criterion 2	Perceived value	Delta
BASELINE	-10 %	**5**	**4**	**4,1**	
Same		5	4	4,9	0,8
Opposite		4	5	5,0	0,9
50/50		5	5	5,5	1,4
100/50		10	5	8,0	3,9
100/100		10	10	11,0	6,9
Perceived weight (baseline)		0,50	0,40		
Perceived weight (alternative)		0,50	0,60		

Risk attitude in the context of preference extraction through questionnaires

Jana Stoklasová*, Jan Stoklasa†, Tomáš Talášek‡, Azzurra Morreale§
*Marital and family counseling centre Prostějov, Czech Republic, jana.stoklasova.cz@gmail.com
†School of Business and Management, LUT University, Lappeenranta, Finland, and
Palacký University Olomouc, Czech Republic, Department of Applied Economics, jan.stoklasa@lut.fi
‡Palacký University Olomouc, Czech Republic, Department of Applied Economics, tomas.talasek@upol.cz
§School of Business and Management, LUT University, Lappeenranta, Finland, azzurra.morreale@lut.fi

***Abstract*—Questionnaires are a frequent source of data in social science and economics. In this paper we consider their use for the extraction of information on preferences of decision-makers and also for the identification of their attitude to risk. A choice between two alternatives is considered - a risky one and a non-risky one with a sure payoff identical to the expected payoff of the risky alternative. We investigate how to obtain a correct information concerning the preferences of the decision-maker in this context. We also inquire how to assess the risk-attitude of the decision-maker based on his/her response. We compare the added/lost value of the use of questionnaire in the extraction of preferences and in the identification of risk-attitude to the possible added/lost value of the use of observation or experiment in this endeavor. We discuss the limitations of our ability to distinguish between indifference and incomparability by the results of a choice-task of the possibility of inferring risk attitude correctly simply based on the choices made by the decision-maker.**

***Keywords*—preferences, risk, risk attitude, incomparability, indecisiveness, indifference, questionnaire, choice, behavior**

I. Introduction

Preferences and risk-attitude are crucial concepts in economics in the theory of consumer and in psychology. These concepts also play a crucial role in decision-making in general. Our ability to extract preferences from a decision-maker is vital for the construction of utility functions. Our ability to assess his/her attitude to risk is, in turn, required for the design of normative models of decision-making under risk.

This paper deals with the issue of preference extraction and risk-attitude extraction, or the limitations thereof, by the frequently used tools of social-science research, such as questionnaires and observation. Field experiments, quasi-experiments and experiments are not directly considered, although some of the issues discussed in the paper might be applicable to the experimental setting as well.

We argue that choice as such can, in certain settings, be a rather weak piece of information for preference extraction (and for the extraction of risk attitude as well). We therefore discuss what information can be inferred from actually made choices - mainly in the context of forced-choice situations or questionnaire items where a single of the proposed alternatives is to be chosen/selected. We also propose a set of modified forced-choice answers (called *quasi-forced-choice answers*) for questionnaires requiring the choice between two (or possibly more) alternatives. This format of answers to be provided to the respondents in a forced-choice setting was introduced in the context of our previous research [1], [2]. We discuss its benefits and limitations in the context of preference and risk-attitude extraction and propose possible further research directions in this area.

II. Forced choice and preferences

Let us now consider a decision-maker (respondent) is faced with a choice between alternative A and alternative B. The only thing we assume is that these alternatives are understood and imaginable for the decision-maker. In other words we assume that the alternatives are meaningful for the decision-maker and can be assessed by him/her. We do not make any assumptions concerning the completeness of comparisons (preferences), i.e. we do not rule out incomparability (as e.g. the assumptions of utility theory do [3]).

Our main point of inquiry for the paper is the following - what can we infer from a choice performed by the decision-maker in terms of his preferences in the set $\{A, B\}$. There are potentially four actual preferences:

- A is preferred over B, denoted $A \succeq B$, with the interpretation that A is not worse than B,
- B is preferred over A, denoted $B \succeq A$, with the interpretation that B is not worse than A,
- A is equivalent to B in terms of the overall criterion used for the comparison (not necessarily in terms of the individual criteria), denoted $A \approx B$, with the interpretation that the decision-maker is indifferent between A and B,
- A and B are incomparable, i.e. the decision-maker is unable to to confirm that A is not worse than B and neither is he/she able to confirm the opposite, i.e. that B is not worse than A. In terms of notation we would have $(A \not\succeq B) \wedge (B \not\succeq A)$.

Now, if we assume a rational decision-maker, the results of the decision-making process expressed in a questionnaire (or demonstrated by the behavioral results of the process) by a choice between A and B in the forced-choice setting, would be the following, in the respective order:

- the choice of A
- the choice of B

- the choice of either of the alternatives A or B; note, that since both are considered equally good in terms of the overall criterion, selection of either of these alternatives optimizes the value of the overall criterion. Even a *random choice* between A and B can, therefore, in this case be considered rational.

- again the choice of either of the alternatives A or B; since the alternatives are incomparable, there is no simple way to consider one to be superior. If a choice is required nevertheless, then either choice is as good as the other one, since there is no simple (or even rational) way to decide which one to choose, so a *random choice* might be the only solution. In this case it is questionable whether such a choice is rational, but it cannot be performed in a more rational way.

From the above it is apparent, that forced-choice questions or observation of behavior resulting in the choice of one of the alternatives (A or B) do not help us uncover the actual preferences at all. In essence the selection of A rules out $B \succeq A$, the selection of B rules out $A \succeq B$ in the forced choice setting. But this is as far as we can get (see also e.g. [4]). The possibility of identification of actual preference, indifference and incomparability are nonexistent, as long as a repeated choice is not performed. In case we have repeated choice, then preferences and indecisiveness (covering incomparability and indifference) can be distinguished [5]. When we do not have the time to observe repeated choice, or when such a choice is not repeatable in its essence, the forced-choice questions are unable to identify preferences.

To deal with this issue, we have proposed a *quasi-forced-choice set of answers* for the investigation of the preference between alternatives A and B in [6] as summarized in Fig. 1. As we can see, the proposed set of answers allows for the identification of clear preference by the first two options, but at the same time allows for the identification of indecisiveness. It is still not possible to distinguish between incomparability and indifference, but at least clear preference can be extracted at no significant cost. Note also, that the final result is still a choice of one of the alternatives.

	Answer
☐	I prefer **A**
☐	I prefer **B**
☐	**I cannot decide**, but if I had to choose one, it would be A
☐	**I cannot decide**, but if I had to choose one, it would be B

Fig. 1. A set of answers for the quasi-forced choice as suggested in [6]

III. Forced choice, quasi-forced choice and risk attitude

Even thought the discussion of the possibilities of extracting preferences based on a questionnaire with (quasi-)forced choice or based on observation has so far been done in the preference extraction framework, the issues are valid also for the identification of attitude to risk. Let us now investigate the link between forced choice and risk attitude. Our question is the following: *Can a forced choice between these two alternatives reveal the attitude to risk of a decision-maker?*

For this purpose let us assume that the alternative A represents a sure gain for the decision-maker and B represents a lottery (a risky alternative) with an expected gain identical to the gain in A. For simplicity let us assume gain only, keeping in mind that losses and the combination of gains and losses might require separate investigation as evidenced e.g. by [7], [8]. Let us adopt the following definitions of risk attitude (adapted in line with [9]):

- The decision-maker is *risk seeking*, if he/she prefers a risky alternative with an expected gain of x over an alternative resulting in a sure gain of x.
- The decision-maker is *risk averse*, if he/she prefers an alternative resulting in a sure gain of x over a risky alternative an expected gain of x.
- The decision-maker is *risk neutral*, if he/she is indifferent between a risky alternative with an expected gain of x and an alternative resulting in a sure gain of x.

It would thus seem that choice can be a manifestation of risk attitude as well as long as we consider the setting suggested in this chapter. Note, that the forced-choice approach suffers from the same shortcomings now as it does in the preference extraction setting. However, since only three possible attitudes to risk are considered, the findings obtainable by this approach are slightly stronger:

- If A is chosen, then risk seeking can be ruled out. The decision-maker can thus be either risk averse or risk neutral (in case of risk neutrality, A could have been chosen randomly as a result of forced choice and indifference between A and its risky counterpart B).
- If B is chosen, then risk aversion can be ruled out. The decision-maker can thus be either risk seeking or risk neutral (in case of risk neutrality, B could have been chosen randomly as a result of forced choice and indifference between B and its sure counterpart A).

If we apply the quasi-forced choice approach, then the choice of A is a clear indication of risk aversion, B is a clear indication of risk-seeking. Having ruled-out two out of the three considered risk attitudes, then both “I cannot decide” answers would point to risk-neutrality by the process of elimination. Note that if our assumption of only three risk attitudes holds, then the quasi-forced choice is actually capable of identifying the risk attitude in this setting.

IV. Conclusion

In this paper we investigate the possibility of extracting preferences and risk attitudes by questionnaires with forced-choice questions and by observation. We conclude, in line with previous research, that in this setting the choice is not sufficient to identify preference and/or risk attitude. We therefore suggest quasi-forced choice questions, that allow for the identification of indecisiveness. The ability of identifying indecisiveness (i.e. incomparability of indifference) allows for the identification of preference. In the risk-attitude setting, it allows for the identification of risk-seeking and risk-aversion, and as long as only three modes of risk-attitude are assumed, of risk-neutrality as well by ruling out the two previous cases.

Acknowledgment

The authors would like to acknowledge the financial support received from the Finnish Strategic Research Council, grant number 313396 / MFG40 - Manufacturing 4.0 and by the Specific University Research Grant IGA_FF_2019_002, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2019.

References

[1] A. Morreale, J. Stoklasa, M. Collan, G. Lo Nigro, and T. Talášek, “The effect of multiple possible outcomes representation on managerial decision-making under uncertainty: an exploratory survey,” in *Nineteenth International Working Seminar on Production Economics 2016, Preprints volume 3*, R. W. Grubbström and H. H. Hinterhuber, Eds., Innsbruck, 2016, pp. 239–248.

[2] A. Morreale, J. Stoklasa, M. Collan, and G. Lo Nigro, “Uncertain outcome presentations bias decisions: experimental evidence from Finland and Italy,” *Annals of Operations Research*, 2016. [Online]. Available: http://link.springer.com/10.1007/s10479-016-2349-3

[3] O. Morgenstern and J. Von Neumann, *Theory Of Games And Economic Behavior*, third edit ed. Princeton: Princeton University Press, 1953.

[4] A. Sen, “Behaviour and the Concept of Preference,” *Economica*, vol. 40, no. 159, pp. 241–259, 1973.

[5] K. Eliaz and E. A. Ok, “Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences,” *Games and Economic Behavior*, vol. 56, no. 1, pp. 61–86, 2006.

[6] A. Morreale, J. Stoklasa, and T. Talášek, “Fuzzy Grouping Variables in Economic Analysis. A Pilot Study of a Verification of a Normative Model for R&D Alliances,” *Fuzzy Economic Review*, vol. 21, no. 2, pp. 19–46, 2016.

[7] D. Kahneman and A. Tversky, “Choices , Values , and Frames,” *American Psychologist*, vol. 39, no. 4, pp. 341–350, 1984.

[8] A. Tversky and D. Kahneman, “The framing of decisions and the psychology of choice.” pp. 453–458, 1981.

[9] D. Kahneman and A. Tversky, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979.

Managerial cognition and firm performance

Tarkiainen, A., Luukka, P., Jantunen, A. Bergman, J-P., Kuivalainen, O., Stoklasa, J.
School of Business and Management,
Lappeenranta-Lahti University of Technology LUT,
Lappeenranta, Finland

***Abstract*— In this study, we empirically assess the relationship between shared managerial cognition and firm performance. We find that cognitive diversity within the firm is positively related to firm performance.**

Keywords—managerial cognition; strategy; performance; cognitive mapping

I. Introduction

Scholars in strategic management have increasingly directed their attention towards differences in managerial cognitive structures in order to understand organizational sense-making processes [1]. This phenomenon is generally referred to as managerial cognition in management studies literature. The diversity in managerial cognition is generally seen as a source of heterogeneity in firm strategies and performance [2]. In this study, we empirically assess the relationship between shared managerial cognition and firm performance. We develop hypotheses and test them with empirical data, which combines results of strategy simulation (CESIM Global Challenge) and measures of shared managerial cognition (within and between firms) based on elicited cognitive maps of students at international business strategy – course.

II. Literature review

When thinking about organizations, it should be noted that managers work in teams gathering, sharing, and attending to relevant information and jointly analyze and integrate it creating shared representations on a certain issue or situation [3]. Therefore, it is actually the differences in shared managerial cognition that is the source of heterogeneity in firm strategies and performance. The differences in shared managerial cognition may occur at least at two different levels of analysis: a) cognitive diversity among the members of management teams, and b) cognitive diversity among the firms that are competing against each other in the same market. Interestingly, the word shared has two opposite meanings [4]: a) it may refer to things that are in common, or b) it may refer to dividing up things. Strategic management of the firm is a complex task, and therefore can be considered as a task that is beyond the cognitive capabilities of one person [5]. In this meaning, shared mental model is built on cognitive diversity within decision-making groups, which is a necessity for emergence of wider horizon of opportunities for organizational renewal and surveillance in competitive environment. However, dark side of this cognitive diversity is that it may lead to difficulties in reaching consensus, which slows down decision-making and diminishes organizational responsiveness to environmental changes [6]. Thus, it seems that also the other meaning of sharedness of mental models, i.e. having something in common, is important. At the level of managerial cognition between organizations, this is related to differences in dominant logic of the market. If a firm's shared managerial cognition differs significantly from the dominant cognition in its markets, it may enable them to introduce radical innovations to the market, and create value more efficiently compared to its competitors. However, the drawback of having different dominant logic is that the firm's strategy may lack fit with the logic of the market it operates in, and consequently lead to poor performance. Our review of extant literature reveals that cognitive diversity has mainly been assessed as between firm diversity. When shared cognition within the firm has been studied, the applied approaches are mainly characterized by focus on similarity rather than diversity. The exceptions that have not focused on similarity have relied on research methods that do not enable assessment of actual knowledge structures (e.g. surveys).

Based on literature review above, we draw four hypotheses:

H1: Cognitive diversity between firms in a market is related to firm performance.

H2: Cognitive diversity within firms is related to firm performance.

H3: Cognitive diversity between firms in a market has a nonlinear U-shaped relationship with firm performance.

H4: Within firm cognitive diversity has a nonlinear inverted U-shaped relationship with firm performance.

III. Methods

A. *Sample and procedure*

The empirical data was collected during years 2014 and 2015 in a course of International Business Strategies in Lappeenranta University of Technology, School of Business and Management. In the course, the students participated on a simulation exercise (4 student groups represented 4 separate competitive markets), where they worked in teams of 4-5 people (see Table I). In teams (which represented firms in each market), students first developed their strategy for a multinational corporation operating in mobile business. Their individual cognitive interpretations of strategic issues were collected by causal mapping technique, which followed the procedure outlined by [3]. The individual maps were aggregated to shared maps at the levels of team, which represented shared maps of organizations. The shared maps of teams in each market (i.e. student group) were further aggregated to shared cognition (or dominant logic) of the market. After returning their strategic plans and cognitive causal maps, students competed against each other (in the four separate markets) for 8 weeks in CESIM's Global Challenge – simulation (https://www.cesim.com/simulations/cesim-global-challenge-international-business-strategy-simulation-game).

TABLE I. SAMPLE OF THE STUDY

Participants		
Student groups (markets)	*Number of teams (firms)*	*Number of students in teams*
2014 Group 1	9	49
2014 Group 2	8	39
2015 Group 1	7	34
2015 Group 2	7	34
Total sample	31	156

B. Measures

Firm performance was taken from the simulation and it was the cumulative total shareholder return after 8 rounds (i.e. fiscal years) of implementing strategy. Codes for cognitive diversity computations were implemented with Mathworks Matlab R2013b software. The measure for between firm cognitive diversity was calculated from aggregated shared map of a market and shared map of the firm using generalized distance formula proposed in [3], which is further developed from [7] distance ratio. It captures the distance between firm's shared cognition and the shared cognition of the market. Between firm cognitive diversity in a market was calculated with Mathworks Matlab R2013a software. The measure for within firm cognitive diversity was calculated average of team members' distances to team's shared map. Again, we used generalized distance formula [3], which is further developed from [7] distance ratio. This measure captures the average distance between team members' cognition and firms shared cognition, and thus represents the amount of cognitive diversity within each team (i.e. firm). The measures are free from common method biases, because dependent variable and the independent variables are collected from different sources.

IV. RESULTS

The hypotheses were tested with OLS-regression. Dummy variables were used to control for market-specific characteristics, such as competitive intensity. Orthogonalized squared terms of the cognitive diversity variables (both between and within) were used, in order to avoid problems with multicollinearity. The results of hypothesis testing are presented in Table II. The effect of Between firm diversity is non-significant, and consequently H1 is not supported. H2 is supported since the effect of Within firm diversity is significant at $p<0.05$. The firms with higher levels of internal cognitive diversity perform better in terms of shareholder return than firms with lower levels of internal cognitive diversity. We did not find support for hypotheses H3 and H4.

TABLE II. OLS RESULTS

Table Head	**Dependent variable: Total cumulative return to shareholders (%)**		
	β	*t-value*	*VIF*
Control variables			
Market dummy1	.487	2.66 [b]	1.68
Market dummy2	.711	4.28 [a]	1.38
Market dummy3	.507	2.91 [a]	1.52
Linear effects			
Between firm diversity (H1)	.170	1.05	1.32
Within firm diversity (H2)	.406	2.25 [b]	1.63
Non-linear effects			
Squared between fd (H3)	-.036	-.24	1.12
Squared within fd (H4)	-.120	-.80	1.14
R^2	.541 [a]		

a p<0.01; b p<0.05; c p<0.10

V. DISCUSSION

We find strong support for the view that cognitive diversity among management team predicts positively firm performance. We do not find support for our hypotheses on cognitive diversity between the firms in the market. Thus, our findings provide empirical support for the view that cognitive diversity within decision-making groups is a necessity for emergence of wider horizon of opportunities for organizational renewal [6]. We also assessed the potential nonlinear effect, which posits that high level of this internal cognitive diversity would slow down decision-making and diminish organizational responsiveness to environmental changes [6], but we do not find support for this phenomenon.

REFERENCES

[1] G.P Huber and K. Lewis, "Cross-understanding: Implications for group cognition and performance", Academy of Management Review, vol. 35, issue 1, 2010, pp. 6-26.

[2] M.S. Gary and R.E. Wood, "Mental models, decision rules, and performance heterogeneity", Strategic Management Journal, vol. 32 issue 6, 2011, pp. 569-594.

[3] J.P. Bergman, P. Luukka, A. Jantunen, A. Tarkiainen, Cognitive diversity, managerial characteristics and performance differences across the cleantech firms, accepted to International Journal of Knowledge-Based Organizations.

[4] S. Mohammed, L. Ferzandi, and K. Hamilton, "Metaphor no more: A 15-year review of the team mental model construct", Journal of Management, vol. 36 issue 4, 2010, pp. 876-910.

[5] E.A.C. Bittner and J.M. Leimeister, "Creating shared understanding in heterogeneous work groups: Why it matters and how to achieve it", Journal of Management Information Systems, vol. 31 issue 1, 2014, pp. 111-144.

[6] J.J. Marcel, P.S. Barr and I.M. Duhaime, "The influence of executive cognition on competitive dynamics", Strategic Management Journal vol. 32 issue 2, 2010, pp. 115-138.

[7] K.M. Langfield-Smith and A. Wirth, "Measuring differences between cognitive maps", Journal of the Operational Research Society, vol. 43 issue 12, 1992, pp. 1135-1150.

First steps of Data Collection Planning for Development of Artificial Intelligence Technology for the European Crowdfunding Aggregator Platform

Olga Bogdanova[1] and Mariia Kozlova[1]
[1]LUT University, School of Business and Management,
Lappeenranta, Finland

***Abstract*—This extended abstract outlines the first planning steps of developing an artificial intelligence system for investment crowdfunding. The system will match investment projects with interested investors taking into account multiple factors, such as project attributes, risk-attitude and return expectations, as well as existing behavioral patterns of the investors. The first stage of the system development is collecting historical data from the cross-European crowdfunding aggregator website. However, with a rare opportunity to collect such an extensive, the design of the data collection should be carefully thought though. This abstract aims to provide a short overview of the data collection planning for the further use in the thorough development of data collection design.**

***Keywords*—crowdfunding, investors' behaviour, data collection.**

I. Introduction

Investment crowdfunding is a relatively new form of financing that allows startups and small and medium enterprises (SMEs) to raise funding in form of equity or debt for their projects through the Internet. The research on investment crowdfunding has been a growing body of the academic and professional literature in the recent years, with the main focus on project and company attributes that drive investors' attention, decision to invest and, ultimately, the success in raising funding. According to [1], the majority of authors investigate the single factors affecting the investors' decisions, and more recently there has been attempts to predict the outcome of the funding round by synthesizing those factors in the models and applying machine learning techniques [1].

Nevertheless, scarce of research exist on behavior of investors [2], mainly due to the lack of data. The investment crowdfunding platforms are not willing to cooperate and to provide their data to the researchers, some databases for the research were created by collecting information from the websites manually and therefore, the investor-level data is incomplete, the sample sizes are small, and many variables are missing [3, 4]. Limited number of observations implies other limitations of the current research, leaving many questions unanswered, such as the effect of the size, the age of the company, and the project type on success in equity crowdfunding. Some researchers [5] expect that the cross-country research to have great potential, however, conducting this research will require creating a bigger database on investment crowdfunding in cooperation with the crowdfunding platforms.

A cross-European database on investors' profiles, transactions and the invested projects would allow to apply the machine learning to better understand, who the investors are, how they select investment opportunities, and what kind of criteria they use in decision making, what kind of factors affect their behavior, especially their individual traits, previous behavior and behavior of others, and how these factors interact. Understanding behavior of various groups of investors and triggers of investment decisions will become a foundation for creation of the model for the recommender system, which will allow to match the projects with the investors and predict the funding behavior of investors.

Developing the process of data collection for such a cross-European investment crowdfunding database is the aim of this study.

II. The project description

The future cross-European database is a part of the project on developing an artificial intelligence (AI) system for the investment crowdfunding. This project will be run in cooperation with the Privanet Group Oyj, research input from LUT University, and access to the databases created at Crowhow platform. Privanet Group Oyj is a Finnish investment service provider, that operates an equity crowdfunding and a real estate lending platforms and online marketplace for trading the securities of unlisted companies through the Group company Privanet Securities Oy. Crowhow AI is a Business Finland funded project, an initiative of the Privanet Group Oyj. The objective of the Crowhow AI project are to create an international crowdfunding aggregator and a secondary market for unlisted shares. The aggregator will work as a service, providing information and possibility to invest to the ongoing funding rounds of the leading European crowdfunding platforms, mainly the members of European Crowdfunding Network. The service aims to help investors find suitable investment opportunities as smoothly as possible from among growth companies looking for crowdfunding by applying AI.

III. Data collection principles

The process of data collection will include several steps. First, it will start from the literature review on the investment (and other types) of crowdfunding, behavioral finance, customer relationship management, and application of machine learning for predicting customers and investors behavior. As an outcome, and based on the objectives of the research, the list of needed variables will be created. These variables will create the following groups of factors:

- investor-level factors: investor profile and the associated transactional data;
- project-level factors: company/project attributes, company founders' attributes, etc;
- contextual factors: behaviour of other investors, communication with the staff on the platform, platform design, text and visual information of the project description, and other factors, affecting the cognitive processes of the investors.

Based on the list of the variables, the following actions will be defined, such as the structures of investor information questionnaires, structure of transactional data, and the methods of collecting and categorizing textual and visual data. As the final step, the data structures will be implemented into the web-site design, and the data collection will start approximately in September 2019 with the launch of the Crowhow platform. The data collection design therefore, is a critical step in data collection, since it will define the structure of data in the very beginning of data collection. This is an unique opportunity to get a better quality international level data, compared to the datasets used in the crowdfunding research before.

Fig. 1. Steps of data collection

References

[1] J. Beckwith, "Predicting Success in Equity Crowdfunding," *ScholarlyCommons*, 2016.

[2] Hoegen, Andreas & Steininger, Dennis & Veit, Daniel., "An Interdisciplinary Review of Investor Decision-Making in Crowdfunding.," in *International Conference on Wirtschaftsinformatik 2017*, St. Gallen, Switzerland, 2017.

[3] L. S. A. Hornuf, "Crowdinvesting – Angel Investing for the Masses?," in *Handbook of Research on Venture Capital: Volume 3. Business Angels*, Edward Elgar, 2014.

[4] Johannes Wallmeroth, Peter Wirtz, Alexander Groh, "Institutional Seed Financing, Angel Financing, and Crowdfunding of Entrepreneurial Ventures: A Literature Review.," *HAL*, 2016.

[5] Vanacker, Tom & Vismara, Silvio & Walthoff-Borm, Xavier, "What happens after a crowdfunding campaign?," in *Handbook of Research on Crowdfunding*, 2019.

Fuzzy similarity and entropy (FSAE) feature selection revisited by using intra-class entropy and a normalized scaling factor

Christoph Lohrmann
School of Business and Management
Lappeenranta-Lahti University of Technology
Lappeenranta, Finland
christoph.lohrmann@lut.fi

Pasi Luukka
School of Business and Management
Lappeenranta-Lahti University of Technology
Lappeenranta, Finland
pasi.luukka@lut.fi

Keywords - Feature Ranking, Filter method, Supervised feature selection, Machine Learning

I. INTRODUCTION

Feature selection is a dimensionality reduction technique that selects a feature subset from the existing features in a dataset [1]. The objective of feature selection is to reduce the dimensionality of a dataset by discarding features that are irrelevant and that deteriorate the ability of a classifier to assign observations well to classes [2], [3]. It exists essentially in two forms, a supervised and an unsupervised version. In the first form, the class of observations is known, whereas in the latter form the class labels are unknown [4]. In the context of this paper, we focus on the supervised feature selection methods introduced by Lohrmann et al. [5] and Luukka [6]. Both of these methods, the fuzzy similarity and entropy (FSAE) feature selection [5] and the feature selection method by Luukka (2011) [6] are so-called filter techniques. Filter techniques are next to wrapper functions and embedded functions, one of the three types that feature selection methods can be divided into [7]. The defining characteristic of supervised filter methods are that they embody a part of the feature pre-processing and discard features premised on each feature's characteristics and consistency [7], [8]. In this paper, we suggest an adaptation of the feature selection method of Luukka (2011) and the FSAE using intra-class similarity and, for the FSAE, a normalization of its scaling factor.

II. METHODS

The feature selection approach by Luukka (2011) [6] computes the similarity between all observations and the ideal vector of each class and, subsequently, deploys an entropy measure such as from De Luca & Termini [9] or Parkash et al. [10] to determine how "informative" each feature is. The features that are least informative, and because of that possess the highest entropy, are suggested for removal.

The FSAE approach was recently introduced and is an adaptation of the feature selection method by Luukka (2011). It deploys a scaling factor to account for the distance between class representatives [5]. The class- and feature-specific scaling factors are multiplied with the feature- and class-specific entropy values to adjust their level of informativity. In other words, they allow for each feature to adjust them according to how far each class representative (ideal vector) is from the other classes' representatives.

In this paper, we suggest to use intra-class similarity values, meaning that only the similarity of observations with their own class's ideal vector is considered. This is in contrast to the previous version of the FSAE and the method by Luukka (2011) and replaces the idea to incorporate the similarity of all observations to all class ideal vectors. Moreover, for the FSAE the scaling factor for the class- and feature-specific entropy values is normalized by the variation in the observations in the different classes. The objective of the first change is to make the entropy of classes only dependent on each classes' variation. Moreover, for the FSAE, the normalization of the scaling factor aims to account for the variation in the classes when calculating the distance between their representatives.

III. DATA

The datasets created to illustrate the improvement of the feature selection method by Luukka (2011) and the FSAE are three artificial examples including two features.

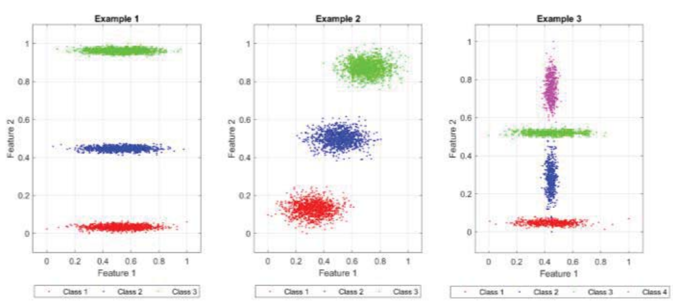

Figure 1: Artificial examples for feature selection

The three artificial examples have in common that the first feature is completely or at least to a large extent overlapping for all classes. Moreover, this feature is less relevant than the second feature that can basically linearly separate all classes.

IV. Results and Conclusion

The original feature selection method by Luukka (2011) and the FSAE are compared with both, the adjusted version of the method by Luukka (2011) and the adjusted FSAE using also intra-class similarity and a normalized scaling factor. The feature removal decision for the first artificial example (see Table 1) shows that the adjustment of the approach of Luukka (2011) changed the removal decision from the second, more discriminating feature to the first, irrelevant one. The FSAE in its original and adjusted version both select the second feature and discard the first one. However, the feature ranking with the adjusted FSAE is more distinct than of the original algorithm.

Example 1	Feature 1	Feature 2	Removal
Original Luukka 2011	2773.9 55.0 %	3385.1 45.0 %	2
Luukka 2011 (adjusted)	924.6 12.7 %	134.1 87.3 %	1
Original FSAE	2766.8 32.9 %	1354.4 67.1 %	1
FSAE (adjusted)	101960.0 0.0 %	4.6 100.0 %	1

Table 1: Results first artificial example

A similar situation is encountered for the second artificial example. The original method by Luukka (2011) suggests the removal of the second feature, which is capable to linearly separate the three classes. The adjusted method by Luukka (2011) and both FSAE methods correctly suggest the removal of the first feature.

Example 2	Feature 1	Feature 2	Removal
Original Luukka 2011	3897.5 51.5 %	4133.7 48.5 %	2
Luukka 2011 (adjusted)	739.6 33.7 %	376.4 66.3 %	1
Original FSAE	2875 42.1 %	2093.5 57.9 %	1
FSAE (adjusted)	566.7 9.5 %	59.5 90.5 %	1

Table 2: Results second artificial example

In the third example, both the original method by Luukka (2011) and the FSAE suggest to remove the second feature, which overlaps for at least 2 classes and for values in the range from 0.4 to 0.5 even for all classes. The two adjusted approaches both correctly select the first feature to be discarded. However, the decision made by the adjusted FSAE is clearer in terms of the variable importance values.

Example 3	Feature 1	Feature 2	Removal
Original Luukka 2011	1517.3 71.6 %	3828.5 28.4 %	2
Luukka 2011 (adjusted)	379.1 38.6 %	238.0 61.4 %	1
Original FSAE	1514.1 60.5 %	2321.7 39.5 %	2
FSAE (adjusted)	33832.0 0.2 %	53.1 99.8 %	1

Table 3: Results third artificial example

As demonstrated, adjusting the feature selection method of Luukka (2011) and the FSAE improves the feature ranking. Using intra-class similarity and normalizing the scaling factor of the FSAE ensures for all artificial examples that the irrelevant / less relevant feature is discarded.

V. References

[1] J. Kittler and K. V Mardia, “Statistical pattern recognition in image analysis,” *J. Appl. Stat.*, vol. 21, no. 1,2, pp. 61–75, 1994.

[2] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Comput. Electr. Eng.*, vol. 40, pp. 16–28, 2014.

[3] G. Dougherty, *Pattern Recognition and Classification: An Introduction*. New York: Springer ScienceBusiness Media, 2013.

[4] J. Liang, S. Yang, and A. Winstanley, “Invariant optimal feature selection: A distance discriminant and feature ranking based solution,” *Pattern Recognit.*, vol. 41, pp. 1429–1439, 2008.

[5] C. Lohrmann, P. Luukka, M. Jablonska-Sabuka, and T. Kauranne, “A combination of fuzzy similarity measures and fuzzy entropy measures for supervised feature selection,” *Expert Syst. Appl.*, vol. 110, 2018.

[6] P. Luukka, “Feature selection using fuzzy entropy measures with similarity classifier,” *Expert Syst. Appl.*, vol. 38, pp. 4600–4607, 2011.

[7] A. L. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” *Artif. Intell.*, vol. 97, pp. 245–271, 1997.

[8] M. Seo and S. Oh, “CBFS: High performance feature selection algorithm based on feature clearness,” *PLoS One*, vol. 7, no. 7, pp. 1–10, 2012.

[9] A. De Luca and S. Termini, “A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory,” *Inf. Control*, vol. 20, pp. 301–312, 1972.

[10] O. Parkash, P. Sharma, and R. Mahajan, “New measures of weighted fuzzy entropy and their applications for the study of maximum weighted fuzzy entropy principle,” *Inf. Sci. (Ny).*, vol. 178, pp. 2389–2395, 2008.

On the use of pairwise comparisons for absolute-type evaluation

Jan Stoklasa
School of Business and Management,
LUT University, Lappeenranta, Finland,
and
Palacký University Olomouc, Czech Republic,
Department of Applied Economics,
jan.stoklasa@lut.fi

Jana Siebert
Palacký University Olomouc, Czech Republic,
Department of Applied Economics,
jana.siebert@upol.cz

***Abstract*—Pairwise comparison methods are a frequently used tool for the calculation of weights of criteria, evaluations of alternatives, for the ordering of objects and for multiple-criteria decision-making and evaluation in general. This paper presents a progress report on the development of a novel absolute-type evaluation method based on pairwise comparisons. We discuss how various types of scales can be converted into artificial absolute scales as a prerequisite for absolute-type evaluation and their further use in the derivation of absolute-type evaluations. We conclude that using the so called (un)saturation thresholds it is possible to use any type of criteria with any type of underlying scale and obtain absolute type evaluations of the alternatives via pairwise comparisons.**

Keywords*—*pairwise comparison, absolute evaluation, multiple-criteria decision-making, priorities, absolute-rating, acceptance, saturation, threshold

I. Introduction

The popularity of pairwise comparison methods lies in their simplicity and ease of use (see e.g. [1], [2], [3], [4]). The requirements of these methods on the abilities and knowledge of the decision-makers are low, yet the methods are able to provide relative-type evaluations required for the selection of the best alternative out of the set of alternatives, for the ordering of alternatives and for the evaluation of these alternatives. The practical use of these methods is wide [5], [6], ranging from the evaluation of scientific monographs [7] and works of arts [8], [9], [10], [11], to various economic and industrial applications [12].

Pairwise comparison methods are based on the simple act of comparing pairs of objects (alternatives, criteria). The set of evaluated objects therefore needs to contain at least two elements to be compared. The evaluations computed by these methods are relative to the set of compared objects. It is therefore clear that some applications of these methods are ruled out. Mainly one of the goals of the decision-making is impossible to achieve by relative-type evaluations. This goal is the decision on acceptability or non-acceptability of an alternative. Simply put, the best alternative in the set of alternatives might still not be good enough to be accepted. There are many real-life situations that require the ability to decide on the (non)acceptability of an alternative - credit scoring for loans, personnel selection, investment decisions, etc. The need for absolute-type methods that would be able to assist in these cases is therefore clear.

Absolute-type methods are being proposed [13], [14] in the literature. The utilization of pairwise comparisons, which are easy to administer and simple to explain to laymen decision-makers, is, however, not frequent.

II. Absolute evaluation

For the purposes of this report we define absolute-type evaluation as such evaluation, that has the following properties:

- it is provided on an absolute-type scale,
- for each alternative separately we are able to decide on its acceptance/non-acceptance based on the evaluation,
- the evaluation of an alternative does not change if the set of alternatives changes, i.e. if an alternative is added to or removed from the set of alternatives,
- the evaluation of an alternative does not change when the evaluations of other alternatives with respect to one or more criteria change.

An absolute scale is such a scale that is bounded and contains a natural maximum and minimum value. It is therefore meaningful to consider absolute-type scales to be cardinal[1] in their nature (see e.g. [15]). Obviously, many real-life scales are not absolute. We therefore suggest a method for the transformation of scales that are not absolute into artificial absolute scales. We suggest how to introduce these artificial absolute scales for the frequently used types of scales[2] including the nominal ones. Our proposal defines the artificial minima and maxima as representative values (or subsets) of the sets of *completely unacceptable* and *completely acceptable* values of the original scales respectively. The sets of completely acceptable and

[1] In some cases also a bounded ordinal scales e.g. the qualitative scale (low, below-average, average, above-average, high) can have the required property of absolute-type scales, i.e. the existence of a natural minimum and maximum. However, the values of such scales (or their numerical representations) might not constitute a cardinal scale. The existence of natural minima and maxima might still suffice to obtain overall evaluations on an absolute-type scale.

[2] Nominal, ordinal, cardinal-ratio-type and cardinal-interval-type non-absolute scales are considered. Other types of scales are left out of the scope of this paper.

unacceptable values of the scale are defined using a fuzzy set[3] representing acceptable values of the scale. Note, that these fuzzy sets can be defined for nominal, ordinal and cardinal scales, and also for all types of criteria, including benefit-, cost- and target-value-type[4] criteria. The kernels of these fuzzy sets constitute the sets of completely acceptable values, the complements to the supports of these fuzzy sets constitute the sets of completely unacceptable values.

Saaty claims in [17] that the absolute level of evaluation is already present in his Analytic Hierarchic Process (AHP), and calls it *absolute measurement AHP*. The absolute measurement AHP is capable of providing evaluations that are stable with respect to the changes in the set of alternatives and with respect to the changes in the evaluations of other alternatives with respect to the criteria. However, the method results in an unnecessary loss of information, and, Saaty does not propose a way to use the results for the decisions on (non)acceptance of alternatives. The resulting evaluation scale is also not absolute in this method.

III. A SUMMARY OF THE PROPOSED ABSOLUTE-TYPE EVALUATION METHOD

Having defined the artificial absolute scales, we can construct pairwise comparison matrices for the comparison of the (artificial) minimal/maximal values of the scales with the evaluation of each alternative with respect to each criterion. The (artificial) minimal/maximal values of the criteria scales thus define an artificial *ideal* and *anti-ideal* alternative. We calculate the local priorities of the alternatives (i.e. of the actual alternative, of the ideal alternative and of the anti-ideal alternative) with respect to each criterion by the geometric means method (GMM) [18]. These local priorities of actual, ideal and anti-ideal alternatives are aggregated into global priorities by weighted geometric mean and interpreted in terms of ratios, as advocated in [18]. Such ratios of priorities represent evaluations on absolute-type scales, which can be used for the decisions on acceptance/non-acceptance of alternatives and also fulfill all the requirements we have introduced for an absolute-type evaluation.

IV. CONCLUSION

In this paper we present the results of our work-in-progress on an absolute-type evaluation method based on pairwise comparisons. We propose a definition of absolute-type evaluation and a method for the conversion of non-absolute scales into artificial absolute ones using fuzzy sets. Finally, we compute the global priorities of the alternatives and an artificial ideal and anti-ideal alternative by GMM and compute the ratios of the evaluations of the alternatives to the evaluations of the ideal and anti-ideal alternatives. These ratios are the required absolute-type evaluations, that can be used for the decisions on acceptability or non-acceptability of an alternative.

[3] We assume standard definitions of a fuzzy set, its kernel and support as proposed by Zadeh in [16].

[4] Criteria which aim on achieving a specific value, e.g. square footage of a flat, or daily time spent at work (it might not be reasonable to commute to spend e.g. 10 minutes at work, and we might not want to spend more than 8 hours there, so something around 4-5 hours might represent a “sweet spot”), etc.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support received from the Finnish Strategic Research Council, grant number 313396 / MFG40 - Manufacturing 4.0 and by the Specific University Research Grant IGA_FF_2019_002, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2019.

REFERENCES

[1] T. L. Saaty, *Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process*. Pittsburgh: RWS Publishers, 2000.

[2] ——, “Decision making with the analytic hierarchy process,” *International Journal of Services Sciences*, vol. 1, no. 1, pp. 83–98, 2008.

[3] ——, “Relative measurement and its generalization in decision making why pairwise comparisons are central in mathematics for the measurement of intangible factors the analytic hierarchy/network process,” *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, vol. 102, no. 2, pp. 251–318, sep 2008.

[4] 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*, 2nd ed. New York, Dordrecht, Heidelberg, London: Springer, 2013.

[5] A. Ishizaka and A. Labib, “Review of the main developments in the analytic hierarchy process,” *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 336–14 345, 2011.

[6] N. Subramanian and R. Ramanathan, “A review of applications of Analytic Hierarchy Process in operations management,” *International Journal of Production Economics*, vol. 138, no. 2, pp. 215–241, 2012.

[7] J. Krejčí and J. Stoklasa, “Fuzzified AHP in the evaluation of scientific monographs,” *Central European Journal of Operations Research*, vol. 24, no. 2, pp. 353–370, 2016.

[8] V. Jandová, J. Krejčí, J. Stoklasa, and M. Fedrizzi, “Computing interval weights for incomplete pairwise-comparison matrices of large dimension - a weak consistency based approach,” *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 6, pp. 1714–1728, 2017.

[9] V. Jandová, J. Stoklasa, and J. Talašová, “Modification of the AHP based model for evaluating artistic production of Czech colleges,” in *Proceedings of the 32nd International Conference on Mathematical Methods in Economics*, 2014, pp. 372–377.

[10] J. Stoklasa, V. Jandová, and J. Talašová, “Weak consistency in Saaty’s AHP - evaluating creative work outcomes of Czech Art Colleges Classification of works of art,” *Neural Network World*, vol. 23, no. 1, pp. 61–77, 2013.

[11] J. Stoklasa, T. Talášek, and J. Talašová, “AHP and weak consistency in the evaluation of works of art a case study of a large problem,” *International Journal of Business Innovation and Research*, vol. 11, no. 1, pp. 60–75, 2016.

[12] 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*. New York: Springer, 2006.

[13] V. Jandová and J. Talašová, “Evaluation of absolute type in the Partial Goals Method,” in *Proceedings of the 33rd International Conference on Mathematical Methods in Economics*, 2015, pp. 321–326.

[14] V. Sukač, J. Talašová, and J. Stoklasa, “Soft’ consensus in decision-making model using partial goals method,” in *Proceedings of the 34th International Conference on Mathematical Methods in Economics*. Liberec: Technical University of Liberec, 2016, pp. 791–796.

[15] S. S. Stevens, “On the Theory of Scales of Measurement,” *Science*, vol. 103, no. 2684, pp. 677–680, 1946.

[16] L. A. Zadeh, “Fuzzy sets,” *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.

[17] T. L. Saaty, “Absolute and relative measurement with the AHP.The most livable cities in the United States,” *Socio-Economic Planning Sciences*, vol. 20, no. 6, pp. 327–331, 1986.

[18] J. Krejčí and J. Stoklasa, “Aggregation in the analytic hierarchy process: Why weighted geometric mean should be used instead of weighted arithmetic mean,” *Expert Systems with Applications*, vol. 114, pp. 97–106, 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0957417418303981

Developing a Decision Tree Block for the Exercise Boundary Fitting Method

Christoph Lohrmann[1] and Mariia Kozlova[1*]
[1]LUT University, School of Business and Management,
Lappeenranta, Finland
*corresponding author, e-mail: mariia.kozlova@lut.fi

***Abstract*—This extended abstract aims at developing a decision tree block for a real option model based on the exercise boundary fitting method. The method allows for multiple real options and multiple sources of uncertainty, while also complying with financial theory. When multiple real options are introduced, the decision tree that represents all the possible combinations of exercising the available options can become complex. It should be automated for the purpose of preserving flexibility of the model. And therefore, the whole process of decision tree construction should be generalized. This abstract presents the solution to this problem. With this development the real option model rises to a new level and enables wider variety of cases to be analyzed.**

Keywords—decision tree, simulation, real options, exercise boundary fitting

I. INTRODUCTION

This extended abstract presents the development of a block for a novel real option approach [1] based on the exercise boundary fitting method [2]. The exercise boundary fitting method adopts an idea of optimizing a threshold price for exercising an option to the real options framework. The method potentially allows incorporating multiple real options and multiple sources of uncertainty into a simulation model that complies with financial theory [2]. This enables modeling complex real-world investment cases, when all other approaches fail, analytical systems cannot be solved and trees and least-squares Monte Carlo models explode in the exponentially growing number of scenarios.

The method has been applied for a mining investment [1] with multiple real options, including delaying and staging the investment. However, one of the hurdles of its implementation was the automatization of the decision tree construction for different number of plant modules and decision points. This extended abstract presents the solution to this problem.

II. THE EXERCISE BOUNDARY FITTING MODEL

The model consists of several blocks presented on Fig.1.

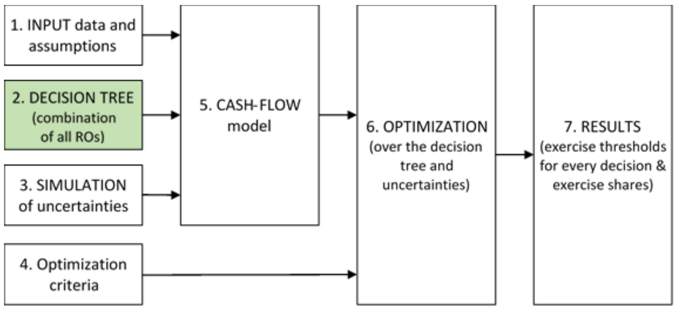

Fig. 1. Schematic representation of the model

Numerical assumptions (block 1), the decision tree (block 2) that represents all possible scenarios of exercising available real options, simulated uncertainty (block 3) e.g. product prices that follow mean-reverting or any other process are all inputs to the cash flow model (block 5) that dictates the logic of revenues and costs for every decision tree branch. The cash flow model assigns each decision tree scenario to every simulated path of uncertainty and calculates the value of every path. Which decision branch is realized at which uncertainty path depends on the overall value maximization (block 6). The optimization parameters are the threshold values of the uncertainty variables (or functions for two and more variables) at each decision node of the tree. As a result, the model delivers (block 7) these exercise boundaries (values or functions) and the corresponding share of paths that have 'fired' at every decision node.

III. THE DECISION TREE

The decision tree (block 2, Fig.1) is a combination of all real options available over the span of the decision points. The real options considered are to delay investment and to stage it. Therefore, at every decision node the investor has a choice to build either nothing or any number of modules up to the maximum possible size. Such a tree for three decision points and maximum size of 2 modules is presented on Fig.2. In case of a single course of uncertainty, a stochastic price, every decision in the tree is characterized by threshold prices. E.g. at the first decision point *t1*, building one module is profitable only for the simulation paths, where current price is higher than S_1, and two modules if it is higher than S_2. In case the optimization results into $S_1 = S_2$, this means that with the current assumptions and uncertainties, building one module at *t1* is not relevant at all and the exercise share will be zero.

Acknowledgements: The authors would like to acknowledge the financial support received from the Finnish Strategic Research Council, grant number 313396 / MFG40 - Manufacturing 4.0.

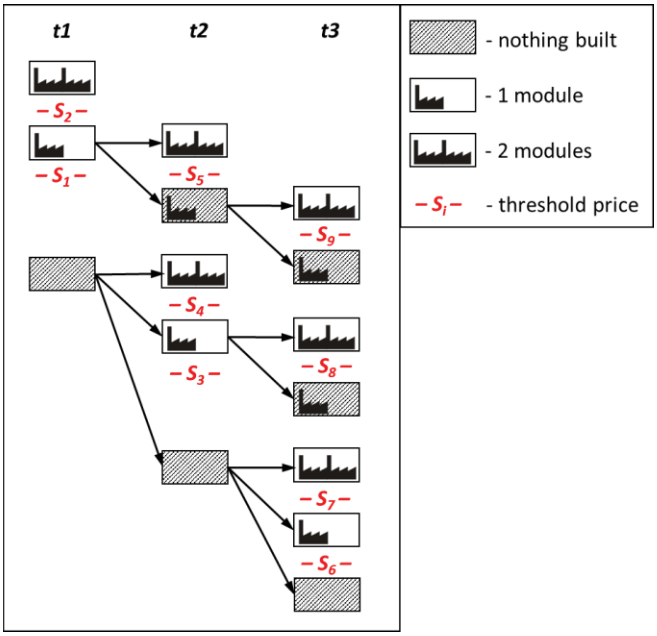

Fig. 2. Decision tree for maximum size of 2 modules and 3 decision times

The programming problem addressed in this abstract is to transfer this tree structure and the corresponding threshold price indices into a matrix form and to generalize it for any number of decision points t and any maximum size of the plant, n. The output matrices should be of the size $t*Nb$, where Nb is the total number of branches of the tree, e.g. it is equal to 10 for the case in Fig.2.

IV. IMPLEMENTATION

The construction of the required matrix is implemented based on the idea to solve the money changing problem repeatedly for different size options (or number of modules) – since n or less modules can be build overall – and for all possible building stages – since the sum of modules built during all stages can be built in a different number of stages (e.g. all n in one or more stages). Subsequently, a number of zeros for the remaining stages are added to each of the solutions so that each solution contains entries (zero or non-zero) for all stages. After that, all permutations of these solution vectors are calculated and added together for all stages and size options. Finally, all entries in the resulting matrix are sorted in descending order to represent the preferred order of the decision tree. The calculations were implemented in Matlab and existing code for the money changing problem was deployed as part of the coded solution [3]. The entire code is presented in Appendix I and is illustrated for the case of a maximum of two modules and three decision points as in Fig.2.

REFERENCES

[1] M. Kozlova, Y. Lawryshyn and A. Bashiri, "Exercise Boundary-based Simulation with Application to Multi-Option Mining Investments," in Annual International Real Options Conference, 2019.

[2] A. Bashiri, M. Davison and Y. Lawryshyn, "Real Option Valuation Using Simulation and Exercise Boundary Fitting–Extended Abstract," in Annual International real Options Conference, 2018.

[3] J. D'Errico, "Partitions of an integer," 2018. Available at https://se.mathworks.com/matlabcentral/fileexchange/12009-partitions-of-an-integer

```
clear all
close all
clc
```

define n and t

```
n=2;

t=3;
```

Calculation of scenarios (using 'money changing problem' and permutations)

```
permMatall=[zeros(1,t)]; % initially one line of zeros
seqn=1:n;

for i=1:min(n,t) % for each number of items used (e.g. only 2 number used)
    for j=max(1,i):n % for each sum of elements (e.g. sum has to be 3)
        plist=partitions(j,seqn,[],i); % "money changing problem"
        % code for 'partitions' function:
https://se.mathworks.com/matlabcentral/fileexchange/12009-partitions-of-an-integer
        if isempty(plist)
        else
            temp=[reshape(repelem(repmat(seqn,1,size(plist,1)), reshape(plist',1,[]))', [],
size(plist, 1))', zeros(size(plist,1),t-i)]; % % Repeat copies of array elements (element-wise
according to vector that defines number of replications); adding suitable number of zeros
            permMat=[]; % empty permMat
            for k=1:size(temp,1)
                permMat=[permMat; unique(perms(temp(k,:)),'rows','stable')]; % unique function
is needed since 0 0 are seen as two different values for permutations (hence 0 0 1 and 0 0 1
are both listed in results)
            end
            permMatall=[permMatall;permMat];
        end
    end
end

permMatallSorted=sortrows(permMatall,[1:size(permMatall,2)],'descend');
transpermMatallSorted=permMatallSorted'; % transposed version time x options
```

Second part - s_low, s_high and inequality matrix

```
s_low=zeros(size(transpermMatallSorted));

s_high=zeros(size(transpermMatallSorted));

clear A_opt
A_opt_counter=0;
b_opt=zeros(1,size(transpermMatallSorted,2)-1);
Snum=0; % starting threshold number for S

for i=1:t % for each time period

    uniquePerm=fliplr(unique(transpermMatallSorted(1:i,:)','rows','stable')'); % find unique
and start from right side (to left)
    uniquePermNoZeroEnd=uniquePerm(:,uniquePerm(i,:)~=0); % all unique combinations that do
not end on a zero (since for them no scenario needs to be counted - otherwise count is to
high) - for s_low
```

```
    uniquePermNoSumN=uniquePerm(:,sum(uniquePerm(1:i,:),1)<n); % all unique combinations that
do not sum up to n (since these either are for at least one round at n, so s_high is zero, or
they just reached n this round and s_high = s_max)

    for j=size(transpermMatallSorted,2):-1:1 % backwards through the scenarios

        % s_low
        if transpermMatallSorted(i,j)==0
            %s_low(i,j) is automatically zero
        else
            % s_low
            % Determine index by comparison with (ordered) unique Permutation matrix
            [~, loc]=ismember(transpermMatallSorted(1:i,j)',uniquePermNoZeroEnd','rows');
            s_low(i,j)=Snum+loc; % previous t's final (so highest) scenario number + current
number of unique combination (from the right, so reversed)

        end

        % s_high
        if sum(transpermMatallSorted(1:i,j))==n & transpermMatallSorted(i,j)==0 % if current
one is zero and already maximum reached
            % s_high(i,j) is zero since current element is zero and already sum is n
        elseif sum(transpermMatallSorted(1:i,j))==n & transpermMatallSorted(i,j)~=0 % if with
this round the maximum of n was reached
            s_high(i,j)=size(transpermMatallSorted,2); % set to maximum
        else % so if maximum is not reached until now (no matter what number < n appeared)
                [~, loc]=ismember(transpermMatallSorted(1:i,j)',uniquePermNoSumN','rows');
                s_high(i,j)=Snum+loc;

        end
    end


    % inequality matrix
    % For the current time step (i) (row) consider all rows were at this
    % time step is no zero e.g. 1 0 is discarded / not considered - and use all unique
combinations

    % uniquePermNoZeroEnd - unique combinations until this time step (ordered from right to
left)
    % uniquePermBeforeNow - unique combinations of uniquePermNoZeroEnd from 1 to i-1 timestep

    if i==1 % special case if first time step since there are no previous paths
        s_consecutive=uniquePermNoZeroEnd(1:i,:)';
        for l=1:size(s_consecutive,1)-1 % rule for consecutive items that share the same
previous path
            A_opt(A_opt_counter+l,:)=zeros(1,size(transpermMatallSorted,2)-1); % so that for
all other values it is zero
            A_opt(A_opt_counter+l,s_consecutive(l))=1;
            A_opt(A_opt_counter+l,s_consecutive(l+1))=-1;
        end
        A_opt_counter=A_opt_counter+size(s_consecutive,1)-1; % adjust A_opt_counter
    else

        uniquePermBeforeNow=unique(uniquePermNoZeroEnd(1:i-1,:)','rows','stable')';
        for k=1:size(uniquePermBeforeNow,2)
            [~, loc]=ismember(uniquePermNoZeroEnd(1:i-1,:)',uniquePermBeforeNow(:,k)','rows');

            if sum(loc)>1 % so if at least two outcomes (that do not end on a zero) have the
same previous path -> inequality / relationship needs to be established
                % index inequality matrix: look up the scenario S in the s_low matrix at the
time i, where the two (or more) scenarios occur that have the same path up to time i-1

s_consecutive=sort(unique(s_low(i,ismember(transpermMatallSorted(1:i,:)',uniquePermNoZeroEnd(:
,find(loc))','rows'))),'ascend')';

                for l=1:size(s_consecutive,1)-1 % rule for consecutive items that share the
same previous path
                    A_opt(A_opt_counter+l,s_consecutive(l))=1;
```

```
                    A_opt(A_opt_counter+l,s_consecutive(l+1))=-1;
                end
                A_opt_counter=A_opt_counter+size(s_consecutive,1)-1; % adjust A_opt_counter
            else
                % nothing done since only a single threshold price for this path that does not
need to be compared
            end
        end

    end

    % Adjust of maximum scenario of previous step
    Snum=Snum+size(uniquePermNoZeroEnd,2);
end


transpermMatallSorted
s_low
s_high
```

```
transpermMatallSorted =

     2     1     1     1     0     0     0     0     0     0
     0     1     0     0     2     1     1     0     0     0
     0     0     1     0     0     1     0     2     1     0

s_low =

     2     1     1     1     0     0     0     0     0     0
     0     5     0     0     4     3     3     0     0     0
     0     0     9     0     0     8     0     7     6     0

s_high =

    10     2     2     2     1     1     1     1     1     1
     0    10     5     5    10     4     4     3     3     3
     0     0    10     9     0    10     8    10     7     6
```

Analysing Peaking Generator Availability in PJM

The Effect of the Reliability Pricing Model

Stein-Erik Fleten, Magne Ledsaak, Sigurd Mehl, Ola Røssum
Norwegian University of Science and Technology
Trondheim, Norway

Benjamin Fram
Norwegian School of Economics
Bergen, Norway

Carl Ullrich
James Madison University
Harrisonburg, VA, USA

***Abstract*— This paper aims to study the effects of capacity payments on the operational decisions of plant managers for peaking units in the PJM Interconnection. We achieve this through a structural estimation of maintenance and switching costs between the operational state, the standby state and retirement of generating units. We have focused on the period from 2001 throughout 2016—a period where we have identified some significant changes in the power market dynamics. We conduct a counterfactual analysis on the level of capacity payments to study the effects of introducing a capacity market in 2007. The reliability of the power system depends crucially on the availability of flexible peaking units to cover load in periods of high demand. Therefore, an understanding of the real costs facing the owners of these units is essential in order to enforce policies that ensure sufficient peak capacity in the power system. Capacity markets are introduced as a means of compensating capacity, and our study aims to analyze the effects of this additional market on switching behavior.**

The empirical data shows less switching between states after the introduction of capacity remunerations. We find that the role of peaking units has changed, with the units being dispatched more often. In the counterfactual analysis, we find a clear connection between the level of capacity payments and switching. We conclude that the current level of capacity payments in PJM incentivizes peaking units to stay in the operational state.

Keywords—electricity markets; security of supply; real options; Capacity remuneration

I. Introduction

In order to design a market where peak generators are compensated appropriately to secure sufficient investment activity, a regulator must have a thorough understanding of the generators' cost structure as well as the market dynamics. Regulators make cost estimates, but empirical testing of such estimates is difficult. Generator costs are influenced by exogenous factors that can be hard to observe. Also, the cost structure of a power producer is business sensitive information, as this determines the lower limit of their bids in market auctions. Therefore, the empirical estimation of generator costs is one of few viable options for investigating the real costs faced by generators. The business decisions of an owner of a peaking unit are readily formulated as a sequential decision process in time, where choices about the operational state of the generator must be made before each consecutive time period. Markov decision processes provide an excellent framework for modeling sequential decision making under uncertainty [1]. Under the assumption that the generator owners act rationally, dynamic programming provides a way of identifying the optimal decision rule for choosing how to operate one's generator. The agent can be represented through a set of economic primitives, describing their utility function, transition probabilities and discount factor for future states. The primitives convey information about the decision process of the generator owner as well as the uncertainty of the decision environment. Structural estimation provides a framework for robustly estimating such primitives.

II. Results and Conclusions

We find evidence that market conditions for peaking units in the PJM has changed significantly after 2007, and identify three market trends influencing the behavior of peaking units. Technological advancements have changed the supply side of the natural gas market, giving a persistent drop in fuel prices for gas-fired turbines. New environmental regulations have forced old coal-fired baseload into retirement, presenting new market opportunities for gas-fired units. We also see that the regulations have led to the retirement of old combustion turbines. The introduction of capacity payments has led to less switching and a higher amount of peaking plants being ready to operate.

The first trend, the penetration of shale gas in the US gas market, significantly reduced the fuel price for many generators. We conclude that this has disrupted the traditional market dynamics where coal-fired plants serve as baseload, and combustion turbines cover peak demand. Gas-fired turbines have become more competitive in serving baseload, and besides, traditional baseload has been punished harder by stricter environmental regulations than gas units. Consequently, peaking units are now dispatched more often, increasing the wear and tear on the mechanical equipment. This is a plausible explanation for the increase in the estimated maintenance cost for generators after 2007. The second effect that influences the switching behavior of peak generators is the introduction of stricter environmental regulation schemes. In years where regulatory changes are expected, our estimates show that the perceived cost of startup decreases and the perceived cost of shutdown increases. This tendency to prefer to operate in years with new regulations must be seen in light of the fact that environmental regulations are imposed on all actors in the power market. Coal-fired baseload is more polluting than most other technologies and is therefore affected more severely by stricter environmental regulations. Gas is cleaner, has become cheap, and gas plants are quick to bring online. This makes it possible for gas-fired units to replace the retiring coal-fired baseload, a fact reflected in the environmental regulation coefficient estimate.

Finally, after the introduction of the RPM, less switching is observed, and the share of operational peaking generators is larger, with few generators being in the standby state. The results from the counterfactual analysis indicate that the switching behavior is affected by the level of the capacity payments. Lowered capacity payments will give more switching, whereas increased payments cause minimal change. Overall, our findings indicate that the system operator is successful in incentivizing peaking generators to stay in an operational-ready state through capacity payments.

References

[1] Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. Econometrica, pages 999–1033.

Extended MobileNet convolutional neural network with Bayesian approach for economic applications

Aleksandr Popkov
Department of Information Systems in Economics
Faculty of Economics St. Petersburg State University
St. Petersburg, Russia;
Email: alr.popkov@gmail.com

Mikhail Zaboev
Department of Information Systems in Economics
Faculty of Economics St. Petersburg State University
St. Petersburg, Russia;
Lappeenranta University of Technology,
LUT School of Business and Management
Email: m.zaboev@spbu.ru

***Abstract*—In this article we present the results of the investigation how variational Bayesian inference technique based on Bayes by Backprop algorithm can perform in a modern specific convolutional neural network (CNN) suggested by Google Inc. also known as MobileNet. Additionaly, we open and broaden possible applications of the results in economic practice.**

I. Introduction

Machine learning technique of discovering valuable information from data individuals in a dataset became one of the widespread technology of modeling and analyzing the current state of various complicated processes. Recent success in this field was achieved thanks to the mathematical model called an artificial neural network [1-3]. This connectionists system became valuable help as one of the alternative problem solvers in many spheres like healthcare, retail, security, mobile application development, and other industry.

The standard artificial neural network uses point estimation of its neurons weights in a network layer. This called frequentist method that assumes a huge amount of data to learn through backpropagation [4]. The frequentist approach had a meaningful impact on the artificial neural networks research and development, data collection and warehousing and still stayed as a successful approach in practice. However, there is a problem of overfitting in the neural network in the general case which is mainly solved by dropout technique [5]. Moreover, the model became overconfident on the inputs that are not related to the types of data the model was trained on. In order to get the model outperforming human vision, we still need lots amount of training data to generalize the estimation of the model and still need to apply the tricks of avoiding model overfitting. In order to smooth the issues in developing artificial neural networks and convolutional neural networks especially, we suggest considering the current state of the Bayesian approach to the problem [6]. We implemented Bayesian version of MobileNetV1 and analyze how developing variational Bayesian inference [7] for convolutional neural networks based on Bayes by Backprop [8] can impact on the results of the frequentist lightweight convolutional neural network MobileNetV1. As a result, we open possible economic applications of this model.

II. The Methodology

There has been a wide range of developing different methods related to the neural networks with a goal to consider such fundamental property like robust, to make a model more certain about its predictions, preventable from overfitting and being able to acquire the obscure properties and generalization of an object without collecting a lot of data [5,9,10,11].In [8] new Bayesian method Bayes by Backprop was proposed. Further explorations of the method led to the adaptation not only for fully connected neural networks but also for recurrent and convolutional neural networks [12,13].

Our extended Bayesian MobileNet was built on the original MobileNet architecture [14]. The versions of MobileNet architecture stand on depthwise separable convolution operations which were used in [15] and also implemented in Inception model [16]. The depthwise separable convolutions were placed as network modules in order to get the model with high speed during its inference and thus to provide best computation solutions. The convolutional neural network MobileNet was invented specifically to exploit it in mobile devices and other mobile technology where such properties like computational power and memory have strict constraints and can be considered today as one of the state–of–the–art models in neural networks.

The key concept of the latter Bayesian method can be considered from the information theory. When we train the weights of a neural network we approximate a distribution $q(w|\theta)$, where θ – the parameters of the distribution we are able to control, w – the weights of a neural network model. In addition to the distribution $q(w|\theta)$, we understand that the real process $p(w)$ is unobserved and unknown. We observe only the distribution $p(w|D)$, where D is the data individuals. To compare $q(w|\theta)$ and $p(w|D)$ we apply one of the most common metrics between distributions named relative entropy or Kullback–Leibler divergence:

$$KL(q(w|\theta) \parallel p(w|D)) = \int q(w|\theta) log \frac{q(w|\theta)}{p(w|D)} dw \quad (1)$$

Then we exploit the Bayesian rule w.r.t. network weights w and data individuals D :

$$p(w|D) = \frac{p(D|w)p(w)}{\int p(D|w)p(w)dw} = \frac{p(D|w)p(w)}{p(D)} \quad (2)$$

In (2) $p(D|w)$ is the likelihood of the evidence that our hypothesis about the weights is correct, p(w) is our prior knowledge and p(D) – the evidence or the likelihood or the evidence under any circumstances. Considering (2) in Kullback-Leibler divergence (1) we get:

$$\int q(w|\theta) log \frac{q(w|\theta)}{p(w|D)} dw = \int q(w|\theta) log \frac{q(w|\theta)p(D)}{p(D|w)p(w)} dw =$$

$$= \int q(w|\theta) log \frac{q(w|\theta)}{p(w)} dw - \int q(w|\theta) log p(D|w) dw + log p(D) \quad (3)$$

As a result of (3), we may get an optimization problem which can be solved using the so called local reparmetrization trick [11]:

$$\theta^{\star} = argmin\ E_{q(w|\theta)} \left(log q(w|\theta) - log p(D|w) - log p(w) \right) \quad (4)$$

In the case of convolutional neural networks, one more feature was introduced in the Bayesian neural network paradigm [13]. Instead of direct weights w sampling the operation of sampling layer activations b was suggested:

$$b = A \star \mu + \varepsilon \odot \sqrt{A^2 \star (\alpha \odot \mu^2)}, \varepsilon \sim N(0,1) \quad (5)$$

In (5) $\star$ means the convolution operation, $\odot$ component-wise multiplication, α the receptive field in a convolution, the variance multiplier.

First, we need to update the outputs of a CNN by the frequentist inference operation. The single point estimate made during frequentist inference is treated as the mean μ of $q(w|\theta)$ in (8). The second convolution operation describes how much neural network deviates from the estimation of μ. This procedure also continues in fully connected layers at the end of the convolutional network. Since the variance part of the expression (5) is also determined by the mean , the second convolution operation is aimed at learning the parameter α.

All layers in the Bayesian version of MobileNet as in the origin MobileNetV1 are followed by batch normalization and ReLU activation function. After the sequence of convolutions, we have an average pooling operation and get fully connected layer without ReLU nonlinearity. BConv in (Table 1) means Bayesian version of convolution and BFullyConnected means Bayesian fully connected neural network layer.

The Bayesian MobileNet was trained in Pytorch framework [7] using the Adam method of stochastic optimization. We also use less regularization technique and limit our experiments by the real dataset sizes and did not artificially augment it. In order to get comparable results we also trained the original frequentist model MobileNetV1 with the same learning parameters as for the extended Bayesian MobileNet.

TABLE I. BAYESIAN MOBILENET BODY ARCHITECTURE.

Type / Stride	Kernel Size
BConv / s2	input channels x3x3x32
BConv d2 / s1	3x3x3x32 dw
BConv / s1	1x1x32x64
BConv dw / s2	3x3x64 dw
BConv / s1	1x1x64x128
BConv dw / s2	3x3x128 dw
BConv / s1	1x1x128x256
BConv dw / s1	3x3x256 dw
BConv / s1	1x1x256x256
BConv dw / s2	3x3x256 dw
BConv / s1	1x1x256x512
5x BConv dw / s1 Conv / s1	3x3x512 dw 1x1x512x512
BConv dw / s2	3x3x512 dw
BConv / s1	1x1x512x1024
BConv dw / s2	3x3x1024 dw
BConv / s1	1x1x1024x1024
Average Pool / s1	Pool 7x7
BFullyConnected / s1	1024X1000
Softmax function / s1	1024 $\rightarrow$ num classes

III. EXPERIMENTS

The experiments with Bayesian MobileNet architecture were conducted on three image datasets: MNIST, Fashion MNIST, and CIFAR10 [17-19].

The MNIST dataset of handwritten digits consists of 60,000 training and 10,000 validation images and has dimension 1x28x28. The Fashion MNIST dataset similarly to MNIST consists of 60,000 training and 10,000 validation images and has the same dimension. The CIFAR-10 dataset of 60,000 RGB images with three color channels in 10 classes, with 6,000 images per class, each image has a dimension 3x32x32. Each of the classes has 5,000 training images and 1,000 validation images (Fig.1).

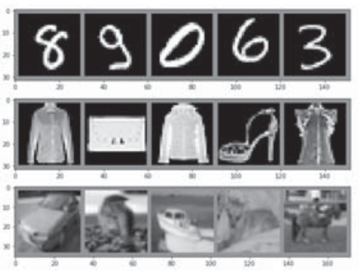

Fig. 1. From top to bottom: MNIST, FashionMNIST, CIFAR10 datasets.

IV. CONCLUSION

First of all, we evaluated the performance of Bayesian MobileNetV1. Table 2. shows the resulting validation accuracy for our model among three datasets in comparison to the original MobileNetV1.

MNIST and Fashion MNIST training batch sizes were set to 32 data individuals, for CIFAR10 we reduce the number of batch size to 16 to successfully run the computing. For MNIST and Fashion MNIST we also limit the number of training epochs to 10, for CIFAR10 we expanded the number of epochs to 50 in order to get more extensive training perspective for multiple channeled data. Additionally, the Bayesian model

TABLE II. CLASSIFICATION COMPARISON OF BAYESIAN MOBILENETV1 AND MOBILENETV1.

Model	Datasets validation accuracy		
	MNIST	F-MNIST	CIFAR10
Batch Size	32	32	16
Bayesian MobileNetV1, Samples = 10	0,99	0,89	0,70
MobileNetV1	0,99	0,92	0,76

takes 10 samples to get the estimation of the model parameters to learn.

All the initial iterations for both Bayesian and non-Bayesian MobileNet start with frequentist estimation so the initial training data is the same for both models. On MNIST dataset for Bayesian MobileNetV1, we have comparable results to MobileNetV1. Despite the high speed of getting 0,99 level of accuracy for the frequentist MobileNetV1 (Fig.2), there is much more valuable information that can be revealed in the Bayesian perspective. The Bayesian perspective gives information about the starting point of overfitting the model on MNIST data. On Fashion MNIST and CIFAR10 dataset, the Bayesian MobileNetV1 did not reach the same results as the frequentist model. However, the plots for Bayesian training on these datasets also have an indication of overfitting. The results achieved for the CIFAR10 indicates that for multiple channeled data more advanced sampling technique can be applied as it was not enough to get less volatile accuracy during training iterations.

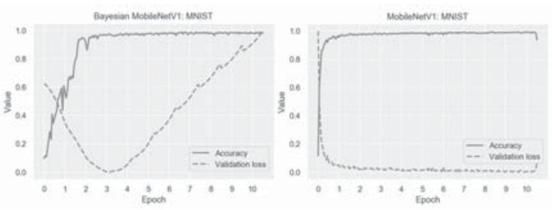

Fig. 2. An example of Bayesian MobileNet training. MNIST dataset.

To sum it up, we discovered that the Bayesian framework introduced in Bayes by Backprop [8] can be considered as one of the valuable frameworks for alternative neural network representation and training consequently. Different complex frequentist network architectures can be modified in order to get more reliable training results and some level of uncertainty in the neural network. At the same time, the Bayesian framework can be less computationally effective as it needs to make samples to estimate weights in a neural network and requires more memory to store the parameters of probability distributions.

The results open a discussion to a wide range of economic applications of Bayesian convolutional neural networks where low level of false positive estimation is required and where collecting huge amount of training data is quite difficult or even impossible. First of all, systems of reliable objects detection and data aggregation based on Bayesian CNN can be valuable in agriculture equipment for automotive field service with cultivated crops. Second, more traditional banking and postal systems can be improved with such models in order to provide reliable processing for documents, cards, securities, parcel information and user-friendly mobile applications for document registration. Third, the retail analytics can be augmented with such applications like goods recognition on the store shelf in order to verify the location of goods. Additionally, Bayesian neural networks can be a reliable model to analyse customers, their in-store behaviour, understanding hot and cold shop zones, canceled or pending goods. All the possible applications mentioned can be considered as explicit analysis for data collection and data understanding. In spite of the explicit analysis we can also consider the application of the model in implicit analysis. Thus, the Bayesian neural networks can be a part of more complicated system. For example, the model can provide new features from images like web advertisements, cinema posters or cinema numbers of shots in order to set up soft and desirable advertisements for users, recommend new cinema from the extracted cinema shots information, to build the customer profile and at the same time meet his or her hidden needs with new recommendations.

REFERENCES

[1] Janai J., Behl A., Geiger A. Computer Vision for Autonomous Vehicles: Problems, Datasets and State–of–the–Art // arXiv:1704.05519 [cs.CV]. 2017.

[2] Gu J. et al. Recent Advances in Convolutional Neural Networks // Pattern Recognit. 2006. P. 1-38.

[3] Rai H.G.N., Krishna P.R. Video Analytics Solution for Tracking Customer Locations in Retail Shopping Malls. 2011. P. 773-776.

[4] Rumelhart D.E., Hinton G.E., Williams R.J. Learning representations by back-propagating errors // Nature. 1986. Vol. 323. P. 533-536.

[5] Hinton G. Dropout: A Simple Way to Prevent Neural Networks from Overfitting // J. Mach. Learn. Res. 2014. Vol. 15. P. 1929-1958.

[6] Mullachery V., Husain A., Khera A. A study of Bayesian Neural Networks // arXiv1801.07710 [cs, stat]. 2015. P. 1-16.

[7] Kingma D.P., Salimans T., Welling M. Variational Dropout and the Local Reparameterization Trick // arXiv:1506.02557 [stat.ML]. 2015. Mcmc. P. 1-14.

[8] Blundell C. et al. Weight Uncertainty in Neural Networks // arXiv:1505.05424 [stat.ML]. 2015. Vol. 37.

[9] Zhang T. et al. Comparison Network for One-Shot Conditional Object Detection // arXiv:1904.02317 [cs.CV]. 2019.

[10] Mangal R., Nori A. V., Orso A. Robustness of Neural Networks: A Probabilistic and Practical Approach // arXiv:1902.05983 [cs.LG]. 2019. i.

[11] Araujo A. et al. Robust Neural Networks using Randomized Adversarial Training // arXiv:1903.10219 [cs.LG]. 2019.

[12] Fortunato M., Blundell C., Vinyals O. Bayesian Recurrent Neural Networks // arXiv:1704.02798 [cs.LG]. 2018. P. 1-14.

[13] Shridhar K. et al. Bayesian Convolutional Neural Networks with Variational Inference // arXiv:1806.05978 [cs.LG]. 2018. P. 1-11.

[14] Howard A.G., Wang W. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications // arXiv:1704.04861 [cs.CV]. 2017.

[15] Chollet F. Xception: Deep learning with depthwise separable convolutions // Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017. 2017. Vol. 2017Janua. P. 1800-1807.

[16] Szegedy C. et al. Going deeper with convolutions // arXiv:1409.4842 [cs.CV]. 2014. P. 1-12.

[17] Posch K., Pilz J. Correlated Parameters to Accurately Measure Uncertainty in Deep Neural Networks // arXiv:1904.01334 [cs.LG]. 2015. Vol. 14, 8. P. 1-14.

[18] Krizhevsky A. Learning Multiple Layers of Features from Tiny Images // Tech. report, Univ. Toronto. 2009.

[19] Vollgraf R. Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms // arXiv:1708.07747 [cs.LG]. 2017. P. 1-6.

Optimizing maintenance scheduling of multi-machine industrial systems with simulation – example from metals mining

Jyrki Savolainen
School of Business of Management
LUT-University
Lappeenranta, Finland
jyrki.savolainen@lut.fi

Michele Urbani
Department of Industrial Engineering
University of Trento
Trento TN, Italy
michele.urbani@unitn.it

Abstract—**This paper deals with the optimal maintenance policy selection in complex industrial systems consisting of several interconnected pieces of equipment. A combined approach of simulation-optimization and system dynamic (SD) modeling is applied. We present a method where optimized maintenance schedules of the available equipment are used as an input in an SD-model of an industrial process. The simulation process is demonstrated using a load-haul-dump (LHD) system from the mining industry, where the objective function is the total throughput of material within a specified period of time.**

Keywords—Predictive Maintenance; System dynamics; Simulation; Optimization

I. INTRODUCTION

Machine-specific maintenance history data allow predicting and flexibly planning future service schedules. The timing of these maintenance events have direct consequences on the production capacity of an overall industrial system. Maintenance in industry is traditionally considered as a *necessary evil*. As stated by [1] and [2], maintenance costs can rise up to 60 % of total production cost. In the case of critical, large-scale industrial equipment, the question of maintenance scheduling is no means a trivial one. However, when considering multi-machine environments (i.e. systems), optimization models are sparse due to the problem complexity. In this research, we present and discuss preliminary results from a simulation model usable for testing maintenance policies in a system setting. The purpose is to study how the combination of machine-specific maintenance policies effect the system's capacity within a planning timeframe.

For a state-of-the art review on simulation-based maintenance optimization, we advise the reader to refer to [3]. In general, maintenance activities can be divided into corrective, preventive, and condition-based. The work of Wang [4] summarizes a multitude of maintenance policies and replacement models available in the literature for single- and multi-component deteriorating systems. A further development of corrective and preventive policies is predictive maintenance. Mobley [2] shows that predictive maintenance is more than a maintenance policy: it serves as a tool for maintenance management, plant optimization, and reliability improvement. Frontiers in modern maintenance are also usually represented by condition-based maintenance (CBM) and prognostic and health management (PHM) [5], [6]. A key concept of predictive maintenance is to fit distributions into the existing historical data of maintenance. Taylor and others [7] illustrate a basic data analyses to determine equipment availability. Lashgari & Sayadi [8], show how to use equipment data to regress estimate of future maintenance costs.

The role of equipment reliability in mining is highlighted by the complexity of new, advanced machinery and pressures from the competitive environment to meet the production targets of an individual mine (see discussion, e.g., [9]). A data-driven analysis of maintenance policy optimization in mining systems faces the problem of reliability behavior of equipment. As a key challenge to maintenance, [10], point out the number of mobile / semi-mobile equipment hindering collection of "clean" datasets. Data collection is also inhibited by electronic-based system hardware failures (e.g. sensors, wiring, connectors, etc.) commonly occurring in harsh mining environments (see [9]).

In this paper, we focus on the predictive maintenance of mining industry. Through simulation, we aim to find out, what is the optimal amount and timing of maintenance to be conducted for an LHD-configuration that serves as a critical part of the overall metal extraction process.

II. Methodology

The simulation model consists of two parts: a maintenance module and a system dynamic (SD) model of a metal mine. The former is created as a Matlab script, which runs the optimization simulation and feeds the input into the latter (SD-model) (see **Figure 1**). By simulating the whole loop for n-times we are able to uncover the effects of maintenance planning under uncertainty.

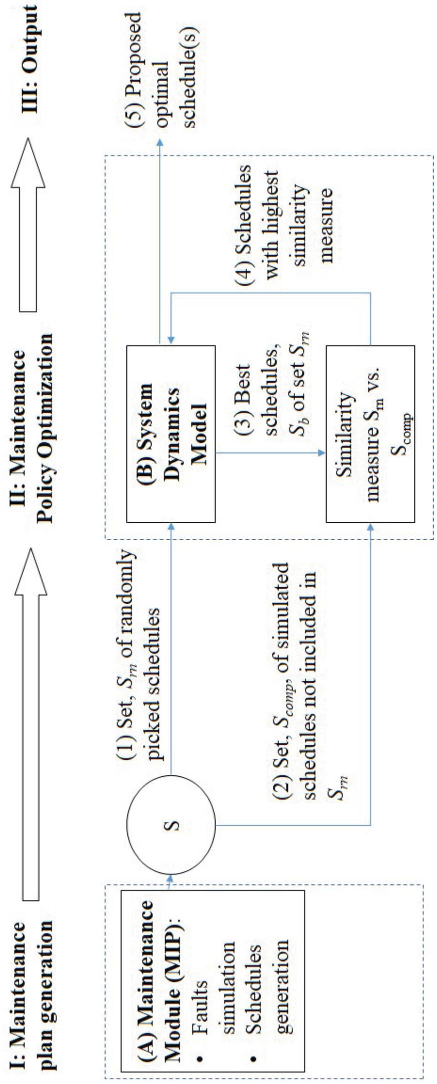

Fig. 1. Illustrative representation of the optimization process

The maintenance module ("(A)" in Figure 1) aims at maximizing the availability of the system under limited capacity of the maintenance workshop. The equipment is owned and managed by the company and the capacity of the workshop (performing maintenance) is limited. The maintenance module produces a set of alternative maintenance schedules, *S*, that are used as an inputs in a system dynamic (SD) model of the overall industrial system ((1)). The SD-model (B) contains feedback-loops and second-order interactions, which means that we have to test, which of the feasible maintenance schedule(s) produce the maximum output of the system as a whole. The candidates for the optimal maintenance schedule are selected based on an n-folded sampling of the possible set of solutions (1). To avoid local optima, we use similarity number –based methods to examine close-to-optimal solutions ((3) and (4)) before giving the simulated suggestion of the proposed solution ((5)).

III. Results and Discussion

We demonstrate the maintenance schedule optimization methodology using a realistic load-haul-dump (LHD) example from the metals mining industry. Monte Carlo simulations (MCS) together with sensitivity analyses (SA) are applied to discuss and verify the attained results.

References

[1] M. Bevilacqua and M. Braglia, "The analytic hierarchy process applied to maintenance strategy selection," *Reliab. Eng. Syst. Saf.*, vol. 70, no. 1, pp. 71–83, 2000.

[2] R. K. Mobley, *An Introduction to Predictive Maintenance*. Elsevier, 2002.

[3] A. Alrabghi and A. Tiwari, "State of the art in simulation-based optimisation for maintenance systems," *Comput. Ind. Eng.*, vol. 82, pp. 167–182, 2015.

[4] H. Wang, "A survey of maintenance policies of deteriorating systems," *Eur. J. Oper. Res.*, vol. 139, no. 3, pp. 469–489, 2002.

[5] N.-H. Kim, D. An, and J.-H. Choi, *Prognostics and Health Management of Engineering Systems*. Springer, 2017.

[6] Y. Peng, M. Dong, and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *Int. J. Adv. Manuf. Technol.*, vol. 50, no. 1–4, pp. 297–313, 2010.

[7] P. Taylor, N. Vagenas, N. Runciman, S. R. Clément, and S. R. Clement, "International Journal of Surface Mining , Reclamation and Environment A methodology for maintenance analysis of mining equipment A methodology for maintenance analysis of mining equipment," no. May 2013, pp. 37–41, 2007.

[8] A. Lashgari and A. R. Sayadi, "Statistical approach to determination of overhaul and maintenance cost of loading equipment in surface mining," *Int. J. Min. Sci. Technol.*, vol. 23, no. 3, pp. 441–446, 2013.

[9] B. S. Dhillon, "Mining Equipment Reliability," in *Mining Equipment Reliability, Maintainability, and Safety*, London: Springer London, 2008, pp. 57–70.

[10] R. Hall and L. K. Danshmend, "Reliability and maintainability models for mobile underground haulage equipment," *CIM Bull.*, vol. 96, no. Jun/Jul, pp. 159–165, 2003.

Using Boundary Fitting and Machine Learning to Value Multi-Stage Real Option Investments

Yuri Lawryshyn

***Abstract*—Real option analysis is recognized as a superior method to quantify the value of real-world investment opportunities where managerial flexibility can influence their worth, as compared to standard discounted cash-flow methods typically used in industry. However, realistic models that try to account for a number of risk factors can be mathematically complex, and in situations where many future outcomes are possible, many layers of analysis may be required. Recently, we introduced a real options valuation methodology geared towards practical use with mining valuation as a context. A key innovation of the methodology is the idea of fitting optimal decision making boundaries to optimize the expected value, based on Monte Carlo simulated stochastic processes that represent important uncertain factors. In this paper we will briefly present theoretical and numerical aspects associated with the methodology in the context of a Bermudan put option, then discuss convergence and stability issues for utilziing the model to value an American put option, and finally, present a deep neural network formulation that could ultimately be expanded to a complex real option mining case study.**

***Index Terms*—Real Options; American Options; Boundary Fitting; Mining Valuation.**

I. Introduction

Real option analysis (ROA) is recognized as a superior method to quantify the value of real-world investment opportunities where managerial flexibility can influence their worth, as compared to standard net present value (NPV) and discounted cash-flow (DCF) analysis. The ability for managers to react to uncertainties at a future time adds value to projects, and since this value is not captured by standard DCF methods, erroneous decision making may result (Trigeorgis (1996)). An excellent empirical review of ex-post investment decisions made in copper mining showed that fewer than half of investment timing decisions were made at the right time and 36 of the 51 projects analyzed should have chosen an extraction capacity of 40% larger or smaller (Auger and Guzman (2010)). The authors were unaware of any mining firm basing all or part of their decision making on the systematic use of ROA and emphasize that the "failure to use ROA to assess investments runs against a basic assumption of neoclassical theory: under uncertainty, firms ought to maximize their expected profits". They make the case that irrational decision making exists within the industry due to a lack of real option tools available for better analysis. A number of surveys across industries have found that the use of ROA is in the range of 10-15% of companies, and the main reason for lack of adoption is model complexity (Hartmann and Hassan (2006), Block (2007), Truong, Partington, and Peat (2008), Bennouna, Meredith, and Marchant (2010), Dimitrakopoulos and Abdel Sabour (2007)).

The focus of this research is the development of a real options valuation methodology geared towards practical use with mining valuation as a context. A key innovation of the methodology is the idea of fitting optimal decision making boundaries to optimize the expected value, based on Monte Carlo simulated stochastic processes that represent important uncertain factors. In this paper we will briefly present theoretical and numerical aspects associated with the methodology in the context of a Bermudan put option, then discuss convergence and stability issues for utilziing the model to value an American put option, and finally, present a deep neural network formulation that could ultimately be expanded to a complex real option mining case study.

II. Methodology

A. Bermudan Put Option

For the Bermudan put option, we consider a GBM stock price process, S_t, as

$$dS_t = rS_t dt + \sigma S_t d\widehat{W}_t, \tag{1}$$

where r is the risk-free rate, σ is the volatility and $\widehat{W}_t$ is a Wiener process in the risk-neutral measure. We assume the payoff of the option to be $\max(K-S_t, 0)$ and can be exercised at times $t = \tau$ and $t = T$ where $\tau < T$. The value of the put option can be written as

$$V_0 = e^{-r\tau} \int_0^\infty \max\left(K - x, P_{BS_{put}}(x, \tau, T, r, \sigma, K)\right) f_{S_\tau}(x|S_0)dx, \tag{2}$$

where $P_{BS_{put}}(x, \tau, T, r, \sigma, K)$ is the Black-Scholes formula for the value of a European put option with current stock price x, maturity $T - \tau$, risk-free rate r, volatility σ and strike K, and $f_{S_\tau}(x|S_0)$ is the density for S_τ given S_0. As can be seen in equation (2), the optimal exercise occurs when

$$K - \theta^* = P_{BS_{put}}(\theta^*, \tau, T, r, \sigma, K), \tag{3}$$

where θ^* is used to denote the exercise price at $t = \tau$. Equation (3) can be solved using numerical methods and thus the option value simplifies to

$$\begin{aligned} V_0 =& e^{-r\tau} \left(\int_0^{\theta^*} (K - x) f_{S_\tau}(x|S_0)dx + \right. \\ & \left. \int_{\theta^*}^\infty P_{BS_{put}}(x, \tau, T, r, \sigma, K) f_{S_\tau}(x|S_0)dx \right), \end{aligned} \tag{4}$$

which, too, can be solved using standard numerical methods.

To explore numerical issues regarding the proposed boundary fitting methodology in the context of the Bermudan put option, we simulate N risk-neutral paths for S_t. For a given exercise price θ at $t = \tau$, the value of the option for the i-th path is given by

$$\begin{aligned} V_0^{(i)}(\theta) =& \mathbb{1}_{S_\tau^{(i)} \leq \theta} \left(K - S_\tau^{(i)} \right) e^{-r\tau} + \\ & \mathbb{1}_{S_\tau^{(i)} > \theta} \max \left(K - S_T^{(i)}, 0 \right) e^{-rT}, \end{aligned} \tag{5}$$

where $S_t^{(i)}$ represents the value of S_t of the i-th simulated path. The optimal exercise price can then be estimated as

$$\theta^* = \arg\max_\theta \frac{1}{N} \sum_{i=1}^N V_0^{(i)}(\theta), \tag{6}$$

and the option value estimate becomes

$$V_0^{sim} = \frac{1}{N} \sum_{i=1}^N V_0^{(i)}(\theta^*). \tag{7}$$

Note that $\lim_{N\to\infty} V_0^{sim} = V_0$, as required.

B. American Put Option

While there is no analytical solution for the value of an American put option, we utilize a binomial tree to estimate the value and the exercise boundary. To utilize the proposed exercise boundary fitting simulation method, we again simulate N risk-neutral paths for S_t and assume the exercise boundary to be function of the form $h(t|\vec{\eta})$, where $t \in [0, T]$ and $\vec{\eta} = [\eta_1, \eta_2, ..., \eta_n]'$ is a vector of constant parameters. We define the first passage of time when the i-th simulation path, $S_t^{(i)}$ (equation 1), hits the boundary as

$$\tau^{(i)} \equiv \min\{t > 0, T \,:\, S_t^{(i)} \leq h(t; \vec{\eta})\}, \tag{8}$$

where T is the time to maturity of the option.

The option value of the i-th path is thus

$$\begin{aligned} V_0^{(i)}(\vec{\eta}) =& \mathbb{1}_{\tau^{(i)} < T} \left(K - S_t^{(i)} \right) e^{-r\tau^{(i)}} \\ & + \mathbb{1}_{\tau^{(i)} = T} \max \left(K - S_T^{(i)}, 0 \right) e^{-rT}, \end{aligned} \tag{9}$$

where K is the strike price. The optimal parameters can be determined by using methods similar to those above, where

$$\vec{\eta}^* = \arg\max_{\vec{\eta}} \frac{1}{N} \sum_{i=1}^N V_0^{(i)}(\vec{\eta}), \tag{10}$$

from which the option value can be estimated as

$$V_0^{sim} = \frac{1}{N} \sum_{i=1}^N V_0^{(i)}(\vec{\eta}^*). \tag{11}$$

C. Using Neural Network Models to Value American Put Options[1]

While fitting a two-dimensional boundary for a single spacial (risk-factor) dimension and time is somewhat straightforward, when we apply the boundary fitting methodology in a more realistic real options setting with multiple risk-factors, the fitting of the boundary becomes significantly more complex. To overcome this issue, we propose to utilize a neural network framework where a multidimensional problem should be relatively straightforward to solve. For proof of concept, in this work we consider only an American put option – i.e. a two dimensional problem.

Our proposed methodology consists of the following main steps in the context of an American put option whose underlying stock value is S_t:

1) Develop a density image representing the probability of the stock process achieving a given value of S at time t.
2) Apply a *deep* neural network where the input is the image and the outputs are discretized (in time, t_i) of S_{t_i} values representing the exercise boundary.
3) Optimize the neural network parameters to maximize the value of the American put option.

The image will be of size $N \times M$ where N is the discretization in stock price S and M is the discretization in time t where $t \in [0, T]$. Each cell will consist of a value between 0 and 1 representing the probability of $S_t \in [S_t - \Delta S, S_t + \Delta S]$, where ΔS is the discretization step in S. These probability values can be determined from the theoretical distribution of S_t or, especially for general processes, through Monte Carlo simulation.

The deep neural network will consist of one or more convolution layers, each followed by ReLU activation and maxpooling, then followed by fully connected layers (number to be determined). The output of the model will be $\hat{S}_{t_i}$ where $t_i \in [0, T]$, $i \in 1, 2, ..., n$ are n user determined time intervals for defining the exercise boundary with $t_n = T$.

Monte Carlo simulated paths will be used to determine the value of the put option for the given based on the $\hat{S}_{t_i}$ values. Maximizing the value by adjusting the neural network parameters will lead to the appropriate valuation and exercise boundary.

D. Results

Histograms of V_0^{sim} of equation (7) resulting from the simulations are presented in Figure 1, where the number of simulation paths was varied from $N = 10^2$ to $N = 10^6$. In each case, 1000 simulations were performed. In Table **??** we present the mean and standard deviation of the 1000 simulation runs for increasing N for V_0 and θ^*. As expected, as N is increased, the values for V_0 and θ^* approach those of the pseudo-analytical solution.

For the American put option, we first explore the use of cubic splines for $h(t; \vec{\eta})$. In Figure 2 we plot the results of the optimization using 4 to 8 node cubic splines. Also plotted in the figure are 20 randomly selected paths for S_t. The nodes are spaced evenly over time. As can be seen, while the error in option calculation is less than 5% on average, the exercise boundaries are oscillatory and exhibit significant error. Next, we explore the use of second, third and fourth order polynomials for $h(t; \vec{\eta})$. The results are plotted in Figure

[1]Note that this work is still under development and full results will be presented at the conference

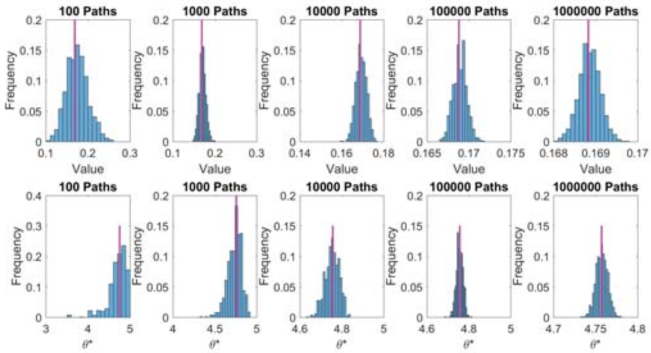

Fig. 1. Histograms of V_0^{sim} for the Bermudan put option (note that each case was simulated 1000 times).

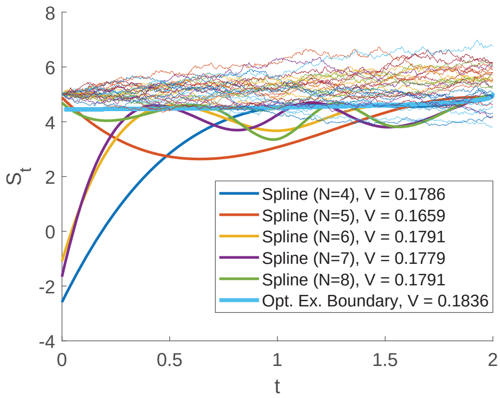

Fig. 2. American put option exercise boundary simulation using cubic splines for $h(t;\vec{\eta})$.

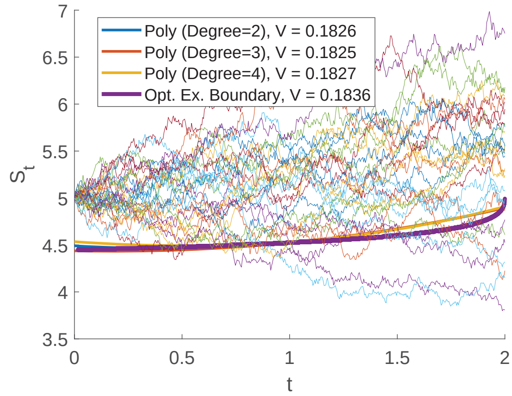

Fig. 3. American put option exercise boundary simulation using polynomial functions for $h(t;\vec{\eta})$.

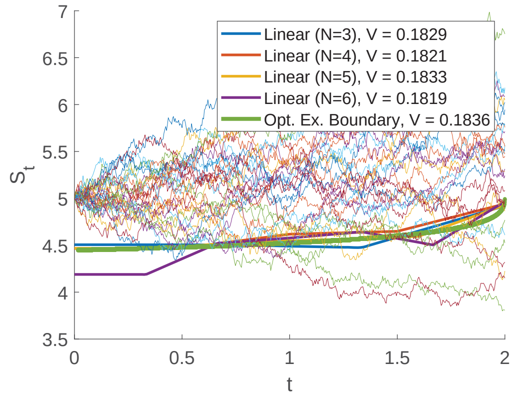

Fig. 4. American put option exercise boundary simulation using piecewise linear functions for $h(t;\vec{\eta})$.

3. The average error is approximately 0.5% and the exercise boundaries are close to the optimal boundary. In Figure 4 we plot the results where we assume a piecewise linear function for $h(t;\vec{\eta})$. As can be seen, as the order increases, the optimal exercise boundary does not necessarily lead to better results. We also note that the exercise boundary was somewhat sensitive to the initial node values assumed for the piecewise function however, while not shown here, the results were further improved by forcing convexity on $h(t;\vec{\eta})$. In general, the results of these trials look promising. We note that the results for the neural network formulation are forthcoming.

References

Auger, F. and J. Guzman (2010). How rational are investment decisions in the copper industry? *Resources Policy 35*, 292–300.

Barone-Adesi, G. (2005). The saga of the american put. *Journal of Banking & Finance 29*, 2909?2918.

Barone-Adesi, G. and R. Whaley (1987). Efficient analytical approximation of american option values. *The Journal of Finance 42*(2), 301–320.

Barraquand, J. and D. Martineau (2007). Numerical valuation of high dimensional multivariate american securities. *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS 30*(3), 383–405.

Bennouna, K., G. Meredith, and T. Marchant (2010). Improved capital budgeting decision making: evidence from canada. *Management Decision 48*(2), 225–247.

Berk, J., P. DeMarzo, and D. Stangeland (2010). *Corporate Finance (Canadian Edition)*. Pearson Canada Inc.

Block, S. (2007). Are real options actually used in the real world? *Engineering Economist 52*(3), 255–267.

Borison, A. (2005). Real options analysis: Where are the emperor's clothes? *Journal of Applied Corporate Finance 17*(2), 17–31.

Brando, L. E., J. S. Dyer, and W. J. Hahn (2012). Volatility estimation for stochastic project value models. *Euro-*

pean Journal of Operational Research 220(3), 642–648.

Broadie, M. and P. Glasserman (1997). Pricing american-style securities using simulation. *Journal of Economic Dynamics and Control 21*, 1323–1352.

Cerny, A. (1999). Minimal martingale measure, capm, and representative agent pricing in incomplete markets. Technical report, Cass Business School. Available at SSRN: http://ssrn.com/abstract=851188.

Collan, M., R. Fullér, and J. Mezei (2009). A fuzzy pay-off method for real option valuation. *Journal of Applied Mathematics and Decision Sciences*, 1–14.

Copeland, T. and V. Antikarov (2001). *Real Options: A Practitioner's Guide*. W. W. Norton and Company.

Datar, V. and S. Mathews (2004). European real options: An intuitive algorithm for the black- scholes formula. *Journal of Applied Finance 14*(1), 45–51.

Del Moral, P., B. Remillard, and S. Rubenthaler (2012). Monte carlo approximations of american options that preserve monotonicity and convexity. In *Numerical Methods in Finance*, pp. 115–143. Springer Berlin Heidelberg.

Dimitrakopoulos, R. and S. Abdel Sabour (2007). Evaluating mine plans under uncertainty: Can the real options make a difference? *Resources Policy 32*, 116?125.

Dixit, A. and R. Pindyck (1994). *Investment under Uncertainty*. Princeton University Press.

Hartmann, M. and A. Hassan (2006). Application of real options analysis for pharmaceutical R&D project valuation?empirical results from a survey. *Research Policy 35*, 343–354.

Jaimungal, S. and Y. A. Lawryshyn (2010). Incorporating managerial information into real option valuation. *Avaiable at SSRN: http://ssrn.com/abstract=1730355*.

Jaimungal, S. and Y. A. Lawryshyn (2011, June). Incorporating managerial information into valuation of early stage investments. In *Proceedings of the Real Options Conference*.

Ju, N. (1998). Pricing an american option by approximating its early exercise boundary as a multipiece exponential function. *The Review of Financial Studies 11*(3), 627–646.

Longstaff, F. and E. Schwartz (2001). Valuaing american options by simulation: A simple least-squares approach. *The Review of Financial Studies 14*(1), 113–147.

Miao, D. W.-C. and Y.-H. Lee (2013). A forward monte carlo method for american options pricing. *The Journal of Futures Markets 33*(4), 369–395.

Nasakkala, E. and J. Keppo (2008). Hydropower with financial information. *Applied Mathematical Finance*.

Pendharkar, P. (2010, March). Valuing interdependent multi-stage it investments: a real options approach. *European Journal of Operational Research 201*(3), 847–859.

Ryan, P. and G. Ryan (2002). Capital budget practices of the fortunre 1000: How things have changed? *Journal of Business and Management 8*(4), 355–364.

Smith, J. and R. Nau (1995, May). Valuing risky projects: Option pricing theory and decision analysis. *Management Science 41*(5), 795–816.

Trigeorgis, L. (1996). *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. Cambridge, MA: The MIT Press.

Truong, G., G. Partington, and M. Peat (2008). Cost-of-capital estimation and capital-budgeting practice in australia. *Australian Journal of Management 33*(1), 95–122.

Tung, H. (2016). Pricing american put options using the mean value theorem. *The Journal of Futures Markets 36*(8), 793?815.

Missing value imputation of the cooling technology for the global thermal power plant fleet using a decision tree classifier

Christoph Lohrmann
School of Business and Management
Lappeenranta-Lahti University of Technology
Lappeenranta, Finland
christoph.lohrmann@lut.fi

Alena Lohrmann
School of Energy Systems
Lappeenranta-Lahti University of Technology
Lappeenranta, Finland
alena.lohrmann@lut.fi

Keywords - Cooling system, Electricity generation, Supervised learning, Machine learning

I. INTRODUCTION

Currently, industrialized countries withdraw a considerable amount of water for cooling purposes in thermoelectric generation. For instance, in the United States freshwater abstractions for thermal power plants account for 40% of all freshwater withdrawals of the country [1]. However, despite this fact, generally no statistical data on the water demand of specific power plants is collected [2]. The data scarcity of the implemented cooling technologies for specific power plants complicates the assessment of the water use in the power generation sector. In previous studies, the assignment of missing information on the cooling technology for specific power plants was conducted by either using (a) market shares of known plants in the region [3], (b) the cooling portfolio mix of the generator-fuel type from known power plants [4], (c) the historical development of different types of cooling systems [5] or (d) the location of specific power plants in relation to different water bodies: large rivers [5], [6], and the ocean coastline [3]. In this paper, we attempt to combine approaches applied previously and propose a new method of “filling up the gaps”, meaning conducting missing value imputation, in the cooling technology data using a decision tree classifier.

II. METHODS

Decision trees are a class of machine learning technique that can address classification and regression problems. The focus in this research will be on classification trees. These trees are applied for classification, which means that they harness the features of observations to assign them to discrete classes [7]. The classes for this application are the five distinct types of cooling technology that any power plant can take. The term ‘feature’ describes an (explanatory) variable or attribute contained in the data that was observed / recorded (e.g. the active capacity of a power plant) or in some form constructed. The objective of a classification tree is to divide and subdivide the feature space [8], [9]. Hence, the algorithm creates starting from the root a tree with branches up to the leaves, where each path to a leaf represents a set of rules for the features that predict the class of each observation [9].

For each of the 145 regions, the top 3 cooling technologies are predicted. The cooling technologies were restricted for each region to the three largest since, first, the top three technologies account for on average 96% of all power plants in all regions, and, second, the minority technologies are frequently so small that they cause a strong class imbalance in the data. For regions with only a single cooling technology or with too few observations for cooling technologies in the minority classes, the assignment to the class was simply conducted via the majority class. For each of the remaining regions, the decision tree is set up based on cross-validation and optimization of the hyperparameter minimum leave size. The results reported in this paper are the mean accuracies for 100 iterations that a decision tree was set up in this manner.

III. DATA

The main source of power plant data for this research is the GlobalData dataset [10], which was complemented and corrected by Farfan and Breyer [11].

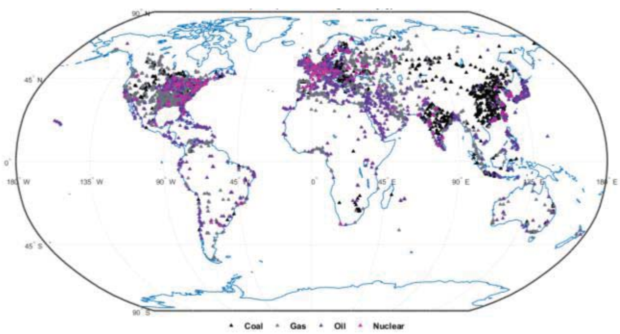

Figure 1: Thermal power plant fleet by fuel type

The presented dataset provides comprehensive information on 13,863 globally active power plants with a total active power plant capacity of 4,182 GW (name, capacity, type of generator, fuel type, commission date, country and region, location coordinates, cooling technology, etc. were reported for each specific power plant). The type of cooling technology was reported for less than 90% of this global thermal power plant capacity and was assigned to one out of five types: dry cooling (or so-called air cooling), once-through cooling (or open loop cooling), recirculating tower cooling systems, recirculating pond cooling and inlet cooling systems.

IV. Results and Conclusion

The methodology of using a decision tree combined with the majority class approach was applied to the 145 global regions in the dataset. In 8 regions the database did not contain information on any power plants in that region, for 40 regions the assignment was conducted via the majority class, and for 97 regions the data was suitable to be used with the decision tree classifier. The benchmark result is a simple assignment of all power plants in a region to the majority class, so the most common cooling technology in that region. This approach yields a mean classification accuracy of 70.85%. Deploying the decision tree together with the majority class was capable to improve this result to 82.78%. In order to highlight the improvement achieved with decision trees, the mean accuracies for the classification using the decision trees versus a simple majority class assignment have to be contrasted. The result of this comparison is displayed in Figure 2.

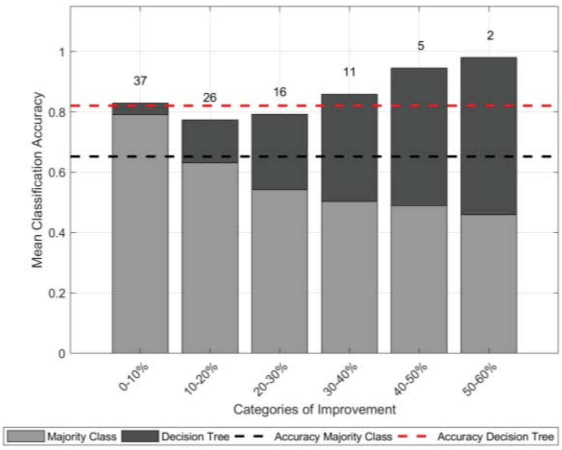

Figure 2: Improvement of the mean classification accuracy

The mean accuracy with the majority class assignment of the 97 regions that can be addressed with the decision tree is 65.23%, which is well below the 70.85% achieved on the entire dataset. The application of decision trees improved the mean accuracy to 82.08%, which is an improvement of almost 17%-points or about 26%. The figure reveals that the largest improvements in accuracy are achieved for those regions where the majority class assignment leads to low mean classification accuracies. Regions with accuracies of on average close to 80% accuracies could only be improved marginally, whereas regions with accuracies of around 50% could be improved by up to 54.04%-points (119.2%).

For this paper, we combined the decision tree classifier together with the simple majority class assignment applied in previous scientific research on the water-energy-nexus. Using this approach allows to combine historical information, the location of power plants, the combustion technology and other information to train a classification model. It was demonstrated that following this methodology with decision trees instead of exclusively using the majority class assignment when sufficient observations are available, can improve the accuracy of the assignment of the cooling technology for specific power plants globally.

V. References

[1] M. A. Maupin, J. F. Kenny, S. S. Hutson, J. K. Lovelace, N. L. Barber, and K. S. Linsey, *Estimated use of water in the United States in 2010*, no. November. 2014.

[2] A. Kohli and K. Frenken, “Cooling water for energy generation and its impact on national-level water statistics,” *Food Agric. Organ. Rome*, 2011.

[3] E. G. R. Davies, P. Kyle, and J. A. Edmonds, “An integrated assessment of global and regional water demands for electricity generation to 2095,” *Adv. Water Resour.*, vol. 52, pp. 296–313, Feb. 2013.

[4] E. S. Spang, W. R. Moomaw, K. S. Gallagher, P. H. Kirshen, and D. H. Marks, “The water consumption of energy production: An international comparison,” *Environ. Res. Lett.*, vol. 9, no. 10, 2014.

[5] S. Vassolo and P. Döll, “Global-scale gridded estimates of thermoelectric power and manufacturing water use,” *Water Resour. Res.*, vol. 41, no. 4, pp. 1–11, 2005.

[6] M. Flörke, E. Kynast, I. Bärlund, S. Eisner, F. Wimmer, and J. Alcamo, “Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study,” *Glob. Environ. Chang.*, vol. 23, no. 1, pp. 144–156, 2013.

[7] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer ScienceBusiness Media, 2006.

[8] W.-Y. Loh, “Classification and regression trees,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 14–23, 2011.

[9] O. Maimon and L. Rokach, “Decision trees,” in *Data mining and knowledge discovery handbook*, 2005, pp. 165–192.

[10] GlobalData Ltd, “GlobalData Power,” *Database*. London, 2014.

[11] J. Farfan and C. Breyer, “Structural changes of global power generation capacity towards sustainability and the risk of stranded investments supported by a sustainability indicator,” *Journal of Cleaner Production*, vol. 141. pp. 370–384, 2017.

A stochastic dynamics model for shaping stock indexes considering self-organization processes, memory, and oscillations

Zhukov Dmitri
Moscow Technological University (MIREA)
Moscow, Russia
zhukovdm@yandex.ru

Tatiana Khvatova
Peter the Great St. Petersburg Polytechnic University
St.Petersburg, Russia
khvatova.ty@spbstu.ru

***Abstract*—Stock market index time series are non-linear and changeable. As such, traditional methods used to model stock market dynamics do not provide sufficient forecasting accuracy. Therefore, new models are required which, by using the technologies of artificial intelligence and machine learning, will enable us to create reliable algorithms for analyzing and forecasting stock index behaviors. Non-parametric models based on the theory of chaos already exist, such as the Huang model. However, in reality, stock markets possess the feature of self-organization and memory of previous events, so they cannot be fully described by the theory of chaos. In this paper, a stochastic dynamics model of stock index formation, which simulates self-organization processes, memory and oscillations, is proposed. An analysis of the Dow Jones and Hang Seng stock indexes was conducted. Based on the data obtained, histograms for the probability density of amplitudes of stock index deviations, depending on the time interval of their calculations, were constructed. These histograms are asymmetric in relation to their maximum points, and oscillations of the probability density were observed. This research demonstrates how and why the dependencies derived cannot be described within the framework of conventional models of diffused or coherent markets. The proposed model fits the data very well regarding the behavior of stock exchange index deviations. The model considers: possible distributional function asymmetry in relation to the maximum; oscillations; shift of the maximal amplitude distribution function; changes of height and width of the distribution which take place over time of amplitude calculation, etc. To create the model, schemes of probabilities of the system's state-to-state transition (changes in stock indexes) were considered. Based on the proposed approach, a non-linear differential equation of the second order was obtained. The boundary problem for defining the probability density function for stock index deviation amplitudes depending on time was resolved. The differential equation contains a term describing potential for self-organization. It also considers the presence of memory. The potential for oscillation occurrence is defined by the initial conditions. This stochastic model can be used for developing investment strategies and for supporting decision-making..**

Keywords—stock exchange index; distribution function; component; self-organization; investment strategy

Features of the Calibration Procedure for AK-type Macroeconomic Models

Lyudmila F. Vyunenko
Lappeenranta-Lahti University of Technology LUT,
LUT School of Business and Management
Lappeenranta, Finland
Liudmila.Viunenko@student.lut.fi
St. Petersburg State University
St. Petersburg, Russia
l.vyunenko@spbu.ru

This study is based on the AK-approach to describing the dynamics of key macroeconomic indicators proposed in [1]. The structure of the model equations allows interpreting them in terms of flows, stocks, feedback loops, and time delays, i.e. to build a system-dynamics simulation model provided the values of all model parameters are somehow determined. Here we concentrate on the problem of calibrating the model under consideration.

The model equations bind gross output Y, capital stock K, consumption C, technological factor A, and capital depreciation rate d. Discrete approximation on the time interval [0, T] with step Δ leads to the equations

$$Y_{t+\Delta} = Y_t + (A_t - d_t)K_t\Delta + A_t K_t \sqrt{\Delta}u_t \quad (1)$$

$$K_{t+\Delta} = K_t + ((A_t - d_t)K_t - C_t)\Delta + K_t \sqrt{\Delta}u_t \quad (2)$$

$$C_{t+\Delta} = C_t + A_t K_t \Delta - (1 + d_t)K_t\Delta + A_t K_t \sqrt{\Delta}u_t \quad (3)$$

where u_t are independent normally distributed random variables with expected value zero and variance $s^2\Delta$.

Y and C initial values in (1), (2), (3) are available in statistical sources. To determine parameters that are not available directly in the statistics, we have developed a model calibration procedure, based on comparing the simulation results to the statistical data. Note that the step Δ is one of the parameters to be determined. On the one hand, it should be small enough to ensure the convergence of the calculation method, on the other – large enough to leave stochastic component meaningful. To build an agreement criterion, we analyzed the values of the most common metrics – Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)

$$MAE_Z = \frac{1}{N}\sum_{i=1}^{N}\left|Z_i - \hat{Z}_i\right|, \quad RMSE_Z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Z_i - \hat{Z}_i\right)^2}$$

for Y and C and their combinations. The resultant criterion was

$$RMSE_Y * RMSE_C \to min. \quad (4)$$

To find the set of parameters delivering the minimum to the criterion (4) we used the numerical method combining the Nelder-Mead algorithm with the random search algorithm proposed in [2].

When the values of all parameters are determined we can construct a system dynamics model (Fig. 1) that can serve as an analyzing and forecasting tool for macroeconomic indicators dynamics.

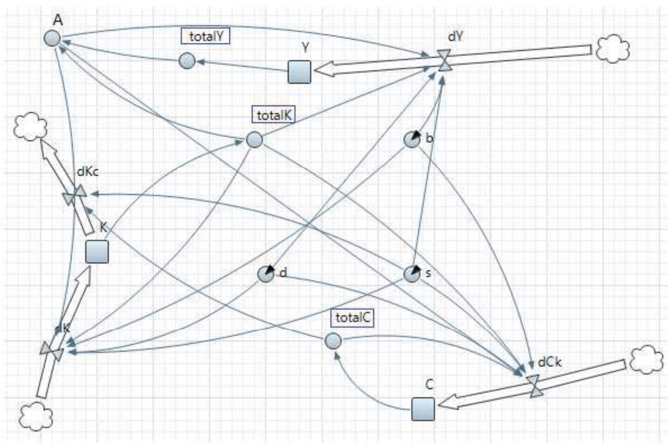

Fig. 1. AK-approach based system dynamics model diagram

As for software tools, system dynamics is supported by several tools that appear to be very much alike. AnyLogic [3] seems to gain some advantages such as model export, cloud model execution, and interoperability with other software tools.

REFERENCES

[1] A.V.Vorontsovskiy and L.F.Vyunenko, "Constructing of economic development trajectories by approximating conditions of stochastic models of economic growth", Vestnik SPbSU, ser. 5, pp.123-147, 2014. Available at http://vestnik.spbu.ru/html14/s05/s05v3/07.pdf (in Russian, accessed 05.05.2019).

[2] L.F.Vyunenko, V.V.Egorov, E.Yu.Morozova, "Algorithm for Determining the Optimal Form of Tightening Sprengel Construction", Surveys of applied and industrial mathematics, vol. 9, № 2, pp. 352-353, 2002 (in Russian).

[3] https://www.anylogic.ru (accessed 05.05.2019).

Examining Inter-Causal Relationships in Shared Cognitive Maps

Mahinda Mailagaha Kumbure*, Anssi Tarkiainen*, Pasi Luukka*, Jan Stoklasa* and Ari Jantunen*
*School of Business and Management
LUT University, Yliopistonkatu 34, Lappeenranta, Finland
Emails: mahinda.mailagaha.kumbure@lut.fi, anssi.tarkiainen@lut.fi, pasi.luukka@lut.fi, jan.stoklasa@lut.fi, ari.jantunen@lut.fi

Keywords—Inter-causal relationship, cognitive map, consistency, coverage

I. Introduction

The relationship between two states (or events) such that one is a cause and other is its effect, is defined as causality (or causation). When one sees a stone flying toward a windshield and breaking it, one also perceives causality between these two events, namely the stone hit and broke the windshield [1]. Precisely, an individual's perceived causality plays a vital role in his/her decision-making, by providing a basis to select the right decisions. In this research, we intend to assess the perceptions of causal relationships in managerial cognition.

Inter-causality is a relationship between one causation to another as a cause and effect, defined based upon a cognitive map. This is a type of behaviors of causes and effects among the elements in the cognitive maps apart from the other actions such as direct or indirect, positive or negative, and forward or backward causal relations as examples. In social science, a cognitive map is defined as a mental model representing relationships among the states (elements) of a particular situation [2]. Using graph theory, Bondy et al [3] originally defined the topology of a cognitive map that it contains nodes indicating the ideas or concepts and directional arcs which connect the nodes indicating the relationship between those nodes with the strength of the relationship. These typical relations are causal relationships in the cognitive map, as Robert Axelrod expressed (see chapter 1, [4]).

In an organizational setting, cognitive mapping has been applied to produce efficient systems which aid strategic planning in businesses and support to keep the stability through the changes in a particular environment. Especially, perceived causality of various factors represented by a cognitive map offers a whole image of inter-relationships between environmental factors and strategic goals which influence in strategic planning systems [5]. In general, today ambiguous business environment is increasing interest in managerial cognition research, especially in top managements perceptions on changes in their environment. Research in managerial cognition and organizational theory has long noticed the fundamental importance of managerial mental structures and their evaluations. Thus, there is still a need for more comprehensive approaches to analyze perceived causal relations to increase knowledge on how individual cognitive constructs scale into a collective level. Along with that, this research topic is to examine the role of perceived causalities on behalf of beliefs on a strategic issue interpretation among the boards of managers (graduate students represent the decision-makers/ managers) making sense of sustainable competitive advantage. To do that, we introduce a comprehensive approach based on the fuzzy set-theoretic qualitative comparative analysis. The method developed with this study serves to scrutinize and interpret the inter-causal relations in the cognitive maps and identify the main concepts which contribute to strong causal relationships which can influence the other causal relations.

II. Data and methods

1) Data: For our study, we used data of shared cognitive maps collected through an eight-week business simulation task on strategy implementation that was conducted by strategic management teams of graduate students. The used data contained 71 individual maps belonging to 16 groups, and the simulation task has been driven at the end in 2016. The causal maps have been produced based on 40 strategic issues and their effects on cumulative total return. Each member of the group was asked to choose 12 strategic constructs most relevant to the cumulative returns according to their knowledge and unique views while creating his/her cognitive map. To applied data into calculations, we first transformed the cognitive maps data into adjacency matrices (40 strategic issues $+$ cumulative total returns define each dimension of the matrix). Each cell value of these matrices represents the strength ($\in [-3, 3]$) of a causal relation between two strategic constructs.

2) Fuzzy set-theoretic consistency and coverage measures: Fuzzy set-theoretic qualitative comparative analysis (fsQCA) initially introduced by social scientist C. Ragin [6], provides powerful tools to find what combinations of characteristics of cases are necessary or sufficient to produce an outcome [7]. The fsQCA attempts to recognize logical connections between causal conditions and outcomes, and the results summarize all sufficient combinations of the causal conditions and the outcomes [7]. There are two important notions in fsQCA, consistency and coverage. Consistency indicates causal configurations percentage of similar structures that exhibit the same results. If the value of consistency is low for a configuration, then empirical evidence does not support the situation. Thus, a threshold for the consistency of at least 0.75 is reasonable to be used in investigations [8]. Coverage indicates the percentage of cases covered by valid configurations. In contrast to consistency, low coverage value of a configuration does not mean less relevance [9]. It is also considered that coverages ranged of 0.25 and 0.65 are more suitable for explanatory solutions [10]. In this work, consistency and coverage measures are the key theoretical aspects applied to explore and examine the complexity of inter-causal relations in cognitive maps. To compute the consistency and coverage measures we used the formulas in [11], [12].

3) Research process: The analysis started with collecting histogram information through the cognitive maps for each causal relation. As we know the range of strengths of causality is from -3 to 3, a frequency on each strength value for a particular causal relationship was counted (that is, for example, $[2, 5, 1, 47, 3, 4, 0]$ for $[-3, -2, -1, 0, 1, 2, 3]$ means that two individuals had -3, five individuals had -2 and so on). In this way, the causal relationships were identified whether they are positive (zero frequency of $-3, -2, -1$ and 0) or negative (zero frequency of $1, 2, 3$ and 0). This task was to examine the inter-causal relationships in a map, where "Inter-causal" refers to one causal relationship causing another one. Accordingly, we first established two assumptions (H_1 & H_2) to determine which causal relations have significant effects on other relationships in the cognitive maps.

H_1: If $C_i \to C_j$ is positive then $C_p \to C_q$ is positive

H_2: If $C_i \to C_j$is negative then $C_p \to C_q$ is negative

Where, C with a subscript indicates a particular element, and $C_i \to C_j$ and $C_p \to C_q$ implies two different causal relations for i, j, p, and q ($i \neq p$ and $j \neq q$) in a map.

To evaluate these hypotheses, the analysis was conducted in the framework of fuzzy sets. Unlike classical sets, fuzzy sets are sets whose elements have degrees of membership. Usually, a fuzzy set is characterized by a membership function [13], for example, trapezoidal membership function. In our analysis, we first produced membership values from the collected histogram information using the trapezoidal membership function. These memberships were based on linguistics levels positive and negative (to be tested hypotheses). From the definition of trapezoidal membership function in [14], each input vector was designed using expert knowledge as (0, 1, 3, 3) for positive memberships, and (-3, -3, -1, 0) for negative memberships, as a representation of the trapezoidal fuzzy number by its significant values. Once the all memberships were generated, consistency and coverage scores were computed. To simplify and evaluate the tested hypotheses, together with ideal consistency ($= 1$) and reasonable coverage values (between 0.25 and 0.65) were considered. Addition to that, we examined the cases which have ideal consistency and ideal coverage supporting the assumptions. Hence, we identified the most clear and strong inter-causal relationships which being in the cognitive maps.

III. Results and discussion

The results obtained with this study have identified the essential causal relations which have been affected by other causalities in the cognitive maps, emerging with H_1 and H_2 based on set-theoretic scores. There is clear evidence that some relations have frequently affected other relationships to determine whether it is positive or negative. Accordingly, it is obvious that which causal relations play a vital role in the causal framework. Besides these, we have found several ideal inter-causal relationships (consistency = 1, coverage = 1) which are significant under H_2. This finding corroborates with Ragin's definitions in [8], [11], who suggested that consistency $(A \Rightarrow B) = 1$ means $A \subseteq B$ and coverage $(B \Rightarrow A) = 1$ means $B \subseteq A$. For example, we first found that perfect consistency and coverage on $(C_1 \to C_2) \Rightarrow (C_3 \to C_4)$ and then also on $(C_3 \to C_4) \Rightarrow (C_5 \to C_6)$. This verification endorses the approach used in this investigation to evaluate causal relations. Notice that, C with each subscription represents an element in strategic points of view in the maps. To sum up, we were able to discover the specific patterns of the inter-causal relations and examine their roles in cognitive maps apart from the implication of the use of fuzzy set-theoretic approaches in the causality analysis.

IV. Conclusion

In this study, we explore the relationships from particular causal relations to other specific causal relations (inter-causal relationships) as causes and effects in the cognitive structures based upon the managerial cognition in a strategic point of views. To the best our knowledge, this is the first study that examines the inter-causal relations in organization cognitive maps. We also introduce a novel method for scrutinizing and expressing the inter-causal relationships in a meaningful way from causal maps using fuzzy set theoretic consistency and coverage scores measuring through the histogram information.

References

[1] U. Hiroyuki, "Causal Context Presented in Subsequent Event Modifies the Perceived Timing of Cause and Effect", *Frontiers in Psychology*, vol. 8, pp. 314, 2017.

[2] R. M. Kitchin, "Psychology", *Journal of Environmental Psychology*, pp 1-19, 1994.

[3] J. A. Bondy and U. S. R. Murty, "Graph theory with applications", 1976.

[4] R. Axelrod, "Structure of decisions: the cognitive maps of political elites", 1976.

[5] K. C. Lee, "A cognitive map knowledge-based strategic planning simulation", *Proc. of the 26th Hawaii Inter. Conf. on System Sciences, Wailea, HI, USA* vol. 3, pp. 249-267, 1993.

[6] M. Vancea and C. C. Ragin, "The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies", *Fuzzy-Set Social Science. Papers: revistade sociologia*, vol. 80, pp. 299-301, 2006.

[7] R. Kent, "Using fsQCA: A Brief Guide and Workshop for Fuzzy-Set Qualitative Comparative Analysis (Teaching Notes)", *University of Manchester*, 2008.

[8] C. C. Ragin, "Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA)", *Configurational Comparative Analysis*, pp. 87-121, 2008.

[9] H. Kane and M. A. Lewis and P. A. Williams and L. C. Kahwati, "Using qualitative comparative analysis to understand and quantify translation and implementation", *Translational Behavioral Medicine*, vol. 4, pp. 201-208, 2014.

[10] B. Rihoux and C. C. Ragin, "Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques", *Thousand Oaks, CA: Sage*, 2009.

[11] J. Stoklasa and P. Luukka and T. Talasek, "Set-theoretic methodology using fuzzy sets in rule extraction and validation - consistency and coverage revisited", *Information Sciences*, pp. 412-413:154173, 2017.

[12] J. Stoklasa and T. Talasek and P. Luukka, "On consistency and coverage measures in the fuzzified set-theoretic approach for social sciences: dealing with ambivalent evidence in the data", *in Proc. of the 36th Inter. Conf. on Mathematical Methods in Economics*, pp. 521526, 2018.

[13] L. A. Zadeh, "Fuzzy Sets", *Information and Control*, vol. 8, pp. 338-353, 1965.

[14] S. Princy and S. S. Dhenakaran, "Comparison of Triangular and Trapezoidal Fuzzy Membership Function", *Journal of Computer Science and Engineering*, vol. 2, 2016.

Analysis of a Digital Transformation Factors's Impact On Company Economic Performance

Gadasina Lyudmila V.
Saint-Petersburg State University
Saint-Petersburg, Russia
LUT University,
Lappeenranta, Finland
gadasina.l.v@gmail.com

Tashlykov Kirill A.
Saint-Petersburg State University,
student,
Saint-Petersburg, Russia
st050020@student.spbu.ru

***Abstract*— This paper proposes an approach for identifying the digital transformation factors that influence an economic performance of a company most effectively in the short term.**

***Keywords*— digital transformation; text mining, cluster analysis**

Over the past decade, worldwide interest in digital business transformation has grown significantly. In fact, the main goal of digital transformation is the modernization of a company's business processes and the transfer of business to a qualitatively new level with the help of new technologies [1-2]. The purpose of the study is to identify the digital transformation factors that influence an economic performance of a company most effectively in the short term. We used Gartner IT Glossary [3] to select the digital transformation factors. Following the selection, a dictionary was compiled, consisting of 199 terms.

94023 documents from LexisNexis database [6] and 6008 from EBSCOhost database [5] documents related to the 2016-2017 period were analyzed concerning applied technologies for 40 companies.

The preliminary results of the analysis allow us to draw the following conclusions:

As leaders in terms of the mentions number in 2016-2017 one can consider such terms as virtual reality (about 10,000 mentions), the Internet of things, cloud services and 5G technology. Artificial intelligence and machine learning are the last of top-10 list.

The absolute leader in the number of companies that mentioned the technology was the Internet of Things, which was mentioned by all 40 companies reviewed in the study. Next come 3D technologies, artificial intelligence, machine learning and virtual reality – 39 companies.

The most innovative companies in terms of the number of technologies applied are IBM (the undisputed leader – 73 technologies), SAP (69 technologies), Oracle, Microsoft and Cisco (67, 63 and 61, respectively). WorkDay, a company that makes ERP systems, Fiat Chrysler and ServiceNow with 20 or less technologies, became anti-leaders. The Russian company Sberbank was interested in 29 technologies according to the collected data.

Cluster analysis of companies using different technologies allowed forming 4 clusters. The greatest influence on the clusters identification had the presence of the following technologies in the companies: 4G, 5G, Big Data, Biometric, Blockchain, Bots, Data Center, Data Processing, SaaS, Smart City. The analysis showed that there is no direct correlation between membership in a cluster and the level of revenue / profit of the company. Thus, the use of advanced digital technologies is insufficient for the successful digital transformation.

REFERENCES

[1] Brennan, L.; Ferdows, K.; Godsell, J.; Golini, R.; Keegan, R.; Kinkel, S.; Srai, J.S.; Taylor, M. Manufacturing in the world: Where next? Int. J. Oper. Prod. Manag. 2015, 35, pp. 1253–1274.

[2] Savastano M., Amendola C., Bellini F. and D'Ascenzo F. Contextual Impacts on Industrial Processes Brought by the Digital Transformation of Manufacturing: A Systematic Review. Sustainability 2019, 11(3), 891.

[3] Gartner IT Glossary. https://www.gartner.com/it-glossary/

[4] EBSCOhost. https://www.ebsco.com/

[5] LexisNexis Academic. https://www.lexisnexis.com/hottopics/lnacademic/

Optimal Saving by Expected Utility Operators

Irina Georgescu
Department of Economic Informatics and Cybernetics
Bucharest University of Economics
Bucharest, Romania
Email: irina.georgescu@csie.ase.ro

Jani Kinnunen
Åbo Akademi
Turku, Finland
Email: jani.kinnunen@abo.fi

***Abstract*—This paper studies an optimal saving model in the possibilistic context defined by an expected utility operator. Among the results of the paper, we mention an approximation formula of the optimal saving level and a sufficient condition such that the presence of a possibilistic risk to lead to an increase in optimal saving.**

I. Introduction

The optimal saving is a topic studied by Drèze and Modigliani [6], Leland [14] and Sandmo [15]. They introduced the notion of precautionary saving to measure the effect of probabilistic risk on the optimal saving and they showed that if the third derivative of the consumer's utility function is positive then the precautionary saving is positive. In [13], Kimball connects the precautionary saving with a notion of "consumer prudence", expressed by the prudence index associated with a utility function. The expected utility operators introduced in [9] allow to define a general notion of possibilistic expected utility which comprises the already existing ones. This paper studies an optimal saving model in the possibilistic context defined by an expected utility operator. The model in the paper extends the one from [3] and, by particularization, other concrete models of optimal saving can be obtained. Among the results of the paper, we mention an approximation formula of the optimal saving level and a sufficient condition such that the presence of a possibilistic risk to lead to a increase in optimal saving.

II. Expected Utility Operators and D-Operators

The classical von Neumann-Morgenstern (=EU-theory) is the framework in which the topics of probabilistic risk have been developed. The notion of expected utility lies in the center of this EU-theory: if X is a random variable and $u : \mathbf{R} \to \mathbf{R}$ is a continuous utility function then the mean value $M(u(X))$ of the random variable $u(X)$ is the expected utility associated with X and u. To define a possibilistic expected utility we need three entities: a)a weighting function $f : [0,1] \to \mathbf{R}$ (f is a non-negative and increasing function that satisfies $\int_0^1 f(\gamma)d\gamma = 1$); b)a fuzzy number A with the level sets $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$; c)a utility function $u : \mathbf{R} \to \mathbf{R}$ of class C^2.

In monograph [9] there are defined two notions of possibilistic expected utilities: $E_1(f, u(A)) = \frac{1}{2}\int_0^1 [u(a_1(\gamma)) + u(a_2(\gamma))]f(\gamma)d\gamma$ (2.1) $E_2(f, u(A)) = \int_0^1 [\frac{1}{a_2(\gamma)-a_1(\gamma)} \int_{a_1(\gamma)}^{a_2(\gamma)} u(x)dx]f(\gamma)d\gamma$ (2.2)

Particularizing $u(x)$ in formulas (2.1) and (2.2) we find some indicators associated with the fuzzy number A.

We will recall from [9], [10], [11] the definition of expected utility operators, some of their properties and some examples. Let $\mathcal{F}$ be the set of fuzzy numbers and $\mathcal{C}(\mathbf{R})$ the set of continuous functions from $\mathbf{R}$ to $\mathbf{R}$. Let $\mathcal{U}$ be a subset of $\mathcal{C}(\mathbf{R})$ satisfying the following properties: (U_1) $\mathcal{U}$ contains the constant functions and first and second degree polynomial functions; (U_2) If $a, b \in \mathbf{R}$ and $g, h \in \mathcal{U}$ then $ag + bh \in \mathcal{U}$. For any $a \in \mathbf{R}$, we denote by $\bar{a} : \mathbf{R} \to \mathbf{R}$ the function $\bar{a}(x) = a$, for any $x \in \mathbf{R}$. We fix a family $\mathcal{U}$ with the properties (U_1), (U_2) and a weighting function $f : \mathbf{R} \to \mathbf{R}$.

Definition 2.1: [10] An *(f-weighted) expected utility operator* is a function $T : \mathcal{F} \times \mathcal{U} \to \mathbf{R}$ such that for any $a, b \in \mathbf{R}$, $g, h \in \mathcal{U}$ and $A \in \mathcal{F}$ the following conditions are fulfilled:

(a) $T(A, 1_\mathbf{R}) = E_f(A)$;(b) $T(A, \bar{a}) = a$;(c) $T(A, ag + bh) = aT(A, g) + bT(A, h)$;d) $g \leq h$ implies $T(A, g) \leq T(A, h)$.

Several times in the paper we will write $T(A, g(x))$ instead of $T(A, g)$. The T-variance $Var_T(A)$ associated with an expected utility operator T is defined by $Var_T(A) = T(A, (x - E_f(A))^2)$.

Example 2.2: Let $T_1 : \mathcal{F} \times \mathcal{C}(\mathbf{R}) \to \mathbf{R}$ defined by $T_1(A, g) = \frac{1}{2}\int_0^1 [g(a_1(\gamma)) + g(a_2(\gamma))]f(\gamma)d\gamma = E_1(f, g(A))$ for any $g \in \mathcal{C}(\mathbf{R})$ and for any fuzzy number A whose level sets are $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$.

Example 2.3: Let $T_2 : \mathcal{F} \times \mathcal{C}(\mathbf{R}) \to \mathbf{R}$ defined by $T_2(A, g) = \int_0^1 [\frac{1}{a_2(\gamma)-a_1(\gamma)} \int_{a_1(\gamma)}^{a_2(\gamma)} g(x)dx]f(\gamma)d\gamma = E_2(f, g(A))$ for any $g \in \mathcal{C}(\mathbf{R})$ and for any fuzzy number A whose level sets are $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$.

Let $g(x, \lambda) : \mathbf{R^2} \to \mathbf{R}$ be a real function in the arguments x and λ. With respect to such a function $g(x, \lambda)$ we consider the following properties: (i) $g(x, \lambda)$ is continuous with respect to x and partially derivable with respect to parameter λ; (ii) For any parameter $\lambda \in \mathbf{R}$, the function $\frac{\partial g(.,\lambda)}{\partial \lambda} : \mathbf{R} \to \mathbf{R}$ is continuous.

Definition 2.4: [11] An expected utility operator $T : \mathcal{F} \times \mathcal{C}(\mathbf{R}) \to \mathbf{R}$ is derivable with respect to parameter λ (shortly, T is a D-operator) if for any fuzzy number A and for any function $g(x, \lambda)$ with the properties (i) and (ii), the following axioms are verified: (D_1) The function $\lambda \longmapsto T(A, g(., \lambda))$ is derivable (with respect to λ); (D_2) $T(A, \frac{\partial g(.,\lambda)}{\partial \lambda}) = \frac{d}{d\lambda} T(A, g(., \lambda))$.

Proposition 2.5: [11] The expected utility operators from Examples 2.2 and 2.3 are D-operators.

III. THE T-MODEL OF OPTIMAL SAVING

In this section we will define a model of optimal saving inside EU-theory associated with an expected utility operator. Our construction has as a starting point the probabilistic model of precautionary saving from [8], p.95 and the possibilistic model of optimal saving from [3].To fix ideas, we will briefly present the probabilistic model from [8]. This is a two-period consumption model, defined by the following entities:a)two utility functions $u(y)$ and $v(y)$: $u(y)$ is the consumer's utility function in the first period 0, and $v(y)$ is the consumer's utility function in the second period 1; b)in period 0 there is a sure income y_0, and in period 1 there is an uncertain income described by a random variable $\tilde{y}$;c)s is the level of saving for period 0; to face the risk in period 1, the consumer transfers to this period a part s of y_0;d)r is the interest rate for saving. The utility functions u and v are of class C^2 and fulfill the conditions $u' > 0$, $v' > 0$, $u'' < 0$, $v'' < 0$. The consumer wishes to choose that value s maximizing the following total utility function:

$U(s) = u(y_0 - s) + M(v((1+r)s + \tilde{y}_0))$ (3.1)

The transformation of the above model in a possibilistic one is done by: a)the uncertain income from period 1 will no longer be a random variable $\tilde{y}$, but a fuzzy number A; b)to define the total utility of the model it is necessary to have a concept of "possibilistic expected utility". Using this thesis, we will define an optimal saving model in the context offered by the expected utility operators.We fix a weighting function $f : [0,1] \to \mathbf{R}$ and an expected utility operator T. The possibilistic optimal saving model associated with T has the following initial data: a)two utility functions u and v, verifying the conditions from the probabilistic model;b)a sure income in period 0 and an uncertain income in period 1, described by a fuzzy number A whose level sets are $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$;c)s is the level of saving for period 0;d)r is the interest rate for saving. The total utility function of the model will be: $V_T(s) = u(y_0 - s) + T(A, v((1+r)s + y))$ (3.2) One notices the way the definition of $V_T(s)$ is inspired from (3.1): instead of $M(v((1+r)s+\tilde{y}))$ appears the generalized possibilistic expected utility $T(A, v((1+r)s+y))$. In case of the model described above, the consumer will choose the level of s resulted from solving the following optimization problem: $\max\limits_s V_T(s)$ (3.3)

To study the existence and the properties of the solutions of (3.3), we will assume that T is a D-operator.

Proposition 3.1: If T is a D-operator then (i) $V_T'(s) = -u'(y_0 - s) + (1+r)T(A, v'((1+r)s + y))$;(ii) $V_T''(s) = u''(y_0 - s) + (1+r)^2 T(A, v''((1+r)s + y))$; (iii) V_T is a concave function.

By the previous proposition, we can consider the solution s_T^* of the optimization problem (3.3). Taking into account Proposition 3.1(i), the first-order condition associated with the optimization problem (3.3) will be written as: $u'(y_0 - s) = (1+r)T(A, v'((1+r)s+y))$ (3.4) We will assume now that the supplementary condition is fulfilled: v is of class C^3. Then the first-order condition (3.4) will be written as:

$u'(y_0 - s) \approx (1+r)[v'((1+r)s + E_f(A)) + \frac{1}{2}v'''((1+r)s + E_f(A))Var_T(A)]$ (3.5)

The solution s_T^* of the optimization problem (3.3) will be called the T-*optimal saving*, and the above model will be called the T-*model of optimal saving*.

Proposition 3.2: An approximate value of s_T^* is given by the following formula: $s_T^* \approx \frac{u'(y_0)-(1+r)[v'(a)+\frac{1}{2}v'''(a)Var_T(A)]}{u''(y_0)+(1+r)^2[v''(a)+\frac{1}{2}v^{iv}(a)Var_T(A)]}$ (3.6) where $a = E_f(A)$.

The approximation formula (3.6) can be written as: $s_T^* \approx \frac{u'(y_0)-(1+r)T(A,v')}{u''(y_0)+(1+r)^2T(A,v'')}$ (3.7) In the above model of T-optimal saving risk appears in the second period as the fuzzy number A. We will modify the model of T-optimal saving in the following way: in period 1, the income will be the possibilistic expected value $E_f(A)$. The other parameters of the first model will be kept for the second one. Then the total utility function of the second model will be: $W(s) = u(y_0 - s) + v((1+r)s + E_f(A))$ (3.8) and the optimization problem will be: $\max\limits_s W(s)$ (3.9) We can consider the solution s^* of the optimization problem (3.9). The first-order condition $W'(s^*) = 0$ will be written as: $u'(y_0 - s^*) = (1+r)v'((1+r)s^* + E_f(A))$ (3.10) The difference $s_T^* - s^*$ will be called T-*precautionary saving*. This indicator measures the variation of optimal saving when we pass from the model without risk (3.9) to the model with risk (3.3). In other words, the T-precautionary saving $s_T^* - s^*$ shows to which extent the optimal saving level increases or decreases in the presence of possibilistic risk A.

Proposition 3.3: Assume that v is of class C^3. If $v''' \geq 0$ then $s_T^* - s^* \geq 0$.

REFERENCES

[1] C. Carlsson, R. Fullér, On possibilistic mean value and variance of fuzzy numbers, Fuzzy Sets Syst., 122, 2001, 315–326

[2] C. Carlsson, R. Fullér, Possibility for decision, Springer, 2011

[3] A. M. Lucia-Casademunt, I. Georgescu, Optimal saving and prudence in a possibilistic framework. In: Omatu S., Neves J., Rodriguez J., Paz Santana J., Gonzalez S. (eds) Distributed Computing and Artificial Intelligence. Advances in Intelligent Systems and Computing, vol 217. Springer, Cham, 2013

[4] A. M. Lucia-Casademunt, I. Georgescu, Connecting possibilistic prudence and optimal saving, International Journal of Interactive Multimedia and Artificial Intelligence, 2(4), 2013, 38-45

[5] C. Courbage, B. Rey, Precautionary saving in the presence of other risks, Economic Theory, 32, 2007, 417-424

[6] J. Drèze, F. Modigliani, Consumption decision under uncertainty, J. Economic Theory, 5, 1972, 308-355

[7] D. Dubois, H. Prade, Possibility theory, Plenum Press, New York, 1988

[8] L. Eeckhoudt, C. Gollier, H. Schlesinger, Economic and Financial Decision under Risk, Princeton University Press, 2005

[9] I. Georgescu, Possibility theory and the risk, Springer, 2012

[10] I. Georgescu, Expected utility operators and possibilistic risk aversion, Soft Computing, 16, 2012, 1671-1680

[11] I. Georgescu, A portfolio choice problem in the framework of expected utility operators, in preparation

[12] I. Georgescu, J. Kinnunen, Precautionary saving with possibilistic background risk, CINTI 2015, 19–21 November 2015, Budapest, Hungary, 165-169

[13] M. S. Kimball, Precautionary saving in the small and in the large, Econometrica, 58, 1990, pp. 53–73

[14] H. Leland, Saving and uncertainty: the precautionary demand for saving, Quarterly Journal of Economics, 82, 1968, 465-473 []

[15] A. Sandmo, The effect of uncertainty on saving decisions, Rev. Economic Studies, 37, 1970, 353-360

Adaptive Market Hypothesis: A Comparative Analysis of Seven Major Digital Currencies

Sheraz Ahmed
School of Business and Management
LUT University
Lappeenranta, Finland
Sheraz.Ahmed@lut.fi

Muhammad Naeem
Faculty of Management Sciences
University of Central Punjab
Lahore, Pakistan

Abstract— **This study tests the adaptive market hypothesis (AMH). Two different methods, namely Generalized Spectral (GS) and Dominguez-Lobato (DL) test are used to analyze AMH in the case of seven main digital currencies; Ripple, Bitcoin, Ethereum, Monaro, Tether, Litecoin, and Steller over the sample period of 2013 — 2018. This study adopts an extremely important category of martingale difference hypothesis (MDH), which uses non-linear methods of dependencies for identifying changing linear and non-linear dependence in the price movement of currencies. The results indicate that price movements with linear and nonlinear dependences varies over time. We find that Bitcoin, Ethereum, Monaro and Litecoin have the longest efficiency periods. We also find that depending on market conditions, different market periods have significant impact on prices fluctuations of cryptocurrencies. Therefore, our findings suggest to use Adaptive market hypothesis (AMH) for testing efficiency of digital currency market as predicting changes in prices over time must take into account the time-varying market conditions for improving the efficiency of forecasting.**

Keywords*— *adaptive market hypothesis; digital currencies; market efficiency

I. Introduction

Innovations are significant contributors of economic advancement and they bring better solutions to market with the help of latest financial innovations (Kristoufek, 2013). Today, digital currencies propose an updated design of existing procedure of financial system. As it enables all internet users to exchange information, which is nearly cost free (Glaser, et al., 2014). Thus, digital currency transactions are gaining media attention due to its unique strands and rapid progression in recent years. Academic interests in anonymous electronic payments started from early 1980s, and Digital cash, the first digital currency was introduced in 1990, which offered reduction in transaction cost with anonymity by cryptographic procedures (Briére et al., 2015). This anonymous electronic monetary system was primarily introduced by Nakamoto (2008) who designed cryptocurrency based on block-chain technology, termed Bitcoin and launched the software that lead to the spectacular development of bitcoin network. Cryptocurrencies are decentralized form of digital currencies, which requires encryption to authenticate transactions (Buchholz et al., 2012).

For cryptocurrency markets, a large body of researchers analyzed the weak form of Efficient Market hypothesis (hereafter EMH) of Fama (1970), mostly in case of Bitcoin Market. However, over the efficiency of cryptocurrency markets, a controversy exists between opponents and proponents of the EMH. For instance, Bariviera (2017), Nadarajah and Chu (2017), and Tiwari et al., (2018) concluded that the Bitcoin market is almost competent as it reflects all known information. In contrast, Al Yahyaee et al., (2018), Cheah et al., (2018), Urquhart (2017), and Urquhart and McGroarty (2016) presented skeptical empirical findings that do not support the EMH for Bitcoin market. One of the main reason of this controversy is that competence of cryptocurrencies' prices varies over time. Grossman and Stiglitz (1980) contributed in literature by giving a hypothetically convincing argument that the presence of a competent market is impossible, because if the prices reflect all available information, then investors would not be concerned about costly acquisition of information. The Adaptive Market Hypothesis (hereafter AMH) was introduced by Lo (2004) to allow market efficiency and inefficiencies to exist in a logically consistent way. The market efficiency evolves over time in AMH in contrast to EMH. Lo (2004) analyzed the time-varying feature of efficiency and determined that AMH is more reliable method than EMH as it states that human errors can lead to arbitrage prospects to exist in specific time periods and these prospects vanish once they are exploited. Likewise, Kurihara and Fukushima (2017) examined the calendar anomalies in digital currencies and found strong evidences against EMH by documenting significant day-of-the-week impact. Albeit, they anticipate that by an increase in trading volume and transactions, the digital currency will be more informationally efficient. Several researchers have also analyzed the AMH on different asset classes and in stock market for example: stocks by Urquhart and McGroarty (2016) and Kim et al. (2011); currencies by Cialenco & Protopapadakis (2011) and Neely et al. (2009) and reported that AMH is an efficient predictor. Cryptocurrencies have become an area of interest for the media, investors and regulators. Practitioners and government authorities appreciate some of their features, such as independence from sovereign authorities and technological innovation. Interestingly, many cryptocurrencies offer outstanding price growth in relatively short periods, exceeding hundred percent levels (Kristoufek, 2013), but with extreme volatility (Chu et al., 2017; Urquhart, 2017; Katsiampa, 2017; Kristjanpoller and Minutolo, 2018). Importantly, there has been a remarkable rise in their market value and the number of transactions. For example, Bitcoin's market value exceeded 250$ billion at the beginning of 2018 and trading volume for the same year reached $2 trillion. Increases in prices and trading activities are seen in other leading cryptocurrencies such as Litecoin and Ethereum. From early 2017 to December 2017, the overall market value of all cryptocurrencies spiked from $18 billion to around $600 billion (Bouri et al., 2018).

In this study, first, we determine the efficiency of AMH for seven digital currencies, namely Ripple, Bitcoin, Ethereum, Monaro, Tether, Litecoin, and Steller. Second, we applied Generalized Spectral (GS) test of Escanciano and Velasco (2006) and Dominguez-Lobato (DL) test of Domínguez and Lobato (2003) to examine martingale difference hypothesis (MDH) and identify changing linear and non-linear dependencies in the selected cryptocurrencies' price movements.

REFERENCES

[1] Al-Yahyaee, K.H., Mensi, W. & Yoon, S.M. (2018). Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets. Finance Research Letters 27, 228-234.

[2] Bariviera, A.F. (2017). The inefficiency of Bitcoin revisited: A dynamic approach. Economics Letters 161, 1-4.

[3] Briére, M., Oosterlinck, K. & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. Journal of Asset Management 16(6), 365-373.

[4] Bouri, E., Lau, C.K.M., Lucey, B. & Roubaud, D. (2019). Trading volume and the predictability of return and volatility in the cryptocurrency market. Finance Research Letters 29, 340-346.

[5] Buchholz, M., Delaney, J., Warren, J. & Parker, J. (2012). Bits and bets, information, price volatility, and demand for Bitcoin. Economics 312, 2-48.

[6] Cialenco, I. & Protopapadakis, A. (2011). Do technical trading profits remain in the foreign exchange market? Evidence from 14 currencies. Journal of International Financial Markets, Institutions and Money 21(2), 176-206.

[7] Cheah, E.T., Mishra, T., Parhi, M. & Zhang, Z. (2018). Long memory interdependency and inefficiency in Bitcoin markets. Economics Letters 167, 18-25.

[8] Chu, J., Chan, S., Nadarajah, S. & Osterrieder, J. (2017). GARCH modelling of cryptocurrencies. Journal of Risk and Financial Management 10(4), 17.

[9] Domínguez, M.A. & Lobato, I.N. (2003). Testing the martingale difference hypothesis. Econometric Reviews 22(4), 351-377.

[10] Escanciano, J.C. & Velasco, C. (2006). Generalized spectral tests for the martingale difference hypothesis. Journal of Econometrics 134, 151–185.

[11] Fama, E. F. (1970). Efficient market hypothesis: A review of theory and empirical work. The Journal of Finance 25(2), 28-30.

[12] Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C. & Siering, M. (2014). Bitcoin – asset or currency? Revealing users' hidden intentions. ECIS 2014 (Tel Aviv). https://ssrn.com/abstract=2425247

[13] Grossman, S.J. & Stiglitz, J.E. (1980). On the impossibility of informationally efficient markets. The American economic review 70(3), 393-408.

[14] Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. Economics Letters 158, 3-6.

[15] Kim, J.H., Shamsuddin, A. & Lim, K.P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data. Journal of Empirical Finance 18(5), 868-879.

[16] Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. Scientific reports 3, 3415.

[17] Kristjanpoller, W. & Minutolo, M.C. (2018). A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis. Expert Systems with Applications 109, 1-11.

[18] Kurihara, Y. & Fukushima, A. (2017). The market efficiency of Bitcoin: a weekly anomaly perspective. Journal of Applied Finance and Banking 7(3), 57.

[19] Lo, A. W. (2004). The adaptive markets hypothesis. The Journal of Portfolio Management 30(5), 15-29.

[20] Nadarajah, S., & Chu, J. (2017). On the inefficiency of Bitcoin. Economics Letters 150, 6-9.

[21] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. http://bitcoin.org/bitcoin.pdf.

[22] Neely, C.J., Weller, P.A. & Ulrich, J.M. (2009). The adaptive markets hypothesis: evidence from the foreign exchange market. Journal of Financial and Quantitative Analysis 44(2), 467-488.

[23] Tiwari, A. K., Jana, R. K., Das, D. & Roubaud, D. (2018). Informational efficiency of Bitcoin—An extension. Economics Letters 163, 106-109.

[24] Urquhart, A. (2017). Price clustering in Bitcoin. Economics letters 159, 145-148.

[25] Urquhart, A. & McGroarty, F. (2016). Are stock markets really efficient? Evidence of the adaptive market hypothesis. International Review of Financial Analysis 47, 39–49.

Enhancing pairs trading strategies with machine learning – a Nordic perspective

Work in progress

Sheraz Ahmed and Mikael Collan
School of Business and Managemen, LUT University
Yliopistonkatu 34, 53851 Lappeenranta, Finland
sheraz.ahmed; mikael.collan; @lut.fi

Andreas Mikkelsen
School of Business and Economics in Alta
The Arctic University of Norway
PO Box 1621, 9509 Alta, Norway
andreas.mikkelsen@uit.no

***Abstract*—This work in progress concentrates on pairs trading strategies and specifically on enhancing the performance of selected pairs trading strategies by using machine learning. The main idea put to action is the optimization of selected parameters of pairs trading strategies by fitting them to data in various ways. The parameters include, for example, the upper and lower trigger distances (or spreads) from the equilibrium, the length of time windows used in the formation period, and the length of time windows used in trading period according to the fitted strategy. Also strategies with a rolling time window used for constant formation (parameter refitting) are studied. The work will also discuss optimizing triggers for stop-loss limits and the possible consequences of transaction costs and thin trading on the performance of pairs-trading strategy.**

The focus is on Nordic markets and pairs from the Norwegian and the Finnish stocks are used as test-examples. Data from selected pairs of stocks from Norwegian companies active in the fishery industry and from Finnish companies with two series (dual class) of stocks is used. The results from the fitted strategies are compared against simple benchmark strategies from previous literature under both growth and recession periods, and periods with high volatility.

Keywords—pairs trading; parameter fitting; machine learning; Nordic countries; algorithmic trading

LUT Scientific and Expertise Publications
Research Reports 97 / August 2019
ISSN 2243-3376

ISBN 978-952-3353-400-5 (PDF)