MULTIPLE CRITERIA DECISION MAKING

ANNUAL

VOL. 14 (2019)

UNIVERSITY OF ECONOMICS IN KATOWICE 2019

Printed by

EXDRUK Wojciech Żuchowski, ul. Rysia 6, 87-800 Włocławek
e-mail: biuroexdruk@gmail.com, www.exdruk.com

ISSN 2084-1531

Edition: 50 copies

Original version of the MCDM is the paper version

Publishing House of the University of Economics in Katowice
ul. 1 Maja 50, 40-287 Katowice, tel. +48 32 25 77 635, fax +48 32 25 77 643
www.wydawnictwo.ue.katowice.pl, e-mail: wydawnictwo@ue.katowice.pl

Contents

MULTIPLE CRITERIA DECISION MAKING
Vol. 14 2019

Helena Gaspars-Wieloch[*]

SPARE PARTS QUANTITY PROBLEM UNDER UNCERTAINTY – THE CASE OF ENTIRELY NEW DEVICES WITH SHORT LIFE CYCLE

DOI: 10.22367/mcdm.2019.14.01

Abstract

The paper presents a new scenario-based decision rule for the spare parts quantity problem (SPQP) under uncertainty with unknown objective probabilities. The goal of SPQP is to ensure the right number of extra parts at the right place at the right time. In the literature, SPQP is usually regarded as a stochastic problem since the demand for extra parts is treated as a random variable with a known distribution. The optimal stock quantity minimizes the expected loss resulting from buying a given number of parts before potential failures.

The novel approach is designed for the purchase of non-repairable spare parts for entirely new seasonal devices, where the estimation of frequencies is complicated because there are no historical data about previous failures. Additionally, the decision maker's knowledge is limited due to the nature of the problem.

The new procedure is a three-criteria method. It is based on the Hurwicz and Bayes decision rules and supported with a forecasting stage enabling one to set the scenario with the greatest subjective chance of occurrence. The method takes into account the decision maker's attitude towards risk and the asymmetry of losses connected with particular stock quantities. We assume that the future unit purchase cost of a service part bought after the breakdown is also uncertain and given as an interval parameter. The approach is designed for short life cycle machines.

Keywords: spare parts quantity problem, new seasonal devices, uncertainty, interval payoffs, unknown objective probabilities, decision maker's preferences, short life cycle.

[*] Poznań University of Economics and Business, Department of Operations Research, Poznań, Poland, e-mail: Helena.gaspars@ue.poznan.pl, ORCID: 0000-0003-0033-3836.

1 Introduction

One of the main goals of the spare parts quantity problem (SPQP) is to ensure that right spare parts and resources are at the right place (where the broken part is) at the right time. Spare parts are kept in an inventory and should be in proximity to a functional item (engine, device, automobile, boat, machine) since they might be used to repair it or to replace failed units. They constitute an important element of logistic engineering and supply chain management. There are many synonyms for "spare parts" such as service parts, extra parts, repair parts, replacement parts and interchangeable parts.

SPQP may be analyzed in the context of repairable and non-repairable spare parts (Guide and Srivastava, 1997; Louit, Banjevic and Jardine, 2005). In this paper we focus on the latter.

In the literature spare parts optimization is usually regarded as a stochastic problem (Aronis et al., 2004; Gu and Li, 2015; Inderfurth and Mukherjee, 2008; Ravindran, 2007; Rodriguez et al., 2013; Sikora, 2008; Wong et al., 1997) since the demand for service parts is treated as a random variable with a known distribution. Here, however, we would like to consider SPQP as a strategic problem, i.e. an uncertain problem with unknown objective probabilities (frequencies). Such a situation may take place when a totally new device is bought and there are no historical data concerning previous breakdowns that could be used to estimate probabilities.

SPQP may be investigated as a single-period (SPP) or a multi-period problem (MPP) (Petrovic et al., 1988). In this contribution we analyze the first situation only since the multi-period horizon cannot be discussed in the context of SPQP under uncertainty with permanently unknown probabilities: in subsequent periods historical data (frequencies) become available and the objective likelihood can be estimated. Thus, in the case of totally new devices MPP would be a mixed problem. The first stage could be analyzed without known probabilities, but further stages could be based on probabilities. The second reason why we do not deal with MPP is that we assume that the purchase of additional spare parts for a given device is made only once for the whole period of usage since the life cycle of the considered device is relatively short due to its seasonal character and a constantly changing environment. It is worth emphasizing that when the chosen decision is supposed to be executed only once (one-shot decision), researchers advise the decision makers (DMs) against the use of probabilities, because only one event has the chance to occur (von Mises, 1949).

The research is related to totally new devices. Therefore, in contrast to the traditionally understood SPQP, we take into account not one but two types of uncertainty. The first one results from the unknown demand for extra parts given as a discrete random variable with an unknown probability distribution (discrete parameter). The second one is caused by the unknown future unit purchase cost of service parts (interval parameter).

The paper is organized as follows. Section 2 briefly presents the main features of the classical SPQP. Section 3 defines a new problem: SPQP for totally new devices, i.e. SPQP under uncertainty with unknown probabilities. Section 4 discusses the characteristics of the loss matrix connected with SPQP for different cases and analyzes the possible usefulness of classical decision rules in that field. Section 5 presents the assumptions of the scenario-based model and a 3-criteria decision rule that may be used for the aforementioned problem. The procedure takes into account DMs' preferences. Section 6 provides an illustrative example. Conclusions are gathered in the last section.

2 The classical spare parts quantity problem: description

In the original version of SPQP the goal is to find the optimal number of extra parts bought with the purchase of the whole device. "Optimal" means "minimizing the expected loss resulting from buying a given number of service parts before potential failures (breakdowns)". If we buy too many parts with the whole machine, we lose the money spent for the purchase of those parts. On the other hand, if we buy not enough spare parts with the whole item, we lose the difference between the current price of a spare part and the previous price of that part. SPQP is mainly related to DMSU − decision making under stochastic uncertainty −, and based upon the assumption of risk neutrality due to the fact that the demand (D) for extra parts is a random variable with a known probability distribution (Sikora, 2008).

The cumulative distribution function (F) of the demand may be continuous or discrete. In this paper we concentrate on the second variant. Within SPQP we can distinguish c_1, denoting the unit purchase cost of the subassembly together with the purchase of the whole device, and c_2, denoting the unit purchase cost of the subassembly just after the failure, where $c_1 < c_2$. Both costs allow us to compute two types of losses: $s_1 = c_1$ and $s_2 = c_2 - c_1$, where s_1 denotes the unit loss from buying a service part with the whole device (loss due to the excess of spare parts) and s_2 is the unit loss from buying an extra part just after the failure (loss due to the shortage of spare parts). The only decision variable in SPQP is q – the order quantity (number of spare parts bought with the device). Usually, the

DM considers possible discrete values of q from the interval $[D_{min}, D_{max}]$, where D_{min}, D_{max} are the lowest and the highest observed demand for spare parts, respectively.

The optimization model enabling one to find the optimal order quantity can be presented in the following way:

$$q^* = \arg\min_q l(q) \tag{1}$$

$$l(q) = \sum_{D=D_{min}}^{D_{max}} l(q,D) \cdot P(D) \tag{2}$$

$$l(q,D) = \begin{cases} s_1(q-D), & if \quad q > D \\ 0, & if \quad q = D \\ s_2(D-q), & if \quad q < D \end{cases}, \tag{3}$$

where q^* is the optimal order quantity, $l(q)$ denotes the expected loss, $l(q,D)$ is the loss incurred when the number of spare parts bought with the device equals q and the demand for spare parts is equal to D. $P(D)$ denotes the probability that the demand will be equal to D. Zero loss occurs if the order quantity is exactly the same as the demand.

There are many possible optimization methods to solve the aforementioned problem, such as the use of optimization software (SAS/OR, Solver in Excel, minizinc, R, cplex, etc.) and formulas (the recurrence equation, the critical ratio or the loss matrix; Sikora, 2008). In this paper we concentrate on loss matrices. Table 1 presents losses $l(q,D)$ for all possible combinations of pairs (q,D), see equation (3). Expected losses are generated in the last column (equation 2). The optimal solution is indicated by the lowest expected loss.

Table 1: Loss matrix for the classical version of SPQP (general case)

$q \setminus D$	$D_{min}=q_{min}$	$D_{min}+1$	…	$D_{max}-1$	$D_{max}=q_{max}$	$l(q)$
q_{min}	0	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	$l(q_{min})$
$q_{min}+1$	$s_1 \cdot (q-D)$	0	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	$l(q_{min}+1)$
…	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	0	$s_2 \cdot (D-q)$	$s_2 \cdot (D-q)$	…
$q_{max}-1$	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	0	$s_2 \cdot (D-q)$	$l(q_{max}-1)$
q_{max}	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	$s_1 \cdot (q-D)$	0	$l(q_{max})$

Source: Prepared by the author.

Interesting overviews of SPQP can be found, for instance, in Kennedy et al., 2002; Qu and Zhang, 2006; Rego and Mesquita, 2011. Extended SPQ models are variations of the classical SPQ model, involving additional losses due to the broken device, the purchase of parts at different moments, etc. (Bartakke, 1981;

Bian et al., 2013; Fera et al., 2010; Fortuin, 1981; Gu and Li, 2015; Papathanassiu and Tsouros, 1986; Pastore et al., 2015; Petrovic et al., 1986; Rodriguez et al., 2013; Rustenburg et al., 2000; Schuh et al., 2015; Sheikh et al., 2000; Verrijdt et al., 1998).

3 Spare parts quantity problem and entirely new seasonal devices

As it has been already mentioned, SPQP is usually regarded in the literature as a stochastic problem. However, in some circumstances it is extremely difficult to estimate the probability distribution (Gaspars-Wieloch, 2016a, 2017a, 2017b, 2018b, 2019a, 2019b).

Here we would like to investigate the case when totally new seasonal devices are bought. This entails: (1) the lack of historical data about previous failures, (2) the lack of sufficient knowledge about the mechanism of particular machines, (3) the inability to precisely define the whole sample space (Kolmogorov, 1993) and (4) perhaps a feeling anticipating new future factors which can radically change the trend up to now. Under such conditions objective probability quantification is impossible.

We focus on machines with very short life cycle. In such a case, the purchase of additional spare parts for these devices is made only once for the whole period of use (until the machine is withdrawn from service). Under such assumptions SPQP can be reduced to a one-shot decision problem (Guo, 2011; Zhu and Guo, 2016), since for each device only one scenario can occur. Czerwiński (1960) and von Mises (1949) state that the mathematical probability (understood as frequency) and expected value cannot be used for a single event, but only for repetitive events. Hence, this is the second reason why the use of probability in SPQP is not always justified.

There is also a general drawback related to the application of likelihood (not necessarily connected with SPQP and new devices). The term “probability” has many discrepant definitions, e.g. objective, subjective, classical, geometrical, frequency, logic, Bayes, Kolmogorov, Springer, Piegat, propensity (Carnap, 1950; de Finetti, 1975; Frechet, 1938; Hau et al., 2009; Knight, 1921; Kolmogorov, 1933; Piegat, 2010; Popper, 1959; Ramsey, 1931; Van Lambalgen, 1996; von Mises, 1949, 1957). The lack of unanimity leads to numerous doubts: what approach should be used? How to estimate the type of probability selected? Caplan (2001) adds that people are even unable to declare subjective probabilities: “they implicitly set them in acting”.

In connection with all those facts, we would like to investigate SPQP for totally new seasonal devices as an SPQP under complete uncertainty, i.e. uncertainty with unknown probabilities (UUP).

It is worth mentioning that the term “uncertainty” also has diverse interpretations and types (Gaspars-Wieloch, 2016a, 2018b). According to decision theory, uncertainty may only be associated with situations where probabilities are unknown (in other situations this theory refers to risk or partial uncertainty). According to the theory of economics, there are diverse degrees of uncertainty but all of them involve situations with non-deterministic parameters (with risk understood as the possibility that some unfavorable or unpredicted event will happen). This paper is based on the second aforementioned theory. We consider both epistemic and aleatory uncertainty (Stirling, 2003; Zio and Pedroni, 2013).

Note that SPQP can be easily combined with scenario planning (Pomerol, 2001) thanks to (1) well-defined discrete sets of decisions (order quantities) and states of nature (demand quantities) and (2) the possibility to compute the loss matrix precisely (see Table 2).

The result of a choice made under uncertainty with scenario planning depends on two factors: which decision will be selected and which scenario will occur. Thus, SPQP under complete uncertainty may be defined by means of a scenario-based decision model with m states of nature (scenarios, events, demand quantities): $S = \{S_1, \ldots, S_i, \ldots, S_m\}$, n possible alternatives (decisions, strategies, order quantities): $A = \{A_1, \ldots, A_j, \ldots, A_n\}$, and $n \times m$ losses ($a_{i,j}$ – loss incurred by the buyer if state S_i occurs and alternative A_j is selected) calculated according to formula (4). The distributions of losses are discrete. The interpretation of S_i is that until the end of the use of a given machine D_i spare parts will be needed.

$$a_{i,j} = \begin{cases} s_1(q-D), & if \quad j > i \\ 0, & if \quad j = i \\ s_2(D-q), & if \quad j < i \end{cases} \tag{4}$$

Table 2: Loss matrix for the SPQP presented as a scenario-based model (general case)

Scen. \ Altern.	**A_1 (q=q_{min})**	**A_2 (q=q_{min}+1)**	…	**A_j**	…	**A_{n-1} (q=q_{max}-1)**	**A_n (q=q_{max})**
S_1 (D=D_{min})	$a_{1,1}$	$a_{1,2}$	…	$a_{1,j}$	…	$a_{1,n-1}$	$a_{1,n}$
S_2 (D=D_{min}+1)	$a_{2,1}$	$a_{2,2}$	…	$a_{2,j}$	…	$a_{2,n-1}$	$a_{2,n}$
⋮	⋮	⋮	⋱	⋮	⋱	⋮	
S_i	$a_{i,1}$	$a_{i,2}$	…	$a_{i,j}$	…	$a_{i,n-1}$	$a_{i,n}$
⋮	⋮	⋮	⋱	⋮	⋱	⋮	
S_{m-1} (D=D_{max}-1)	$a_{m-1,1}$	$a_{m-1,2}$	…	$a_{m-1,j}$	…	$a_{m-1,n-1}$	$a_{m-1,n}$
S_m (D=D_{max}1)	$a_{m,1}$	$a_{m,2}$	…	$a_{m,j}$	…	$a_{m,n-1}$	$a_{m,n}$

Source: Prepared by the author.

Researchers discuss the pros and cons of using probability data in scenario planning in various papers (Gaspars-Wieloch, 2019b) and SPQP usually refers to the probability calculus. However, in this contribution, we do not assign a likelihood to the demand, since the novelty degree of the decisions considered is very high.

The original version of SPQP is based upon the assumption of risk neutrality. In this work, however, we would like to take into account various preferences of the DMs (predictions, attitudes towards future results) which can be measured by the coefficients of optimism (β) and pessimism (α): $\alpha, \beta \in [0,1]$ and $\alpha + \beta = 1$ (α is close to 1 for extreme pessimists – risk averse behaviour, β is close to 1 for radical optimists – risk prone behaviour). Thanks to these parameters, we can adjust the final decision to the DM's nature. Additionally, the estimation of the coefficients is relatively little time-consuming (less time-consuming than the estimation of scenario probability).

4 Classical decision rules and specificity of SPQP loss matrices

In this section we analyze the specificity of SPQP loss matrices and investigate the usefulness of classical decision rules applied to scenario planning and decision making under complete uncertainty (i.e. max-max rule, Wald rule, Hurwicz rule, Bayes rule, Savage rule and max-min joy criterion). This research may be helpful in constructing a suitable procedure for SPQP with totally new devices.

Table 3 presents losses for three possible situations: 1. s_1–> s_2, 2. $s_1 = s_2$, 3. $s_1 < s_2$) which can occur in real-life situations. Values for prices c_1 and c_2 are ficticious, but in each case the first one is lower than the second one. The first situation is the least dangerous for the DM since the difference between prices is the lowest. The highest risk is connected with the last situation where a spare part bought in the future is much more expensive than a current extra part.

Table 3: Loss matrices for SPQP ($q_{min} = D_{min} = 0$, $q_{max} = D_{max} = 4$) and rankings generated on the basis of classical rules – examples 1-3

Ex.	***1. c_1=50, c_2=51, s_1=50, s_2=1***					***2. c_1=5, c_2=10, s_1=5, s_2=5***					***3. c_1=1, c_2=51, s_1=1, s_2=50***				
S \ A	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
S_1	0	50	100	150	200	0	5	10	15	20	0	1	2	3	4
S_2	1	0	50	100	150	5	0	5	10	15	50	0	1	2	3
S_3	2	1	0	50	100	10	5	0	5	10	100	50	0	1	2
S_4	3	2	1	0	50	15	10	5	0	5	150	100	50	0	1
S_5	4	3	2	1	0	20	15	10	5	0	200	150	100	50	0

Table 3 cont.

Ex.	1. c_1=50, c_2=51, s_1=50, s_2=1					2. c_1=5, c_2=10, s_1=5, s_2=5					3. c_1=1, c_2=51, s_1=1, s_2=50				
S \ A	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W	4	50	100	150	200	20	15	10	15	20	200	150	100	50	4
H *α=0.2*	0.8	10	20	30	40	4	3	2	3	4	40	30	20	10	0.8
H *α=0.8*	3.2	40	80	120	160	16	12	8	12	16	160	120	80	40	3.2
B	2	11.2	30.6	60.2	100	10	7	6	7	10	100	60.2	30.6	11.2	2
S	4	50	100	150	200	20	15	10	15	20	200	150	100	50	4

Source: Prepared by the author.

Conclusions regarding the specific structure of SPQP loss matrices are as follows:

a) for s_1 sufficiently larger than s_2 the average of losses is the smallest for $q = q_{min}$ and the range between $a_{j,min}$ (the smallest loss related to alternative A_j) and $a_{j,max}$ (the largest loss related to A_j) is an increasing function $f(q)$;
b) for s_1 close to s_2 the average of losses is the smallest for the middle q;
c) for s_1 sufficiently lower than s_2 the average of losses is the smallest for $q = q_{max}$ and the range between $a_{j,min}$ and $a_{j,max}$ is a decreasing function $f(q)$;
d) loss distributions connected with particular orders are usually asymmetric;
e) loss distributions are always symmetric for extreme alternatives (i.e. for the smallest and the largest numbers of spare parts);
f) for each decision particular losses $a_{1,j}$, …, $a_{i,j}$, …, $a_{m,j}$ are always ordered in the form of a convex function and the minimum loss is equal to zero.
g) ranges vary significantly for cases where s_1 significantly differs from s_2;
h) each decision is Pareto-optimal!

Hence, we see that in SPQP distributions of losses are usually asymmetric and loss ranges for particular order quantities can be extremely diverse.

Now, let us check whether classical decision rules may be applied to SPQP. The **max-max rule** is designed for radical optimists only: it does not satisfy the assumption from the previous section since it is unable to adjust the decision to the DM's nature. Note that in the case of SPQP, the max-max rule has to be transformed prior to its use to a min-min rule because the matrix contains losses expressed as positive numbers. And then we can notice that, due to the very specific structure of the loss matrix it is impossible to generate a ranking on the basis of that procedure since all decisions are always treated as the best ones, regardless of the problem analyzed (the best value for each decision is equal to 0), see Table 3 (row *M*)!

The **Wald** (Wald, 1950) **decision rule** (max-min rule for profits and min-max rule for losses expressed as positive numbers) is designed for radical pessimists only, so again, this approach does not allow to consider diverse types of decision makers, either. In the case of SPQP, this method always suggests the decision with the smallest range of losses and focuses on extreme states, i.e. scenarios for which the demand is equal to D_{min} or D_{max} (other events are not significant), see Table 3 (row W). Those states are connected with the largest loss.

The next well-known decision rule is the **Hurwicz criterion** (Hurwicz, 1952). Here, the DM declares his/her coefficient of optimism/pessimism and two extreme scenarios are always taken into account: one with the highest loss and one with the lowest loss. In the case of SPQP the event with the highest loss is related to D_{min} or D_{max}. The event with the lowest loss is different for each decision and occurs when the order quantity is equal to the demand. The idea of the Hurwicz rule consists in (1) calculating for each decision the sum of two products: coefficient of optimism (β) multiplied by the highest profit (the lowest loss) and coefficient of pessimism (α) multiplied by the lowest profit (the highest loss), and (2) selecting the decision with the highest profit weighted average or the lowest loss weighted average. Theoretically, the Hurwicz rule may be applied by different decision makers (optimists, pessimists, moderate DMs). Nevertheless, the structure of the SPQP loss matrix is so unusual that the maximal profit (i.e. the minimal loss) is always equal to zero. Therefore, regardless of the level of α and β (with one exception: $\alpha = 0$), decisions recommended by the Hurwicz rule are exactly the same as alternatives suggested by… the Wald rule, see Table 3 (rows H $\alpha = 0.2$ and H $\alpha = 0.8$). Hence, as a matter of fact, there is no possibility to take into consideration different types of decision makers, although each strategy is Pareto-optimal! For instance, according to the Hurwicz rule, alternative A_1 is better than A_2 in Example 1 even for $\alpha = 0.2$, which is quite astonishing as A_2 dominates A_1 in the case of four out of five states! Even when the coefficient of pessimism decreases, the Hurwicz rule applied to SPQP indicates variants suitable for pessimists.

Additionally, we can easily notice that when computing weighted indices for each decision, the status of particular scenarios varies depending on the alternative (see, for instance, example 1, Table 3: S_1 is the best scenario for A_1, but it is the worst state for A_5), which may be quite surprising in SPQP, where we rather tend towards the view that the most optimistic (pessimistic) scenario is that with the lowest (highest) demand for extra parts, regardless of losses connected with particular decisions. Perhaps, a global status for each state would be more appropriate than a local one.

A general remark concerning the Hurwicz rule: the procedure does not take into account the nature of the outcome distribution connected with particular alternatives, which leads to illogical recommendations for decision problems with asymmetric profits (or losses) (Gaspars-Wieloch 2014a, 2014b, 2016a, 2017b). This drawback is worth considering since in SPQP losses are usually asymmetric.

As opposed to previous approaches, the **Bayes (Laplace) criterion**, thanks to the use of the arithmetical average, analyzes both extreme and intermediate losses (not only extreme ones), which is significant in the case of asymmetric outcomes (Table 3, row *B*). However, the Bayes rule is not suitable for SPQP under complete uncertainty, since it does not allow to declare our coefficients of pessimism and it is designed for multi-shot decisions (hence for a multi-period horizon) only, while in this paper we assume that the purchase of additional spare parts at cost c_1 for a given device is made once for the whole period of use (until the machine is withdrawn from service).

There are also two other classical decision rules for which the position of a given outcome in the profit (loss) matrix is extremely important. This is a feature characteristic of the Savage rule (Savage, 1961) and the max-min joy criterion (MJC) (Hayashi, 2008). The goal of MJC is to show the superiority of particular outcomes connected with a given scenario to its worst result, while in the Savage rule the aim is to demonstrate the inferiority of particular payoffs related to a state of nature to its best result (Gaspars-Wieloch, 2014c, 2018a).

The **Savage rule** (min-max rule) requires the DM to generate a relative loss matrix (regret matrix), but due to the occurrence of zero losses for each scenario in SPQP, the original loss matrix may be treated as a relative loss matrix (without any transformation). Hence, in the case of SPQP, rankings obtained by means of the Savage approach correspond to rankings offered by the Wald rule (Table 3, row *S*) and, again, that method is appropriate for pessimists only: there is no possibility to adjust recommendations to the DM's nature.

The idea of the max-min joy criterion (MJC) is very similar to the reasoning characteristic of the Savage procedure, but instead of a regret table a relative profits matrix is applied and the solution is set on the basis of the worst relative profits connected with particular alternatives. MJC is designed only for people exhibiting a risk-averse behavior. Note that in the case of SPQP the worst relative profits are always related to the first (q_{min}) or the last decision (q_{max}), which means that even for significant differences between s_1 and s_2 those extreme alternatives (q_{min} and q_{max}) will never be optimal in accordance with MJC (they have at least one zero value in their column): Table 4, MJC. This conclusion seems illogical since in some real-life situations the choice of extreme order quantities is desirable.

Table 4: Relative profit matrices and rankings generated by MJC – examples 1-3

Ex.	***1. c_1=50, c_2=51, s_1=50, s_2=1***					***2. c_1=5, c_2=10, s_1=5, s_2=5***					***3. c_1=1, c_2=51, s_1=1, s_2=50***				
S \ A	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
S_1	200	150	100	50	0	20	15	10	5	0	4	3	2	1	0
S_2	149	150	100	50	0	10	15	10	5	0	0	50	49	48	47
S_3	98	99	100	50	0	0	5	10	5	0	0	50	100	99	98
S_4	47	48	49	50	0	0	5	10	15	10	0	50	100	150	149
S_5	0	1	2	3	4	0	5	10	15	20	0	50	100	150	200
MJC	**0**	**1**	**2**	**3**	**0**	0	5	**10**	5	0	0	**3**	2	1	0

Source: Prepared by the author.

Due to all these factors, we can conclude that the aforementioned decision rules should not be applied to SPQP (lack of possibility to consider the DM's nature; lack of application to one-shot decisions or asymmetric distribution of losses; generation of irrational rankings). Besides classical decision rules, there are of course many extended decision rules designed for uncertain decision making, but they refer to the probability calculus (e.g. Basili and Chateauneuf, 2011; Ellsberg, 2001; Etner et al., 2012; Garcia et al., 2012; Ghirardato et al., 2004; Gilboa, 2009; Gilboa and Schmeidler, 1989; Hildebrandt and Knoke, 2011; Marinacci, 2002; Pereira et al., 2015; Perez et al., 2015; Tversky and Kahneman, 1992).

In the next section we are going to describe in detail the problem to be solved, and suggest a new decision rule for that purpose.

5 Three-criteria decision rule for SPQP and entirely new devices

In previous sections we have demonstrated that (1) SPQP under complete uncertainty was worth investigating and (2) classical and extended decision rules were not appropriate to solve that problem. In this section all the assumptions connected with the chosen problem are gathered and a novel procedure is proposed.

The scenario-based SPQP model contains the following assumptions:

1) states and the loss matrix are known, but the probability (frequency) of particular scenarios is not known (entirely new devices, lack of historical data);
2) cost c_2 is not treated as a deterministic parameter since it concerns the future: it is given as an interval parameter, which means that parameter s_2 is also interval and the loss matrix is partially interval (Table 5 presents fictitious illustrative prices);

3) the problem concerns a one-period horizon (short life cycle devices) and the period ends when the machine is withdrawn from service (spare parts at cost c_1 are purchased only together with the purchase of the machine);
4) the final recommendation takes into account the DM's nature, i.e. his/her attitude towards a given problem (coefficients α and β);
5) the optimal decision is performed only once (one-shot decision): in the future, due to new experiences, the DM's nature and the loss matrix may change;
6) the optimality is checked using the weighted and arithmetical averages and the standard deviation of all losses (each loss connected with a given alternative has an impact on the final choice, not only extreme losses);
7) each order quantity may be optimal (depending on coefficients α and β) since each one is Pareto-optimal (which is not the case for classical decision rules);
8) the model is useful for both active and passive DMs (we assume that an active DM is a person who intends to analyze all the values very carefully and even influence the particular steps of the algorithm; while a passive DM only declares his/her coefficient of optimism and waits for the final recommendation);
9) the status of each scenario is defined globally, not locally.

Table 5: Partially interval loss matrices – examples 4-5

Ex.	***4.*** $c_1=50$, $c_2\in[51,52]$, $s_1=50$, $s_2\in[1,2]$					***5.*** $c_1=1$, $c_2\in[41,51]$, $s_1=1$, $s_2\in[40,50]$				
S \ A	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
S_1	0	50	100	150	200	0	1	2	3	4
S_2	[1,2]	0	50	100	150	[40,50]	0	1	2	3
S_3	[2,4]	[1,2]	0	50	100	[80,100]	[40,50]	0	1	2
S_4	[3,6]	[2,4]	[1,2]	0	50	[120,150]	[80,100]	[40,50]	0	1
S_5	[4,8]	[3,6]	[2,4]	[1,2]	0	[160,200]	[120,150]	[80,100]	[40,50]	0

Source: Prepared by the author.

The investigation of SPQP under complete uncertainty with interval unit purchase costs of the subassembly just after the failure (see assumption 2) is desirable because that price is related to the future and future purchase times and circumstances are not known exactly, especially in the case of totally new devices. The interval cost c_2 influences particular states of nature to a different extent (compare, for instance, scenarios S_1 and S_4, Table 5). Intervals in the matrix have different widths, i.e., differences between their endpoints (e.g., 50-40 = 10 and 150-120 = 30), and they occur only in the bottom left corner of the matrix. Losses connected with the first scenario and the last decision are given as point values.

The analysis of the standard deviation (assumption 6) is crucial because the ranges of losses related to particular order quantities vary rather significantly.

The procedure developed for the aforementioned problem refers to several other approaches described in the literature.

First, we are going to apply some elements of the (H+B) rule presented in Gaspars-Wieloch (2014a, 2015b, 2016b), which is a hybrid of the Hurwicz and Bayes decision rules. That method was originally worked out for profit matrices, but it can be easily modified for loss matrices. The hybrid, thanks to parameters $\alpha \in [0,1]$ and $\beta = 1 - \alpha \in [0,1]$, takes into account the DM's preferences (as does the Hurwicz rule). In (H+B) rule, in contrast to the Hurwicz, Wald, Hayashi, and Savage approaches, all the outcomes influence the value of the final measure, which is quite advantageous for cases where alternatives contain many payoffs almost equal to the extreme values. The general idea of H+B is to assign, for a pessimist, α to the last term of the non-increasing sequence of all the payoffs related to a given decision and β to the remaining terms of that sequence. For an optimist, weights are set in a different way: β is connected with the first term of the sequence and α with the remaining ones. The assignment of parameters α and β to particular payoffs, depending on the level of optimism, is justified in Gaspars-Wieloch (2014a, 2017b) where the author suggests a significant modification of the classical Hurwicz decision rule and adds to that procedure certain features characteristic for the Bayes rule. The idea of the hybrid presented in Gaspars-Wieloch (2014a) is to recommend, for a strong pessimist, an alternative with a relatively high payoff $a_{j,min}$ or with quite frequent payoffs (almost) equal to $a_{j,max}$ since the pessimist fears the worst, regardless of the decision selected, and that is why such a DM needs an alternative which is attractive even if the worst state occurs and which gives a feeling of security. On the other hand, that rule suggests, for a strong optimist, an alternative with the highest (or almost the highest) payoff $a_{j,max}$, but its highest payoffs do not have to be frequent since the optimist is almost or even completely sure that the best scenario will occur regardless of the decision selected.

Second, due to the fact that in SPQP the ranges of losses related to particular alternatives vary rather significantly, we will support the (H+B) rule with an additional auxiliary decision tool, which analyzes the deviations between outcomes (Gaspars-Wieloch, 2015a, 2017b; Ioan and Ioan, 2011).

Third, we perceive a necessity to refer to the SF+AS (scenario forecasting and alternative selection) procedure recommended in Gaspars-Wieloch (2015a). Its general idea is to (1) forecast the set of scenarios with the largest subjective chance of occurrence (according to the DM's level of pessimism/optimism), see assumption 9, and (2) select a suitable alternative on the basis of a reduced payoff matrix. The use of certain SF+AS features is crucial in SPQP under complete uncertainty since, due to the existence of zero losses for each decision, the original (H+B) decision rule, just like the Hurwicz rule, unfortunately recommends the same optimal order quantities as the Wald rule does, regardless of the DM's nature.

Fourth, we intend to choose a tool enabling one to analyze interval values (see parameter c_2). One may apply, for instance, (1) fuzzy numbers and sets (which requires the estimation of additional parameters, such as degrees of membership), (2) the average cost c_2, (3) the level of c_2 which corresponds to the DM's nature, (4) a meta loss matrix (containing scenarios with the same demand and different values of c_2). Here, we decide to create two loss matrices for extreme cases (i.e. endpoints of interval $[c_{2,min}; c_{2,max}]$) and to compare the recommended solutions.

The suggested three-criteria rule for SPQP and totally new devices consist of the following steps:

1) Define $q_{min} = D_{min}$, $q_{max} = D_{max}$, m (number of scenarios), n (number of decisions), the set of alternatives (A) and the set of scenarios (S); this is performed mainly by experts;
2) Estimate cost c_1 as a point value and cost c_2 as an interval value: $[c_{2,min}; c_{2,max}]$. Compute s_1, s_2 and generate the partially interval loss matrix; this is performed mainly by experts;
3) Determine α and β (subjectively or on the basis of psychological tests). The coefficients should describe the DM's attitude towards a demand for spare parts. If $\alpha \in [0,0.5[$, then $\alpha = \alpha_o$, $\beta = \beta_o$ (α_o and β_o are optimist's coefficients). If $\alpha \in]0.5,1]$, then $\alpha = \alpha_p$, $\beta = \beta_p$ (α_p and β_p are pessimist's coefficients);
4) Assign an interval for the coefficient of optimism to each scenario. The width w of the range for each state of nature is defined as follows:

$$w = \frac{1}{m} \tag{5}$$

The extreme values (b_i and t_i) of interval $[b_i; t_i]$ set for scenario S_i, i.e. its endpoints, are computed according to Equations (6)–(7):

$$b_i = \frac{D_{\max} - D_i}{D_{\max} - D_{\min} + 1} \qquad i = 1,\ldots, m \tag{6}$$

$$t_i = \frac{D_{\max} - D_i + 1}{D_{\max} - D_{\min} + 1} = b_i + w \qquad i = 1,\ldots, m \tag{7}$$

Apart from the interval for the highest demand (i.e. the last scenario), the intervals are left-open, i.e. $]b_i; t_i]$ for $i = 1,2,\ldots, m$-1 and $[b_i; t_i]$ for $i = m$.

5) Find the scenario which corresponds to the coefficient of optimism declared by the DM (according to intervals computed in step 4). Let us denote this state of nature by S^*_i, and the losses connected with S^*_i by $a^*_{i,1}, \ldots, a^*_{i,n-1}, a^*_{i,n}$;
6) Create two loss matrices: matrix I containing losses calculated on the basis of $c_{2,min}$ and matrix II for $c_{2,max}$. Perform steps 7-10 separately for each matrix;
7) Calculate, for each decision, index hb_j (hb^p_j, hb^o_j or $hb^{0.5}_j$ depending on the parameter α). If $\alpha \in]0.5,1]$, calculate hb^p_j (index for pessimists) from Equation (8). If $\alpha \in [0,0.5[$, compute hb^o_j (index for optimists) following formula (9). If $\alpha = 0.5$, calculate $hb^{0.5}_j$ using Equation (10), where b_j denotes the Bayes criterion, i.e. the average of all losses.

$$hb_j^p = \frac{\alpha_p \cdot a^*_{i,j} + \beta_p \cdot \left(\sum_{i=1}^{m} (a_{i,j}) - a^*_{i,j} \right)}{\alpha_p + (m-1) \cdot \beta_p} \tag{8}$$

$$hb_j^o = \frac{\alpha_o \cdot \left(\sum_{i=1}^{m} (a_{i,j}) - a^*_{i,j} \right) + \beta_o \cdot a^*_{i,j}}{(m-1) \cdot \alpha_o + \beta_o} \tag{9}$$

$$hb_j^{0.5} = hb_j^p = hb_j^o = b_j = \frac{1}{m} \cdot \sum_{i=1}^{m} a_{i,j} \tag{10}$$

The denominators in Equations (8)-(9) are introduced so that the final values of the particular indices belong to the interval $[a_{j,min};a_{j,max}]$. Denominators are not crucial and can be omitted when preparing the ranking;

8) Choose alternative A^*_j fulfilling condition (11). Options A^*_j chosen on the basis of matrices I and II belong to sets A^*_{I} and A^*_{II}, respectively. If, within a given matrix, there are alternatives with indices hb_j very close to the smallest one, they may also be selected by the DM as elements of sets A^*_{I} and A^*_{II};

$$A_j^* = \arg\min_j (hb_j) \tag{11}$$

9) If selected decisions A^*_j fulfill Equations (12)-(19), $A^*_j = A^{**}_j$. Go to step 11. Otherwise, go to step 10.

$$b_j^* \leq \beta \cdot (b_{\max} - b_{\min}) + b_{\min} \tag{12}$$

$$b_j = \frac{1}{m} \cdot \sum_{i=1}^{m} a_{i,j} \quad j = 1, \ldots, n \tag{13}$$

$$b_{\max} = \max_j \{b_j\} \tag{14}$$

$$b_{\min} = \min_j \{b_j\} \tag{15}$$

$$\sigma_j^* \le \beta \cdot (\sigma_{\max} - \sigma_{\min}) + \sigma_{\min} \tag{16}$$

$$\sigma_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(a_{i,j} - b_j\right)^2} \quad j = 1, \ldots, n \tag{17}$$

$$\sigma_{\max} = \max_j\{\sigma_j\} \tag{18}$$

$$\sigma_{\min} = \min_j\{\sigma_j\} \tag{19}$$

10) Find the nearest decision (by gradually increasing or decreasing the order quantity) satisfying Equations (12)-(19) and denote it by $A^{**}{}_j$;
11) Decisions $A^{**}{}_j$ chosen on the basis of matrices I and II belong to sets $A^{**}{}_{\mathrm{I}}$ and $A^{**}{}_{\mathrm{II}}$, respectively. If both sets are singleton sets and $A^{**}{}_{\mathrm{I}} = A^{**}{}_{\mathrm{II}}$, then decision $A^{**}{}_j$ (equivalent for both sets) is the suitable one (let us denote it by $A^{***}{}_j$); Otherwise, go to step 12;
12) If at least one set ($A^{**}{}_{\mathrm{I}}$ or $A^{**}{}_{\mathrm{II}}$) is a multi-element one and both sets contain exactly one common decision $A^{**}{}_j$, then that decision is the suitable one (i.e. $A^{***}{}_j$). Otherwise, go to step 13;
13) If both sets contain more than one common decision $A^{**}{}_j$, choose option $A^{***}{}_j$ according to Equations (20)-(22). Otherwise, go to step 14;

$$A_j^{***} = A_j^{**} \mid \begin{cases} \left\lceil \dfrac{j_{\min}^{**} + j_{\max}^{**}}{2} \right\rceil, & if \quad \beta \in [0;0.5[\\ \left\lfloor \dfrac{j_{\min}^{**} + j_{\max}^{**}}{2} \right\rfloor, & if \quad \beta \in [0.5;1] \end{cases} \tag{20}$$

$$j_{\min}^{**} = \min_{A_j^{**} \in A_I^{**} \wedge A_j^{**} \in A_{II}^{**}}\{j\} \tag{21}$$

$$j_{\max}^{**} = \max_{A_j^{**} \in A_I^{**} \wedge A_j^{**} \in A_{II}^{**}}\{j\} \tag{22}$$

14) If the two sets are disjoint, choose option $A^{***}{}_j$ according to formulas (23)-(25).

$$A_j^{***} = A_j^{**} \mid \begin{cases} \left\lceil \dfrac{j_{\min}^{**} + j_{\max}^{**}}{2} \right\rceil, & if \quad \beta \in [0;0.5[\\ \left\lfloor \dfrac{j_{\min}^{**} + j_{\max}^{**}}{2} \right\rfloor, & if \quad \beta \in [0.5;1] \end{cases} \tag{23}$$

$$j_{\min}^{**} = \max_{A_j^{**} \in A_I^{**}}\{j\} \tag{24}$$

$$j_{\max}^{**} = \min_{A_j^{**} \in A_{II}^{**}}\{j\} \tag{25}$$

In the last part of section 5 we explain in detail steps, terms and equations of the above algorithm.

Steps 4 and 5 refer to the SF+AS procedure (Gaspars-Wieloch, 2015a), which consists in predicting the scenario with the greatest subjective chance of occurrence on the basis of the coefficient of optimism, but this time, instead of dominance cases used in the original version, a new method is applied. The reasoning is as follows: the more optimist the DM is, the more probable is the minimal demand for extra parts, so the largest values of β are assigned to scenario D_{min}. The use of a different approach (as compared to the original SF+AS procedure) is justified below. In SPQP the situation is very specific: each successive state of nature is connected with a greater number of failures, hence with worse conditions. Therefore, the status of the particular scenarios can be assessed even without the knowledge of all the losses connected with a given event.

In steps 7 and 8 we refer to the hybrid of Hurwicz and Bayes rule. However, this time, we assign the highest coefficient (α or β) not to the extreme value (the lowest or the highest one), but to the value connected with the scenario with the highest subjective chance of occurrence. Such a modification results from the fact that in SPQP the status of particular scenarios can be evaluated in a global way, thus it does not change depending on the order quantity considered. The idea to treat the scenario status globally (not locally) has been already suggested by Milnor (1954) who stated that each decision rule theoretically designed for games against nature, which treats nature as a conscious opponent who is altering strategies depending on the outcomes, is wrong and unsatisfactory.

Steps 7 and 8 use the first criterion in the three-criteria decision rule, i.e. the weighted average of losses. The second and third criteria (arithmetical average and standard deviation, see Equations 12 and 16, step 9) are introduced in order to find a relatively safe strategy (i.e. an alternative with a relatively small range of losses and as few high losses as possible), which is particularly important in the case of cautious DMs. Of course, the arithmetical average and the standard deviation are just suggestions. One may use other measures, such as ranges of losses connected with particular order quantities. Note that the last two criteria are applied in the algorithm only to decisions satisfying the first criterion. However, if there are other decisions with indices hb_j very close to the lowest one, we recommend calculating and comparing the values of the second and third measures for the whole subset containing the best strategies according to the first criterion. We purposely do not define the acceptable distance between the lowest index hb_j and the other ones: we leave it to the DM.

Step 11 results from the fact that if for $c_{2,min}$ the only solution A^{**}_j is the same as for $c_{2,max}$, then for any value from interval $[c_{2,min}; c_{2,max}]$, solution A^{**}_j will be the same.

As a matter of fact, the ceiling and floor in Equations (20) and (23) (steps 13-14) are useful only if one device is bought. When more than one machine is bought, the ceiling and floor in that formula are not crucial, since the final result may be non-discrete (e.g. we might buy 5.5 spare parts on average, i.e. for some devices 5 and for others, 6). The non-discrete average is appropriate only if all devices are identical and purchased for the same project (company). In the case of the purchase of one machine, we assume that for optimists (pessimists) we search for the floor (ceiling) of that ratio (Equation 20 or 28) since optimists (pessimists) expect a low (high) demand and a low (high) cost c_2.

6 Example

In this section we are going to solve Example 6 (Table 6) by means of the three--criteria decision rule. Let us assume that an engine with a totally new technology is bought. All steps are analyzed below:

1) $q_{min} = D_{min} = 0$, $q_{max} = D_{max} = 4$, $n = m = 5$, $S = \{S_1,S_2,S_3,S_4,S_5\}$, $A = \{A_1,A_2,A_3,A_4,A_5\}$;
2) $c_1 = 10$; $17 \le c_2 \le 25$; $s_1 = 10$; $7 \le s_2 \le 15$. The loss matrix is shown in Table 6.
3) $\alpha = 0.8$, $\beta = 0.2$ (the DM is a moderate pessimist) $\rightarrow$ $\alpha = \alpha_p$, $\beta = \beta_p$;
4) $w = 1/m = 0.2$. Intervals: [0;0.2] for S_5,]0.2;0.4] for S_4,]0.4;0.6] for S_3,]0.6;0.8] for S_2 and]0.8;1.0] for S_1;
5) The scenario with the greatest chance of occurrence is $S_i^* = S_5$ since $\beta_p = 0.2 \in [0;0.2]$. The most "probable" losses are: $a^*_{i,1} = [28,60]$, $a^*_{i,2} = [21,45]$, $a^*_{i,3} = [14,30]$, $a^*_{i,4} = [7,15]$, $a^*_{i,5} = 0$;
6) Matrices I and II contain losses equal to the left and right interval endpoints, respectively (Table 7);

Table 6: Partially interval loss matrix – example 6

Ex.	***6. c_1=10, $c_2 \in$ [17,25], s_1=10, $s_2 \in$ [7,15]***				
S \ A	***A_1***	***A_2***	***A_3***	***A_4***	***A_5***
S_1	0	10	20	30	40
S_2	[7,15]	0	10	20	30
S_3	[14,30]	[7,15]	0	10	20
S_4	[21,45]	[14,30]	[7,15]	0	10
S_5	[28,60]	[21,45]	[14,30]	[7,15]	0

Source: Prepared by the author.

Table 7: Matrices I and II (losses and computations) – example 6

Ex.	***Matrix I. c_1=10, c_2=17, s_1=10, s_2=7***					***Matrix II. c_1=10, c_2=25, s_1=10, s_2=15***				
S \ A	A_1	A_2	A_3	A_4	A_5	A_1	A_2	A_3	A_4	A_5
S_1	0	10	20	30	40	0	10	20	30	40
S_2	7	0	10	20	30	15	0	10	20	30
S_3	14	7	0	10	20	30	15	0	10	20
S_4	21	14	7	0	10	45	30	15	0	10
S_5	28	21	14	7	0	60	45	30	15	0
HB_p	19.25	14.37	11.62	11.00	12.50	41.25	29.38	20.63	15.00	12.50
Constraints	Average <= 12.16; st. deviation <= 12.49					Average <= 18.00; st. deviation <= 13.69				
Average b_j	14.00	10.40	10.20	13.40	20.00	30.00	20.00	15.00	15.00	20.00
Standard deviation	11.07	7.83	7.50	11.74	15.81	23.72	17.68	11.18	11.18	15.81
HB_p (revised)	19.25	14.37	11.62	11.00	12.50	41.25	29.38	20.63	15.00	12.50

Source: Prepared by the author.

7)-10) Computations for both matrices are also presented in Table 7. As we can see, A_4 is selected in step 8 in matrix I (due to the lowest value hb_j): $A^*_{\mathrm{I}} = \{A_4\}$, but the average b_4 for that decision exceeds the allowed one (13.4 > 12.16). Thus, although its standard deviation satisfies Equation (16): 11.74 < 12.49, one should find another alternative. The nearest acceptable is $A_j^{**} = A_3$, since 10.20 < 12.16 and 7.5 < 12.49. Hence $A^{**}_{\mathrm{I}} = \{A_3\}$. A similar procedure is applied to matrix II. This time, $A^*_{\mathrm{II}} = \{A_5\}$, but the average and standard deviation are too high: 20.00 > 18.00 and 15.81 > 13.69. Therefore, we have to search for A_j^{**}: $A^{**}_{\mathrm{II}} = \{A_4\}$;

11)-14) Sets A^{**}_{I} and A^{**}_{II} contain one element each, but they are disjoint. That is why we move directly to step 14 and choose the final decision: $A^{***} = \{A_4\}$ since $j^{**}_{\min} = 3$, $j^{**}_{\max} = 4$ and $\beta = 0.2$. The optimal order quantity is 3.

At the end of this section we may check the results given by the Hurwicz rule and the original (H+B) rule which theoretically take into account the DM's nature. They also recommend A_3 (matrix I) and A_4 (matrix II), but note that their recommendations will not change after the modification of the coefficient values! If e.g. $\alpha = 0.2$, $\beta = 0.8$ (moderate optimist), the solution suggested by both procedures will be still the same and that is alarming (the reason has been given in previous sections: rankings do not change due to the occurrence of a zero loss for each decision). Fortunately, such a situation will not occur if we apply the three-criteria approach. For a moderate optimist that method recommends A_2.

7 Conclusions

The spare parts quantity problem (SPQP) under complete uncertainty has not been discussed yet in the literature, but we perceive the necessity to investigate this issue since in some cases the probability (frequency) estimation may be onerous (devices with a new technology). We have demonstrated that, due to a very specific structure of the loss matrix, classical decision rules designed for decision making under uncertainty with unknown probabilities cannot be applied to this problem, especially if one intends to take into account the decision maker's attitude towards risk. This paper contains a description of a three--criteria procedure that may be useful for the uncertain version of SPQP with totally new seasonal devices. The novel approach combines a hybrid of the Hurwicz and Bayes decision rules with the average and standard deviation criteria. It also refers to a two-stage procedure (SF+AS) consisting in forecasting the scenario with the largest subjective chance of occurrence before the final selection of the appropriate decision. Another method for spare parts demand forecasting has been already proposed by Romeijnders, Teunter and van Jaarsveld, 2012), for instance.

The three-criteria approach has several significant advantages. First, it takes into consideration the decision maker's attitude towards a given problem and leads to logical results for each kind of decision maker. Second, it may be applied even if the distribution of losses connected with particular alternatives is not symmetric since it examines each loss (not only extreme ones). Third, it has been worked out for the case where the future unit purchase cost of a spare part is given as an interval parameter. Fourth, it analyzes two kinds of uncertainties: uncertain demand for spare parts (discrete random variable with unknown probability distribution) and uncertain future cost of missing parts (interval value). Fifth, it does not require any information about the likelihood, which is useful in the case of new machines and one-shot decisions. It only applies certain secondary probability – like quantities which are not estimated by the DM, but are generated using the coefficient of optimism. Sixth, depending on the DM's commitment, the procedure may be applied by both active and passive decision makers. Seventh, the method is designed for one-shot decisions (i.e. single-period problems), but the obtained recommendation can be used simultaneously for each identical device belonging to a given company.

Note that the new procedure can support the SPQP decision making process, but it is reasonable to use it only in the case of expensive purchases. Otherwise, simple reasoning seems sufficient.

In the future, it would be desirable to analyze SPQP in the context of the length of the period of use (when is the machine going to be withdrawn from service?). This factor is also uncertain and may affect the final decision as well. A similar problem is discussed, e.g., in de Jonge et al. (2015).

References

Aronis K.P., Magou I., Dekker R., Tagaras G. (2004), *Inventory Control of Spare Parts Using a Bayesian Approach: A Case Study*, European Journal of Operational Research, 154, 730-739.

Bartakke M.N. (1981), *A Method of Spare Parts Inventory Planning*, Omega, 9(1), 51-58.

Basili M., Chateauneuf A. (2011), *Extreme Events and Entropy: A Multiple Quantile Utility Model*, International Journal of Approximate Reasoning, 52, 1095-1102.

Bian J., Guo L., Yang Y., Wang N. (2013), *Optimizing Spare Parts Inventory for Time-varying Task*, Chemical Engineering Transactions, 33, 637-642.

Caplan B. (2001), *Probability, Common Sense, and Realism: A Reply to Hulsmann and Block*, The Quarterly Journal of Austrian Economics, 4(2), 69-86.

Carnap R. (1950), *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Czerwinski Z. (1960), *Enumerative Induction and the Theory of Games*, Studia Logica, 10.

De Jonge B., Klingenberg W., Teunter R., Tinga T. (2015), *Optimum Maintenance Strategy under Uncertainty in the Lifetime Distribution*, Reliability Engineering and System Safety, 133, 59-67.

Ellsberg D. (2001), *Risk. Ambiguity and Decision*, Garland Publishing, New York.

Etner J., Jeleva M., Tallon J.-M. (2012), *Decision Theory under Ambiguity*, Journal of Economic Surveys, 26(2), 234-270.

Fera M., Lambiase A., Nenni M.E. (2010), *A Proposal for Estimating the Order Level for Slow Moving Spare Parts Subject to Obsolescence*, iBusiness, 2010, 2, 232-237.

de Finetti B. (1975), *Theory of Probability. A Critical Introductory Treatment*, Wiley & Sons, London.

Fortuin L. (1981), *Reduction of the All-time Requirement for Spare Parts*, International Journal of Operations and Production Management, 2, 29-37.

Fréchet M. (1938), *The Diverse Definitions of Probability*, Lecture read at the fourth International Congress for the Unity of Science, Erkenntnis.

García G.M.J., Hernández R.J.G., Hernández G.G.J. (2012), *Making Decisions under Risk and Uncertainty of the Virtual Channel Manager of the Logistic Model Based on Positions*, [in:] G. Dukic (ed.), Proceedings ICIL'2012, Croatia, University of Zagreb, 448-455.

Gaspars-Wieloch H. (2014a), *Propozycja hybrydy reguł Hurwicza i Bayesa w podejmowaniu decyzji w warunkach niepewności* [*A Hybrid of the Hurwicz and Bayes Rules in Decision Making under Uncertainty*], [in:] T. Trzaskalik (ed.), *Modelowanie preferencji a ryzyko'14*, Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach 178, 74-92 [in Polish].

Gaspars-Wieloch H. (2014b), *Modifications of the Hurwicz's Decision Rules*, Central European Journal of Operations Research, 22(4), 779-794.

Gaspars-Wieloch H. (2014c), *Modification of the Maximin Joy Criterion for Decision Making under Uncertainty*, Quantitative Methods in Economics, 15(2), 84-93.

Gaspars-Wieloch H. (2015a), *On a Decision Rule Supported by a Forecasting Stage Based on the Decision Maker's Coefficient of Optimism*, Central European Journal of Operations Research, 23(3), 579-594.

Gaspars-Wieloch H. (2015b), *Innovative Products and Newsvendor Problem under Uncertainty without Probabilities*, [in:] S.L. Zadnik, J. Zerovnik, M. Kljajic Borstnar, S. Drobne (eds.), Proceedings of the 13th International Symposium of Operational Research SOR'15. Slovenian Society INFORMATIKA (SDI). Section for Operational Research (SOR), 343-350.

Gaspars-Wieloch H. (2016a), *Resource Allocation under Complete Uncertainty – Case of Asymmetric Payoffs*, Organization and Management, 96, 247-258.

Gaspars-Wieloch H., Michalska E. (2016b), *On Two Applications of the Omega Ratio: maxOmegamin and Omega(H+B)*, Research Papers of Wrocław University of Economics, 446, *Metody i zastosowania badań operacyjnych*, Wrocław, 21-36.

Gaspars-Wieloch H. (2017a), *A Decision Rule Based on Goal Programming and One-stage Models for Uncertain Multi-criteria Mixed Decision Making and Games Against Nature*, Croatian Operational Research Review, 8(1), 61-76.

Gaspars-Wieloch H. (2017b), *Newsvendor Problem under Complete Uncertainty: A Case of Innovative Products*, Central European Journal of Operations Research, 25(3), 561-585.

Gaspars-Wieloch H. (2018a), *The Impact of the Structure of the Payoff Matrix on the Final Decision Made under Uncertainty*, Asia-Pacific Journal of Operational Research, 35(1).

Gaspars-Wieloch H. (2018b), *Podejmowanie decyzji w warunkach niepewności – planowanie scenariuszowe, reguły decyzyjne i wybrane zastosowania ekonomiczne* [*Decision Making under Uncertainty – Scenario Planning, Decision Rules and Selected Economic Applications*], Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań [in Polish].

Gaspars-Wieloch H. (2019a), *Project Net Present Value Estimation under Uncertainty*, Central European Journal of Operations Research, 27, 179-197.

Gaspars-Wieloch H. (2019b), *Role of Scenario Planning and Probabilities in Economic Decisions Problems – Literature Review and New Conclusions*, Contemporary Issues in Business, Management and Education – Conference Proceedings.

Gilboa I. (2009), *Theory of Decision under Uncertainty*, Cambridge University Press, Cambridge, New York.

Gilboa I., Schmeidler D. (1989), *Maxmin Expected Utility with Non-unique Prior*, Journal of Mathematical Economics, 18, 141-153.

Gu J., Li K. (2015), *Efficient Aircraft Spare Parts Inventory Management under Demand Uncertainty*, Journal of Air Transport Management, 42, 101-109.

Guide V.D.R., Srivastava R. (1997), *Repairable Inventory Theory: Models and Applications*, European Journal of Operational Research, 102, 1-20.

Guo P. (2011), *One-shot Decision Theory*, IEEE Transactions on Systems, Man and Cybernetics, Part A, 41(5), 917-926.

Hau R., Pleskac T.J., Hertwig R. (2009), *Decisions from Experience and Statistical Probabilities: Why They Trigger Different Choices than a Priori Probabilities?* Journal of Behavioral Decision Making, 23(1), 48-68.

Hayashi T. (2008), *Regret Aversion and Opportunity Dependence*, Journal of Economic Theory, 139(1), 242-268.

Hildebrandt P., Knoke T. (2011), *Investment Decisions under Uncertainty − A Methodological Review on Forest Science Studies*, Forest Policy and Economics, 13, 1-15.

Hurwicz L. (1952), *A Criterion for Decision Making under Uncertainty*, Technical Report 355, Cowles Commission.

Inderfurth K., Mukherjee K. (2008), *Decision Support for Spare Parts Acquisition in Post Product Life Cycle*, Central European Journal of Operations Research, 16(1), 17-42.

Ioan C., Ioan G. (2011), *A Method of Choice of the Best Alternative in the Multiple Solutions Case in the Games Theory*, Journal of Accounting and Management, 1(1), 5-8.

Kennedy W.J., Patterson J.W., Fredendall L.D. (2002), *An Overview of Recent Literature on Spare Parts Inventories*, International Journal of Production Economics, 76, 201-215.

Knight F.H. (1921), *Risk. Uncertainty. Profit*, Hart. Boston MA, Schaffner & Marx, Houghton Mifflin Co.

Kolmogorov A.N. (1933), *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Julius Springer Verlag, Berlin.

Louit D., Banjevic D., Jardine A.K.S. (2005), *Optimization of Spare Parts Inventories Composed of Non-repairable or Repairable Parts*, Proceedings of the International Conference of Maintenance Societies, ICOMS, paper 23, Hobart, Australia.

Marinacci M. (2002), *Probabilistic Sophistication and Multiple Priors*, Econometrica, 70, 755-764.

Milnor J. (1954), *Games Against Nature*, [in:] R.M. Thrall, C.H. Coombs, R.L. Davis (eds.), *Decision Processes*, Wiley, New York, 49-60.

von Mises L. (1949), *Human Action. A Treatise on Economics*, The Ludwig von Mises Institute, Auburn, Alabama.

von Mises R. (1957), *Probability, Statistics and Truth*, The Macmillan Company, New York.

Papathanassiu B., Tsouros C. (1986), *Determining the Optimal Place and Capacity of Spare Parts Warehouses in the Construction Industry*, European Journal of Operational Research, 27(1), 91-94.

Pastore E., Alferi A., Zotteri G. (2015), *Managing Different Demand Classes in a Spare Parts Inventory: A Practical Dynamic Allocation Strategy*, Proceedings of the 13th International Symposium on Operational Research, SOR'15, 79-84.

Pereira Jr J.G., Ekel P. Ya, Palhares R.M., Parreiras R.O. (2015), *On Multicriteria Decision Making under Conditions of Uncertainty*, Information Sciences, 324, 44-59.

Perez D.E., Hernandez J.G., Garcia M.J., Hernandez G.J. (2015), *Hurwicz Method Modified and the Amplitude Model (TAM)*, [in:] N. Delener et al. (ed.), GBATA2015 Reading book, USA : GBATA, 559-566.

Petrovic R., Senborn A., Vujosevic M. (1986), *Hierarchical Spare Parts Inventory Systems*, Studies in Production and Engineering Economics 5, Elsevier, Amsterdam.

Petrovic R., Senborn A., Vujosevic M. (1989), *A New Adaptive Algorithm for Determination of Stocks in Spare Parts Inventory Systems*, Engineering Costs and Production Economics, 15(1), 405-410.

Piegat A. (2010), *Uncertainty of Probability*, Workshop on Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics, Warsaw, Poland.

Pomerol J.C. (2001), *Scenario Development and Practical Decision Making under Uncertainty*, Decision Support Systems, 31(2), 197-204.

Popper K. (1959), *The Propensity Interpretation of Probability*, The British Journal of the Philosophy of Science, 10(37), 25-42.

Qu L., Zhang Q. (2006), *An Overview of Inventory Management for Spare Parts*, Research and Exploration in Laboratory, 7, 004.

Ramsey F. (1931), *Truth and Probability*, [in:] F. Ramsey, *The Foundations of Mathematics and Other Logical Essays*, Kegan Paul, London: Routledge, 156-198.

Ravindran A.R. (2007), *Operations Research and Management Science Handbook*, CRC Press, Boca Raton, London, New York.

Rego J.R., Mesquita M.A. (2011), *Spare Parts Inventory Control: A Literature Review*, Producao, 21(4), 656-666.

Rodriguez M.A., Vecchietti A., Harjunkoski I., Grossmann I.E. (2013), *Optimal Supply Chain Design and Management over a Multi-period Horizon under Demand Uncertainty*, Part I: MINLP and MILP models, Computers and Chemical Engineering, 62, 194-210.

Romeijnders W., Teunter R., van Jaarsveld W. (2012), *A Two-step Method for Forecasting Spare Parts Demand Using Information on Component Repairs*, European Journal of Operational Research, 220(2), 386-393.

Rustenburg W.D., van Houtum G.J., Zijm W.H.M. (2000), *Spare Parts Management for Technical Systems: Resupply of Spare Parts under Limited Budgets*, IIE Transactions, 32, 1013-1026.

Savage L. (1961), *The Foundations of Statistics Reconsidered*, Studies in Subjective Probability, Wiley, New York, 173-188.

Schuh P., Schneider D., Funke L., Tracht K. (2015), *Cost-optimal Spare Parts Inventory Planning for Wind Energy Systems*, Logistics Research, 8, 4.

Sheikh A.K., Younas M., Raouf A. (2000), *Reliability Based Spare Parts Forecasting and Procurement Strategies*, [in:] M. Ben-Daya, S.O. Duffuaa, A. Raouf (eds.), *Maintenance, Modeling and Optimization*, Kluwer Academic Publishers, Boston.

Sikora W. (ed.) (2008), *Badania Operacyjne* [*Operations research*], Polskie Wydawnictwo Ekonomiczne, Warszawa [in Polish].

Stirling W.C. (2003), *Satisficing Games and Decision Making. With Applications to Engineering and Computer Science*, Cambridge University Press, New York.

Tversky A., Kahneman D. (1992), *Advances in Prospect Theory: Cumulative Representation of Uncertainty,* Journal of Risk and Uncertainty, 5, 297-323.

Van Lambalgen M. (1996), *Randomness and Foundations of Probability: von Mises' Axiomatization of Random Sequences*, Statistics, Probability and Game Theory: Papers in Honor of David Blackwell, Institute of Mathematical Statistics, University of Amsterdam.

Verrijdt J., Adan I., Dekok T. (1998), *A Trade off Between Emergency Repair and Inventory Investment*, IIE Transactions, 30, 119-132.

Wald A. (1950), *Statistical Decision Functions*, Wiley, New York.

Wong J.Y.F., Chung D.W.C., Ngai B.M.Y., Banjevic D., Jardine A.K.S. (1997), *Evaluation of Spares Requirements Using Statistical and Probability Analysis Techniques*, Transactions of Mechanical Engineering, I.E.Aust., 22, 77-84.

Zhu X., Guo P. (2016), *The One-shot Decision Theory Based Production Planning Models*, IEEE International Conference on Industrial Engineering and Engineering Management (IEEM).

Zio E., Pedroni N. (2013), *Methods for Representing Uncertainty, A Literature Review*, Apports de la recherche, 2013-3: Risk Analysis, Les Cahiers de la Sécurité Industrielle, FONCSI.

MULTIPLE CRITERIA DECISION MAKING

Vol. 14 2019

Przemysław Juszczuk[*]
Lech Kruś[**]

A FUZZY MULTICRITERIA APPROACH FOR THE TRADING SYSTEMS ON THE FOREX MARKET

DOI: 10.22367/mcdm.2019.14.02

Abstract

The paper relates to the trading systems supporting traders making decision on the forex market. Typical trading systems using technical analysis generate a buy or sell signal when the technical indicator crosses a given oversell or overbought levels. The paper extends the approach in which the above strict crisp conditions are replaced by fuzzy relations. The indicators are treated not independently as it is in the typical systems but jointly. Currency pairs are compared in the mulitichriteria space in which each criterion is defined by a membership function referring to a given indicator. New formulations of the membership functions for different indicators are proposed. General ideas of the algorithm generating non-dominated alternatives in the multicriteria space are presented. The algorithm has been implemented in an experimental system. Computational results for different time windows using real-world data from the forex market are presented and discussed.

Keywords: multicriteria decision making, trading systems, fuzzy sets.

[*] University of Economics in Katowice, Faculty of Informatics and Communication, Department of Knowledge Engineering, e-mail: przemyslaw.juszczuk@ue.katowice.pl, ORCID: 0000-0001-7893-5410.

[**] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland, e-mail: Lech.Krus@ibspan.waw.pl, ORCID: 0000-0002-5946-9757.

1 Introduction

Fully automatic trading systems along with various social trading platforms are nowadays becoming one of the most popular possibilities to invest in a range number of markets. A still growing number of instruments available to invest creates a wide range of opportunities to create Intermarket portfolios. At the same time, the high accessibility of various financial analysis tools opens possibilities for decision makers to create their own strategies. Unfortunately, the lack of transparency of risk-related concepts often leads to a situation in which many investors use very risky strategies without any prior domain knowledge. This is especially true for the high-volatility markets related to currency pairs (such as forex market) and cryptocurrencies. Extreme volatility, and at the same time little predictability of possible crucial events on the market can easily lead to losses in the account balance.

The concepts presented here can be adapted to any market, but, due to the large variety of instruments, correlations among instruments and high liquidity on the market, we select forex as our experimental environment. Among some indisputable advantages of the forex market, we could mention the very low transactions costs. This supports the approach that this element of additional costs can be omitted. Thus this paper is entirely directed towards the efficiency of trading systems itself. More information about the forex market with emphasis on trading systems can be found in Chaboud et al. (2014).

Tools present on various trading platforms can be successfully adapted to various instruments. For some more liquid markets, however, it is more natural that such an environment attracts more attention from the decision makers. In the sense of financial analysis, we can note the two most popular concepts used as an order-opening trigger on the market. The first one is called fundamental analysis and includes the use of various economic factors that have a great potential impact on the price direction. The second concept emerges from the Dow theory and is called technical analysis. Assuming that history repeats itself, the various mathematical formulas are in this case used to calculate market indicator values. Finally, these indicators are analyzed to seek patterns and similarities. With the use of some earlier predefined rules we assume that the behavior of the market will be repeated.

Unfortunately, to seek very promising signals, a number of different indicators should be analyzed at the same time. That obviously leads to a very small number of signals generated directly to the decision maker. The existing crisp trading systems based on certain binary rules rarely give signals.

On the other hand, simple trading systems focused on a few indicators generate a large number of poor quality signals.

This paper presents current results of the research dealing with multicriteria fuzzy approach supporting decisions on the forex market. The initial results of the research have been discussed in Juszczuk and Kruś (2017). New results in this paper include the implementation of three classes of trading systems: traditional (crisp), fuzzy and multicriteria fuzzy. A new formulation of the fuzzy activation function (membership function), general for different indicators is proposed and applied in the second and third classes of the systems. In the third class, the dominance-based algorithm generating non-dominated alternatives is used. Numerical experiments on real-world data from the forex market have been performed for different sets of indicators using the implemented systems. The efficiency of the signals generated by the systems is derived and compared.

This paper is organized as follows. In the next section, we give a short review of the most often cited papers related to trading systems and fuzzy trading systems on the forex market. Section 3 describes in detail a crucial aspect of the proposed approach, that is, the concept of the fuzzification for the selected rules, and also recalls the algorithm used for the selection of non-dominated alternatives. Section 4 describes three cases related to the differences between crisp and fuzzy multicriteria approaches. Section 5 includes numerical experiments, while section 6 consists of conclusions.

2 Related research

The fully automatic trading system is a concept closely related to high-frequency trading. Despite the approach selected (fully automatic or decision support), the forex market seems a perfect fit for a rule-based trading system. One of the important advantages on the market are low transaction costs and the possibility to use the leverage. This supports the approach for which the system could be effective achieving at least a small advantage over random trades (Sewell and Yan, 2008).

Systems including various elements of machine learning have been always commonly used as an investing tool. Approaches related to support vector machines (Lu and Wu, 2009) or widely discussed neural network methods (Kamruzzaman and Sarker, 2003) are commonly used and can be very efficient. One of the most recent papers discussing the use of neural networks along with the gene expression approach can be found in Sermpinis et al. (2012). The main drawback of these methods, however, is their lack of detailed explanation of the signal and of possibility to use this knowledge in the future. The so-called "black-box" approach derives only the result, while the question "how" the signal was derived remains unresolved.

Applications of fuzzy sets to the market are not new. There have been attempts to derive an effective trading system based on the fuzzy sets theory. Some of the first concepts based on Mamdani's rules were introduced in Dourra and Siy (2002). A similar concept assuming the use of rough sets theory was presented in Wang (2003). On the other hand, there is still a wide range of papers proposing new trading rules. For example, Neely et al. (1997) focuses on deriving the new rules using genetic programming. In Neely and Weller (1999) the authors use the same method to generate rules on the basis of a 10-years time span.

Certain machine learning methods adapted for use in forex trading can be found in Booth et al. (2014) whose authors proposed a concept based on the random forest approach. An interesting approach for deriving a support decision system based on the wisdom of the crowd was presented in Gottschlich and Hinz (2014). The authors used the crowd's recommendations as an element of the investment decision. First experiments have shown that this approach could outperform the market benchmark.

A trading system with elements of fundamental analysis was presented in Nassirtoussi et al. (2015). In their analysis, the authors applied text-mining techniques to estimate the direction of daily currency pair changes using news headlines.

All the methods described above apply different approaches combining concepts related to fuzzy sets, decision support systems, automatic rule generation, machine learning techniques, etc. Our proposed approach takes the best concepts from crisp trading systems and expands them so that more signals can be included. The proposed fuzzy trading systems offer more opportunities to open the transactions without significant decrease of the quality of the signals.

3 The development of fuzzified trading systems

Any trading system: fully automatic or a decision support system, can be represented as a flow between independent modules. In this section, we present crucial aspects of our approach. Among these elements, we can find the set of rules related to different market indicators, the process of the fuzzification of criteria and, finally, the dominance-based approach to the derivation of the set of solutions for the decision maker. We developed a system which also allows for a simple comparison of crisp, fuzzy and multicriteria fuzzy approaches described further in this section. The general schema of such systems can be found in Figure 1.

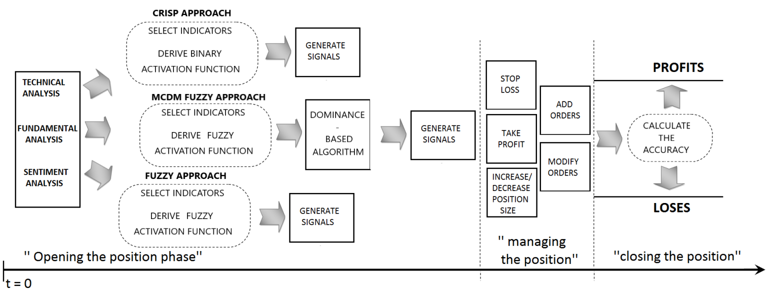


Figure 1: General schema of the trading system

Please note that the entire paper is devoted to deriving the phase responsible for the derivation of the signals. We have also implemented in detail the last phase of a system which allows for the initial evaluation of the signals derived.

3.1 Trading rules and the proposed fuzzy approach

Technical analysis includes dozens of different indicators based on different rules and assumptions. In this paper, we have selected a few indicators most commonly used in modern financial analysis. It is important to note that the concept of fuzzification presented further in this section can be easily adapted to different indicators as well. Additionally, we focus on BUY signals, but opposite rules can be adapted to SELL signals as well. The general trading rule for all indicators considered in this research is as follows:

$$cond_{ind_{Buy}} = true \text{ if } (ind_n(t-1) < c) \wedge (ind_n(t) > c), \quad (1)$$

where ind is any of the indicators used in our research, $t-1$ and t are two successive, discrete time readings, while c is the threshold value for the indicator to generate the signal. The c values for different indicators, in the traditional (crisp) systems, are as follows:

- CCI (Commodity Channel Index), threshold value equal to -100;
- RSI (Relative Strength Index), threshold value equal to 0.3;
- DM (DeMarker), threshold value equal to 0.3;

- Stoch (Stochastic oscillator), threshold value equal to 20 – for the main signal line.

Several readings taken to calculate the values of the indicators were the default and equal to 14. More on the example rules and technical analysis indicators can be found, for example, in Kirkpatrick and Dahlquist (2010).

In the traditional (crisp) trading system the signal is generated only when the indicator value increases and crosses the strict threshold value c according to rule (1). Otherwise, the signal is not generated. In the proposed fuzzification we assume that the threshold value is fuzzy. This means that a fuzzy signal can be generated also when the indicator value crosses a threshold lower or higher than c. In such a case the strength of the signal is lower than in the case of the crisp system.

The most important assumption in the fuzzification process, as in crisp systems, is the dependency between two successive indicator values. The process is conducted only when $ind(t) > ind(t-1)$, thus $\Delta ind > 0$, where $\Delta ind = ind(t) - ind(t-1)$. The strength of the fuzzy signal is calculated by a fuzzy activation function, that is, a membership function with values from [0, 1]. It takes value 1 when the crisp threshold c is crossed. Otherwise, it takes values below 1. One should note that the original signal generated by the crisp version is always present, while the proposed fuzzy approach extends the possibility of the occurrence of additional signals.

Figure 2 shows the example fuzzification process for the CCI indicator. The upper part of the window includes the original CCI data. The indicator value used in the crisp rules is described as "crisp" on the right-hand side of the chart. Two additional boundaries labelled as "upper bound" and "lower bound" are used to set the maximal and minimal indicator values, for which the fuzzy activation function receives a value different from 0. Black vertical lines seen on the upper chart show those fragments of the chart which will be transformed into non-zero values of the fuzzy activation function.

The lower part of Figure 2 corresponds to the generated fuzzy activation functions, which are further used in the experiments.

To sum up:

- in the crisp approach, the binary activation function is present, and possible criterion (market indicator) values are equal to 0 or 1;

- in the fuzzy approach, the initial indicator values are transformed into the fuzzy activation function, for which non-zero values are observed only if the difference between the second and the first indicator values is positive, and at the same time the indicator is within the range $\langle lowerbound; upperbound \rangle$.

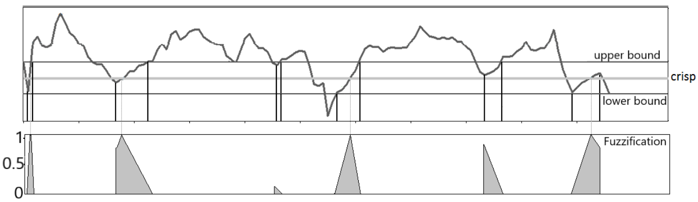


Figure 2: Calculation of the fuzzy membership function values for the example indicator *ind* for successive readings

3.2 Dominance based-algorithm

The traditional (crisp) systems and the proposed fuzzy system generate several signals independently for different currency pairs and different indicators. In general, the systems do not provide information about relations between the signals which would allow to select the best (that is, the most promising) alternatives to make a decision on the market. Therefore, a multicriteria optimization approach is proposed. In this approach, each currency pair is treated as an alternative evaluated by a vector of criteria related to different indicators. The value of each criterion is defined by the fuzzy activation function of the respective indicator and takes a value from the interval $\langle 0, 1 \rangle$. All the currency pairs for which the fuzzy signals are generated can be analyzed and compared in the multicriteria space. A multicriteria optimization problem can be formulated in which we look for non-dominated alternatives. Thus we use the concept originally introduced by the authors of Juszczuk and Kruś (2017). The proposed algorithm which uses dominance relations generates a set of non-dominated alternatives for the decision maker. The concept of the reservation point x is used to initially exclude certain alternatives which do not satisfy the minimal requirements derived by the decision maker.

Input parameters for the dominance-based algorithms are as follows:

- the initial position of the reservation point x;
- the full set of alternatives to be analyzed, denoted by Y;
- the set Y_-, which will include all the alternatives removed by the algorithm;
- the set Y_+, which includes alternatives non-dominated by the point x and analyzed further in the algorithm;
- the set ND, which will include all the non-dominated alternatives selected from the set Y_+.

The whole procedure can be divided into three separate phases. In the first phase, we search for alternatives equal to the aspiration point u (the point for which all criteria take value 1), which could result in an immediate termination of the algorithm and in the derivation of such an alternative by the decision maker. This phase is represented in lines 3-4. The second, longest phase corresponds to the sequential analysis of all alternatives from the set Y_+. All non-dominated alternatives will be moved to the set ND. This phase is represented in lines 5-19. Finally, when there are no more alternatives left to analyze in Y_+, the set ND is given as well as the number of signals generated according to this set.

To calculate the quality of the derived signals we introduce an efficiency measure based on the accuracy formula:

$$acc(p, \epsilon) = \frac{TP}{FP + TP}, \tag{2}$$

where TP is the number of all the signals for which after p readings we have observed at least an ϵ price rise, while FP is the number of all other signals. In general, the efficiency of a trading system is measured as a ratio of the number of acceptable signals to all signals.

4 Case analysis

In this section, we present analysis, discussion and explanation of three different possible cases that could occur on the market in the crisp, fuzzy, and multicriteria approaches.

4.1 Case 1: the existence of alternatives equal to the aspiration point

The existence of alternatives equal to the aspiration point u in the crisp version as well as in the fuzzy version leads to an immediate derivation of the set ND by the algorithm for the decision maker. Each element of the set dominates all remaining alternatives. This situation is presented in Figure 3. Alternative $y_2 = u$ is selected by the algorithm and derived for the decision maker. It dominates all other alternatives. It is selected from all the alternatives generated by the fuzzy system, shown on the left-hand side of the figure. The crisp system generates signals shown on the right-hand side of the figure.

Algorithm 1: Dominance-based algorithm

begin
  Fix the aspiration point u, create sets Y and $ND = \emptyset$
  Set the point x and derive sets Y_- and Y_+
  if *there exists* $y \in Y$ *such that* $y = u$ **then**
    $ND = \{y\}$ End of the algorithm
  for *each alternative* y *in* Y_+ **do**
    if $y \in Y_-$ **then**
      Delete y from further analysis, i.e. $Y_+ = Y_+ \setminus \{y\}$
    else if $y \notin Y_- \wedge ND = \emptyset$ **then**
      Add y to set ND and Update sets Y_- and $Y_+ = Y_+ \setminus \{y\}$.
    else
      for $z \in ND$ **do**
        if $y \succ z$ **then**
          Delete z from ND
        else if $z \succ y$ **then**
          Mark y as dominated, delete it from Y_+, and BREAK
      if y *is non-dominated* **then**
        Add y to ND
        Update set $Y_- = Y_- \cup (y + \mathbb{R}^2_- \setminus \{0\})$
        Delete y from further analysis, i.e. $Y_+ = Y_+ \setminus \{y\}$
  if $Y_+ = \emptyset$ **then**
    End of the algorithm

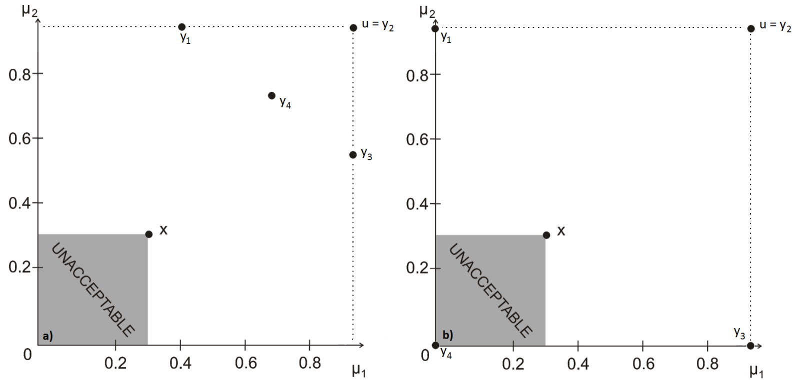


Figure 3: Location of alternatives in the decision space:
a) Fuzzy approach proposed; b) Crisp approach

4.2 Case 2: the existence of dominated alternatives on the boundaries of the search space

The crisp approach does not allow to differentiate the alternatives observed within the boundaries of the search space. Thus every such alternative will be treated as a possible solution derived for the decision maker. This situation for the crisp approach is shown in Figure 4. In this example, the fuzzy system generates four alternatives. Two non-dominated alternatives y_2 and y_3 are selected by the algorithm. The traditional crisp system generates four signals (the right-hand side of the figure) without providing information as to which one is better or worse.

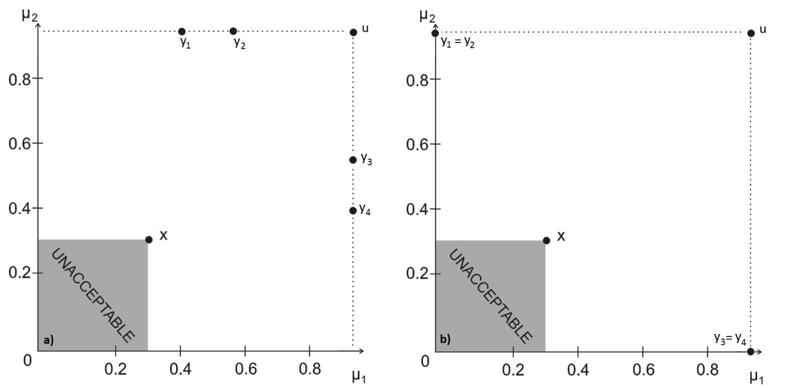


Figure 4: Location of alternatives in the decision space. Alternatives on the boundaries of the search space: a) Fuzzy approach proposed;
b) Crisp approach

At the same time the fuzzy approach proposed here allows to differentiate these alternatives and non-dominated alternatives. The number of such alternatives for the boundary case will be no greater than $n - 1$, where n is the number of criteria available in the system. In general in this case the dominance-based algorithm derives a lower number of non-dominated alternatives than the number of signals generated by the crisp system.

4.3 Case 3: the lack of alternatives on the boundaries of the search space

This case describes the situation for which none of the market indicators present in the system allows in the crisp system to derive a single alternative observed on the boundaries of the search space. The fuzzy approach allows to indicate the alternatives which are relatively close to the boundaries of the search space, where $crit_i < 1$ for any i. By setting the position of the reservation point x, the decision maker can adjust the potential risk related to the situation, by eliminating all alternatives worse than x.

The boundary situation for this case is $x = u$, where no alternatives other than one equal to the aspiration point are good enough for the decision maker. Moving x towards u expands the search space in which alternatives of potential interest for the decision maker can be found. The second boundary situation, in which $x = 0$, is the situation in which all non-dominated alternatives present in the set of alternatives are derived for the decision maker. This situation is shown in Figure 5.

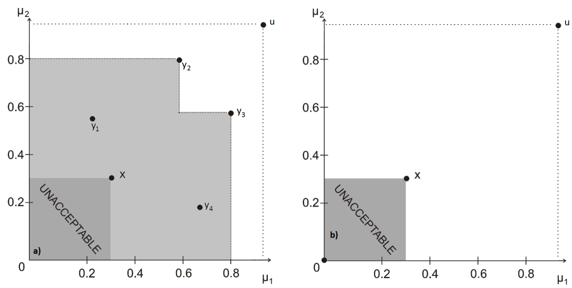


Figure 5: a) Step s^2 of the algorithm; b) the crisp system does not generate signals

To sum up, for the crisp approach the possible signals are generated only when at least one of the critieria considered is equal to 1. For the proposed fuzzy system, this case is extended for the situation in which the value of at least one criterion is greater than x. In general, the fuzzy approach includes all signals generated by the crisp approach and also additional signals, for which $\exists crit_i, x < crit_i < u, i = 1, 2, ..., n$, where n is the number of criteria considered. Finally, in the proposed multicriteria fuzzy approach we assume that dominance relations in the criteria space are included. Thus only non--dominated alternatives from the previous fuzzy approach are selected and derived for the decision maker.

5 Numerical experiments

In this section, we perform a set of experiments using real-world data. To estimate the efficiency of different systems and different trades we analyzed 15 different currency pairs and a time span equal to 16 months (from September 2016 to December 2017). All experiments were performed on 10 different trading systems and three different approaches: crisp, fuzzy, and multicriteria fuzzy approaches. We selected all possible combinations of indicators described in section 2 (with at least two indicators). The position of the reservation point x for all cases was set to 0.85. Below we discuss selected results of these systems:

- The number of signals derived by each trading system and each approach,
- The efficiency of the trading system,
- The efficiency of the trading system in which only alternatives equal to u were analyzed.

First of all, in Table 1, we analyzed the number of signals generated by the systems. The crisp approach is used as a benchmark for the proposed fuzzy approaches. For the crisp approach, we assumed that at least one criterion should be equal to 1, so the ideal alternatives (equal to point u) are included in this comparison as well.

A similar assumption was used for our proposed fuzzy approach. However, instead of the classical binary activation functions, fuzzy activation functions were used. It is obvious that the number of signals generated will be slightly higher than the numbers included for the crisp approach; compare with Case 3 in section 4.3.

Finally, to limit the number of signals generated for the decision maker and to improve their efficiency, we applied our proposed dominance-based algorithm, which was used to derive only non-dominated signals. This number is shown in the last column of Table 1. We note the limited number of signals where the dominance concept is adapted.

In Table 2 we analyzed the average accuracy for all the trading systems and three different approaches. The accuracy for all cases was measured

Table 1: Number of signals derived by all systems for three different approaches

Trading system	Crisp	Fuzzy	MCDM Fuzzy
CCI RSI Stoch DM	4126	4354	2713
CCI RSI Stoch	3417	3643	2150
CCI RSI DM	3112	3249	2043
RSI Stoch DM	2610	3013	1875
CCI RSI	2334	2405	1442
CCI Stoch	3165	3349	1934
CCI DM	2853	3130	1806
RSI Stoch	1753	1934	1226
RSI DM	1387	1458	1075
Stoch DM	2320	2504	1637

after 3 readings ($p = 3$). The minimal price difference between the actual price and the price observed 3 readings ago was equal to $\epsilon = 10$. Note that we did not use any additional money management method, thus we focus mostly on comparing results for different approaches. Thus, as it was expected, the accuracy is fairly small.

Note that despite a larger number of signals derived by the fuzzy approach (as compared with the crisp approach), the accuracy dropped only for certain selected currency pairs. That implies that small deviations from the crisp approach as well as moving the reservation point x away from the aspiration point u does not necessarily mean a significant drop in accuracy.

At the same time, the results derived for the MCDM fuzzy approach are better than those for the two remaining cases. This is mostly due to the fact that only non-dominated solutions have been included. However, there is still room for improvements, because there was no trading system which could achieve the accuracy value above 50%. The results are presented as the average for all currency pairs, so we could still assume that for the specific instruments threshold of 50% could be broken.

Finally, in Table 3 we present the accuracy for all the alternatives equal to the aspiration point u. This is obviously the same for all the approaches, so the results are included in one column only. The most important observation is that the alternatives for which all the criteria are equal to 1 provide better results. For most of the trading systems we observed an accuracy above 50%. At the same time, one should note that for the MCDM fuzzy approach the number of market indicators involved in the trading system does not significantly affect the accuracy. However, it obviously affects the number of signals generated for the decision maker.

Table 2: Accuracy of different trading systems for the three approaches analyzed

Trading system	Crisp	Fuzzy	MCDM Fuzzy
CCI RSI Stoch DM	46.04%	43 04%	48.34%
CCI RSI Stoch	43.68%	45.68%	47.9%
CCI RSI DM	42.27%	43.27%	48.51%
RSI Stoch DM	49.13%	43.18%	49.15%
CCI RSI	48.20%	48.20%	48.58%
CCI Stoch	41.96%	43.96%	48.00%
CCI DM	44.31%	45.31%	48.36%
RSI Stoch	42.71%	40.71%	47.83%
RSI DM	45.96%	44.96%	49.16%
Stoch DM	49.11%	49.11%	48.78%

Table 3: Accuracy for the case in which an alternative equal to aspiration point u exists

Trading System	Signals	Efficiency
CCI RSI Stoch DM	3	66.67%
CCI RSI Stoch	16	81.25%
CCI RSI DM	20	50.00%
RSI Stoch DM	11	54.55%
CCI RSI	122	52.86%
CCI Stoch	246	44.87%
CCI DM	189	60.23%
RSI Stoch	82	53.85%
RSI DM	79	44.67%
Stoch DM	101	53.00%

6 Conclusions

In this paper, we have described the concept of fuzzy trading systems for the forex market. The original crisp trading system involving the use of crisp rules has been transformed into the fuzzy approach. Each criterion present in the crisp version of the system has been changed to corresponded to the concept of the single fuzzy membership function. Next, the rank-domination based algorithm has been discussed. This method allowed to derive a set of non-dominated alternatives which can be further presented to the decision maker.

In the experimental section of this paper, we have studied three different cases related to the position of alternatives for different trading systems (crisp, fuzzy, and multicriteria fuzzy approaches). The results of the observations were confirmed in the numerical section of the paper, where we analyzed the number of alternatives derived for the user by all three approaches. The efficiency of different indicators was measured using the classical accuracy measure. Finally, the efficiency of the ideal alternatives was verified and estimated.

This research shows a simple transition from the classical crisp trading systems to the multicriteria approach with the use of fuzzy sets.

References

Booth A., Gerding E., McGroarty F. (2014), *Automated Trading with Performance Weighted Random Forests and Seasonality*, Expert Systems with Applications, 41, 3651-3661.

Chaboud A.P., Chiquoine B., Hjalmarsson E., Vega C. (2014), *Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market*, The Journal of Finance, 69(5), 2045-2084.

Dourra H., Siy P. (2002), *Investment Using Technical Analysis and Fuzzy Logic*, Fuzzy Sets and Systems, 127, 221-240.

Gottschlich J., Hinz O. (2014), *A Decision Support System for Stock Investment Recommendations Using Collective Wisdom*, Decision Support Systems, 59, 52-62.

Juszczuk P., Kruś L. (2017), *Supporting Multicriteria Fuzzy Decisions on the Forex Market*, Multiple Criteria Decision Making, 12, 60-74.

Kamruzzaman J., Sarker R.A. (2003), *Comparing ANN Based Models with ARIMA for Prediction of Forex Rates*, ASOR Bulletin, 22(2), 1-11.

Kirkpatrick II Ch. D., Dahlquist J.R. (2010), *Technical Analysis: The Complete Resource for Financial Market Technicians*, FT Press, Old Tappan, New Jersey.

Lu C.-C., Wu C.-H. (2009), *Support Vector Machine Combined with GARCH Models for Call Option Price Prediction*, 2009 International Conference on Artificial Intelligence and Computational Intelligence, Shanghai, China, 35-40.

Nassirtoussi A.E., Aghabozorgi S., Wah T.Y., Check D., Ngo L. (2015), *Text Mining of News-headlines for FOREX Market Prediction: A Multi-layer Dimension Reduction Algorithm with Semantics and Sentiment*, Expert Systems with Applications, 42(1), 306-324.

Neely C.J., Weller P.A. (1999), *Technical Trading Rules in the European Monetary System*, Journal of International Money and Finance, 18, 429-458.

Neely C.J., Weller P.A., Dittmar R. (1997), *Is Technical Analysis Profitable in the Foreign Exchange Market? A Genetic Programming Approach*, Journal of Financial and Quantitative Analysis, 32, 405-426.

Sermpinis G., Laws J., Karathanasopoulos A., Dunis C.L. (2012), *Forecasting and Trading the EUR/USD Exchange Rate with Gene Expression and Psi Sigma Neural Networks*, Expert Systems with Applications, 39, 8865-8877.

Sewell M.V., Yan W. (2008), *Ultra High Frequency Financial Data*, GECCO '08: Proceedings of the 10th Annual Conference Companion on Genetic and Evolutionary Computation, ACM Press, 1847-1849.

Wang Y.F. (2003), *Mining Stock Price Using Fuzzy Rough Set System*, Expert Systems with Applications, 24(1), 13-23.

Dorota Kuchta[*]

MULTICRITERIA FUZZY EVALUATION OF PROJECT SUCCESS IN R&D PROJECTS

DOI: 10.22367/mcdm.2019.14.03

Abstract

The paper reviews the state of art in project success evaluation, especially with respect to R&D projects, with emphasis on the ambiguity of such evaluation, its strong dependence on the con-text and on the evaluator. We also recall basic information about fuzzy numbers and propose the concept of fuzzy rules which are used in the suggested procedure of R&D project evaluation. The procedure is described and illustrated by means of a real--world case study.

Keywords: project success, research and development project, fuzzy success evaluation.

1 Introduction

Project success is a notion which is not always unequivocal. It depends strongly on the context and on the evaluator. This phenomenon is especially striking as regards research and development projects. Is an archaeology project, which revealed approximate chances of interesting excavations, successful or not? Or a medical project, in which the medication investigated turned out to be inefficient, but another substance was found which might be efficient, but this can be checked only by conducting another project? Or a project which showed that the assumed procedure is incorrect? Or another one, whose objectives were not attained, but which opened prospects for many new research projects? Or a project which was deemed as being exactly within budget and time, which led to a required number of publications which, however, were written with the full awareness on the part of the authors that the results presented in them are not really valuable?

[*] Wrocław University of Science and Technology, Wrocław, Poland, Dorota.kuchta@pwr.edu.pl, ORCID: 0000-0002-9747-0759.

The author's experience as research and development project manager, as well as the results of a research project (directed by the author) in which R&D projects were examined, show that it is necessary to be very prudent when evaluating R&D projects. First, a variety of criteria has to be used, many of which would not be used in case of e.g. engineering or IT projects. Second, we have to be aware that the evaluation of the success of an R&D project may differ strongly depending on the evaluator and the context. Third, an important issue are soft limits between evaluation grades: it is difficult to say definitely where a research project ceases to be very successful and begins to be only successful.

The objective of this paper is thus to propose a procedure scheme for the evaluation of R&D projects which take into account the issues listed above. The proposed procedure will use the notion of fuzzy sets and fuzzy rules, which help to model soft notions, ambiguity and subjectivity. The proposed procedure will be illustrated with a real-world example.

The content of the consecutive sections is as follows: In section 2 the notion of project success is discussed, for projects in general. In Section 3 this notion is discussed in the context of R&D projects. In section 4 fuzzy sets and fuzzy rules are briefly presented, using an example from the banking sector, as examples from the field of R&D project evaluation are not available. In section 5 the proposed evaluation scheme for R&D projects is described. The paper terminates with conclusions.

2 Project success

Project success (to begin, we will consider any projects, not only R&D projects) is defined in the literature in many different ways. Many authors suggest that a project is successful if it meets the specification (scope), cost (budget) and time (deadline). This is viewed as the most basic level of project success (Greer, 1999), although many authors (e.g. Cheng et al., 2012) use only time and budget as project success measures. However, numerous authors (e.g. Ashley et al., 1987; Baccarini 1999; Baker et al., 1988; Belassi and Tukel, 1996; Camilleri, 2011; Chan and Chan, 2004; Kerzner, 1992; Lim and Mohamed, 1999; Mir and Pinnington, 2013; Pinto and Slevin, 1988; Raz et al., 2002; Thomsett, 2002; Turner, 1994; Wateridge,1998; de Wit, 1988) expand this definition substantially, introducing other project success measures (or, which for us will be synonymous, success criteria).

First of all, the additions to the list of success criteria are a consequence of the following, by now already accepted, opinion: "There have to be two groups of project success measures: objective measures (such as time or cost) and subjective measures (such as the satisfaction of different project stakeholders)" (Chan and Chan, 2004). Subjective measures are necessary, because the perception of project success depends strongly on the assessor (e.g. Davis, 2014). On top of that, other contexts are taken into account, for example (Khan et al., 2013, Freeman and Beale, 1992): organisational benefits, project influence, future prospects gained thanks to the project, technical parameters of the project product, personal development and business profits.

It has to be pointed out that some of the criteria mentioned above are in fact sets or categories of several criteria. For example, the criterion "satisfaction of the stakeholders" comprises individual criteria of several stakeholders' satisfaction, and even this can be split into satisfaction with various aspects of the project. Also, e.g., the criterion "organisational benefits" may incorporate various benefits. So, and this is clearly shown by the literature, there exist a huge set of different success criteria, which, for transparency reasons, are grouped in success criteria categories. For example, Shenhar et al. (1997) propose a grouping of 13 success criteria in four categories: 1. achieving planned objectives, 2. benefits for the customer, 3. commercial success, 4. future potential.

Some project success criteria are objective, some subjective. For the former ones there exists a natural measurement scale, for the latter ones an evaluation scale has to be elaborated. But independently of the subjectivity or objectivity of the criterion there is always a subjectivity component due to the decision maker as to the interpretation of the linguistic expressions such as "high" success, "low" success, "full" success, "partial" success, etc. This depends in each case on the individual decision, which in turn is a consequence of the situation of the project in question and the preferences of various project stakeholders. In the literature, various approaches are proposed. For example, Yourdon (1997) is of the opinion that 50% "less" or "more" than planned in a negative direction (e.g., less profit more cost etc.) in any of project success parameters means a complete failure (e.g. 50% less scope, more money, more time or less quality). Cheng, Tsai and Sudjono (2012) propose a lower bound of 80%-90% in the achievement of the quality and scope as the minimal requirement for a project to be considered successful. Nahod et al. (2013) formulate the following proposal in this respect (they consider time, budget, scope and customer satisfaction as the aspects which constitute project success):

Table 1: Definition of project success in terms of time

Failure	Almost failure	Almost success	Success
PAD>115%PPD	115%PPD ≥PAD>105%PPD	105%PPD ≥PAD>100%PPD	PAD≤100%PPD

(PAD = project actual duration; PPD = project planned duration)

The definition of project success in terms of budget (i.e. failure, almost failure, almost success and success) is defined analogously to Table 1.

Table 2: Definition of project success in terms of scope

Failure	Almost failure	Almost success	Success
AS<80%PS	90%PS ≥AS>80%PS	100%PS >AS>90%PS	AS≥100%PS

(AS = actual scope; PS = planned scope)

The overall evaluation of the *project outcome* for the project manager should be expressed linguistically: failure, almost failure, partial success and success, and these expressions should be defined with respect to the selected evaluation scale. The overall evaluation of the *customer satisfaction* with project results should also be expressed linguistically: low, rather low, rather high and high, and these expressions should also be defined with respect to the selected evaluation scale.

Another issue in project success evaluation is the method of aggregating the various criteria adopted. In other words, should the various project success aspects be synthesised into one measure, which would evaluate the overall project success, by giving one number, expressing the degree to which the project was successful, according to the suggestion of Sutton (2005) or Shashi et al. (2014)? Or maybe should the different aspects remain separated? Several authors do not synthesise the success measures in various project success aspects (Cheng et al., 2012; Chow and Cao, 2008). We think that this approach is useful for practical applications, but in the end each organisation or each decision maker should find its or his/her own aggregated decision as to whether the project in question was a success or not. However, this aggregation will strongly depend on the decision maker. Each decision maker has preferences regarding his/her projects, often enforced by the projects' environment or the specific situation in which they are realised. The different aspects (such as time, cost, satisfaction of individual project stakeholders, etc.) are never all equally important and their importance should be judged by the decision makers.

Another problem linked to the definition of project success is that projects implemented in different areas have their own project success aspects and groupings. Here we are dealing with R&D projects, which will be the subject of the next section.

3 Success of research and development projects

R&D projects will be understood here as either research projects (i.e. projects undertaken with the objective of acquiring or generating new knowledge) or research and development projects (projects which use the existing knowledge in order to create new products or processes (Klaus-Rosińska, 2019). Although the "classical" success measures (time, cost, scope) are or can be important depending on the context (for example, time and cost criteria are important in order to formally account for the project grant), the understanding of R&D project success is much more nuanced, even more nuanced than it was described in the previous section for projects in general.

On the basis of interviews with over 60 managers of R&D projects (a detailed description of the survey and its basic results can be found in Klaus-Rosińska 2019 and Kuchta et al. 2017) we can formulate the thesis that the evaluation of the success of R&D projects can be based on the following groups of criteria (of course, these groups are arbitrary and each decision maker can formulate his or her own proposal):

A. Short-term research success,
B. Long-term research success prospects,
C. Short-term financial success,
D. Long-term financial success prospects,
E. Short-term personal development success,
F. Long-term personal success prospects,
G. Satisfaction of external (to the project team) stakeholders.

It has to be stressed that, according to the project managers interviewed, a successful project does not have to be successful in all the above criteria groups. For example, numerous interviewees share the opinion that the success in one or two of the groups is sufficient, where these "one or two groups" can be any of the above listed ones. This shows that the evaluation of the overall R&D project success has to be very flexible and nuanced, based on the experts' opinion and allowing for different views of different stakeholders (Davis, 2014). Moreover, this shows that an aggregated, universal evaluation of the success of a project might not always be desirable: each stakeholder can choose other criteria groups for the project success evaluation.

We will use the notation $\{\mathcal{g}_s\}_{s=1}^{t}$ for the criteria groups, where t stands for the number of criteria (for example, the groups A – G listed above). There are dozens of criteria from these groups (e.g. Klaus-Rosińska, 2019; Elkadi, 2013; Eilat et al., 2008; Revilla et al., 2003). Some of them are:

A. Short-term research success
   a. Achievement of the planned research results,
   b. Achievement of other than planned research results,
   c. Demonstration that the selected research direction was incorrect,
   d. Publications with good bibliometric parameters (one criterion per one team member);

B. Long-term research success prospects
   a. Ideas for new research projects,
   b. New cooperation possibilities,
   c. Promising results in progress;
C. Short-term financial success
   a. Satisfying remuneration for individual members of the project team,
   b. Satisfying net cash flow for the organisation where the project was implemented,
   c. New patents developed;
D. Long-term financial success prospects
   a. Prospects of patents,
   b. Prospect of academia-industry cooperation;
E. Short-term personal development success
   a. Satisfaction of the individual project team members with the improvement of personal skills obtained thanks to the project (one criterion per one team member),
   b. Satisfaction of the senior researchers with the development of skills of the junior researchers (one criterion per each couple senior researcher / junior researcher working under his or her supervision);
F. Long-term personal development prospects
   a. New cooperation possibilities (one criterion per each project team member),
   b. New ideas for unassisted research projects (one criterion per each junior project team member);
G. Satisfaction of external (to the project team) stakeholders
   a. Satisfaction of the financing institution,
   b. Satisfaction of the institution(s) which implemented or were partners in the project.

As for the notation of the elements of the criteria groups $\{\mathcal{g}_s\}_{s=1}^{t}$, for each s=1,…,t we will have $\mathcal{g}_s = \{G_r^s\}_{r=1}^{w_s}$, where w_s is the number of elements of the criteria group $\mathcal{g}_s$.

It has to be stressed that all of the above criteria and criteria groups can be satisfied fully, partially or not at all. In order to express this, in the next section we introduce the notion of fuzzy sets.

4 Fuzzy sets and fuzzy rules

According to Zadeh (1965), we can define fuzzy sets to model human understanding of various concepts. Fuzzy sets defined on the set of real numbers are called fuzzy numbers. Many types of fuzzy numbers exist, but here we will consider only two types: triangular and one-sided.

Definition 1: A triangular fuzzy number $\tilde{A} = \left(\underline{a}, \hat{a}, \overline{a}\right)$ is a feature expressed by means of the membership function $\mu_A{:}\mathcal{R} \to [0,1]$ such that $\mu_A(x)$ expresses the degree (determined by an expert or a group of experts) to which *x* possesses this feature and

$$\mu_A(x) = \begin{cases} 0 \text{ for } x \leq \underline{a} \\ \frac{x-\underline{a}}{\hat{a}-\underline{a}} \text{ for } x\epsilon\left(\underline{a}, \hat{a}\right) \\ 1 \text{ for } x = \hat{a} \\ \frac{\overline{a}-x}{\overline{a}-\hat{a}} \text{ for } x\epsilon(\hat{a}, \overline{a}) \\ 0 \text{ for } x \geq \overline{a} \end{cases} \quad (1)$$

Definition 2: One-sided fuzzy numbers (left-sided or right-sided) are triangular fuzzy numbers such that

a. For a left-sided fuzzy number, $\underline{a} = -\infty$

b. For a right-sided fuzzy number, $\overline{a} = \infty$.

It is important to indicate for which values the membership function takes positive values, that is, to define the support of fuzzy numbers:

Definition 3: The support of the triangular fuzzy number $\tilde{A} = \left(\underline{a}, \hat{a}, \overline{a}\right)$ is the open interval $\left(\underline{a}, \overline{a}\right)$, the support of the left-sided fuzzy number is the half-line $(\infty, \overline{a})$ and that of the right-sided fuzzy number is the half-line $\left(\underline{a}, \infty\right)$.

For example, in the banking sector (Korol, 2012) the experts of each bank can define their understanding of such terms as “very low”, “low”, “average”, “high” and “very high” financial security, or “very low”, “low”, “average” , “high” and “very high” yearly income of the potential borrower. Financial security is not easily measurable, so the experts would be asked to use a predefined scale (e.g. from 0 to 5) to define the first five terms. The borrower’s yearly income is measurable, so the scale for the last five items would correspond directly to the income values. Figure 1 presents example definitions of the terms, generated on the basis of expert opinions.

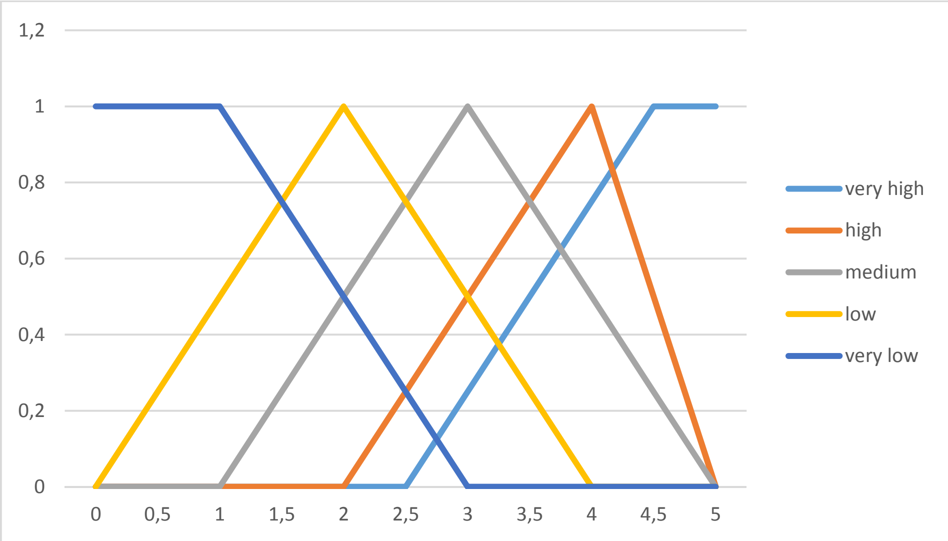


Figure 1: Examples of fuzzy concepts

Source: Author's own elaboration.

Figure1 shows five examples of fuzzy numbers: three triangular, one left-sided and one right-sided. They represent several possible features of financial security or income (measured on the horizontal axis using units which must be predefined, such as hundreds of thousands of dollars for the income or an arbitrary scale for financial security). Each value of the income may have one of the six features fully, partially to a certain degree or not at all. For instance, income 1.5 is not very high at all, not high at all, it is medium to the degree 0.25, low to the degree 0.75 and also very low to the degree 0.75 (according to the experts)

In general, there will be a set of concepts $\mathbb{C} = \{C_i\}_{i=1}^{n}$ (Figure1 refers only to one concept/notion, e.g. either financial security or income) that corresponds to evaluation criteria (in the case of the evaluation of a bank and a potential borrower we would have, for example, the following set $\mathbb{C}$: {age, education, income, type of employment, length of employment, value of the borrower house}) (Korol, 2012). We will also have a set of features (for example, the set {very small, small, medium, big, very big}) $F = \{\tilde{F}_j\}_{j=1}^{M}$, $\tilde{F}_j = (\underline{a}_j, \hat{a}_j, \overline{a}_j)$, with each feature defined by the membership function μ_j^i being either a triangular or a one-sided fuzzy number, according to Definition 1. It is important that for each $\tilde{F}_j$ the corresponding membership function μ_j^i can be different for each element C_i of $\mathbb{C}$ (i.e. the concept "big" can be defined in a different way, for, e.g., income and financial security, it may have other units on the horizontal axis and use a different membership function each time).

Moreover, it is required that for each $i = 1, \dots, n$ and for each element x of $\mathcal{R}^+$ (in the general case the whole set of real numbers can be considered, but here, because of the nature of applications discussed, only non-negative values are taken into account) there exist at least one j_0 such that x belongs to the support of $\mu^i_{j_0}$. This assumption means that for each concept the decision maker has covered all the possible values $\mathcal{R}^+$ with the features: each x has at least one feature from the set F to a positive degree. This condition is satisfied in Figure 1.

In this context it is possible to consider, for each $x \in \mathcal{R}^+$, expressions similar to those used in natural languages, for example "Concept C_i is $\tilde{F}_j$ " for a certain i, , $i = 1, \dots, n$ and j, $j = 1, \dots, k_i$, or $-$ to take a more specific example of the bank – "Income is high".

However, a procedure to generate such sentences in an unambiguous way should be designed, because, as we can see e.g. in Figure 1, a concept may have various features to various degrees. The proposed procedure is as follows: we will say that, by definition, for $x \in \mathcal{R}^+$, concept C_{i_0} is $\tilde{F}_{j_0}$ if $\max_{j=1,\dots,M} \mu^{i_0}_j(x)$=$\mu^{i_0}_{j_0}(x)$. If more than one j_0 with this property exist, the decision maker will be asked to choose one of them. Thanks to these assumptions, $\mu^{i_0}_{j_0}(x)$ selected in this way will always be positive and unambiguous, thus the statement "Concept C_{i_0} is $\tilde{F}_{j_0}$ " will always be to a certain degree justified and also unequivocal. Using the example given in Figure 1, the income 1.5 will be described as low or very low – this ambiguity will have to be resolved by an expert.

An additional assumption is that the features $F = \{\tilde{F}_j\}_{j=1}^{M}$ are ordered in the sense that either for each i the decision maker prefers C_i to be $\tilde{F}_{j+1}$ than to be $\tilde{F}_j$ for j=1,..M-1 or he or she prefers C_i to be $\tilde{F}_{j-1}$ than to be $\tilde{F}_j$ for j=0,..M) and that the corresponding membership functions are defined correctly. In the case of Figure1 the order follows clearly from the meaning of the concepts used as examples (financial security or yearly income) and the membership functions are "defined correctly" in the sense that for a fixed i, the values $\underline{a}^i_j, j = 1, \dots n$ (and analogously the other parameters defining the membership functions, i.e. $\widehat{a}^i_j, j = 1, \dots n$ and $\overline{a}^i_j, j = 1, \dots n$) defining the membership functions of $\tilde{F}_j$ for C_i are ordered in the sense of the usual ordering of real numbers.

Having defined fuzzy features, we can define fuzzy rules.

Definition 4 (author's own definition): A fuzzy rule R is a statement of the following form:
"If for each j=$1, \dots, m$, for at least s_j^R elements i from the set $\{1, \dots, n\}$ the statement 'Concept C_i is $\tilde{F}_j$ or more' is true, and for no more than t_j^R elements i from the set $\{1, \dots, n\}$ the statement 'Concept C_i is $\tilde{F}_j$ or less' is true, then take a specific decision" (implicitly: if any of the elements of the selected set is false, do not take any decision yet). The words "or more" and "or less" refer to the assumed order of $\{\tilde{F}_j\}_{j=1}^{M}$.

Let us present an example of a fuzzy rule from the banking sector, which might be used by commercial banks to make decisions about the credit risk of a potential borrower:

Example 1: Let us assume that three criteria are taken into account by the bank: the potential borrower's income, his/her financial security and the interest rate of the credit. Each of the criteria may be very low, low, medium high and very high. Then we might have the following decision rule: "If at least two criteria are medium or more and at least one criterion is high or more and no more than two criteria are low or less, then set the credit risk to low".

In the next section fuzzy numbers and fuzzy rules will be applied to the evaluation of research and development projects.

5 The proposed approach to R&D project success evaluation

In our opinion, it is important to facilitate the evaluation of R&D project success through grouping of the many possible evaluation criteria (presented in Section 3) into homogenous groups (for example, those presented in Section 3) and performing a separate evaluation of project success in each group. An optional aggregation will be performed later as a second step.

Hence, we consider t homogenous groups of R&D success evaluation criteria $\{\mathcal{G}_s\}_{s=1}^{t}$. For each $r = 1, \dots, w_s$ we have $\mathcal{G}_s = \{G_r^s\}_{r=1}^{w_s}$. We assume that, for each $s = 1, \dots, t$, $\mathcal{G}_s$ can be fully identified (i.e. all the assumptions are stated) with $\mathbb{C}$ from the previous section. We will use the following procedure (for a selected R&D project $\wp$):

I. SET s:=1;
II. SET $\mathbb{C} := \mathcal{G}_s$
III. For each G_r^s, $1, \dots, w_s$ find out for which $\tilde{F}_{j_r^s}$ the sentence "G_r^s is $\tilde{F}_{j_r^s}$" (j_r^s=*1,…,M*) is true according to the procedure described in the previous section;
IV. IF s=t THEN STOP, OTHERWISE SET s:=s+*1* and GO TO step II.

The outcome of this procedure is a set of statements of the form “For project $\wp$ G_r^s is $\tilde{F}_{j_r^s}$ ", for each $s = 1, \dots, t$ and for each criterion $\{G_r^s\}_{r=1}^{w_s}$.

As the next step consider, for each s, the criteria group $\mathcal{g}_s$ and the corresponding values $\tilde{F}_{j_r^s}$ and ask the experts which rules should be applied to obtain an aggregated evaluation of the fulfilment of the criteria from $\mathcal{g}_s$. The same features $F = \left\{\tilde{F}_j\ \right\}_{j=1}^{M}$ can be used here. Fuzzy rules should be applied here, for example made specific for this case as follows:

Rule R: “If for each j=$1, \dots, m$ for at least s_j^R elements r from the set $\{1, \dots, w_s\}$ the statement ‘G_r^s is $\tilde{F}_j$ or more’ is true and for no more than t_j^R elements r from the set $\{1, \dots, w_s\}$ the statement ‘G_r^s is $\tilde{F}_j$ or less’ is true then perform a certain action” (implicitly: if any of the elements of the selected set is false, do nothing). The words “or more” and “or less” refer to the assumed order of $\left\{\tilde{F}_j\ \right\}_{j=1}^{M}$.

The consistency of the rules should be verified by the experts, and metarules deciding which rule to apply if the rules lead to different conclusions should be formulated.

In this step the project $\wp$ will have been evaluated in the homogenous criteria groups $\mathcal{g}_s$, $s = 1, \dots, t$. The decision maker may stop here and not perform any further aggregation or he or she may decide to use fuzzy rules again in order to obtain an aggregated evaluation of the success of the project.

The proposed approach allows to aggregate evaluations of individual criteria from a homogenous groups into one evaluation for the group (and possibly later into an aggregated evaluation of the whole project success) in a simple (not requiring any complicated mathematical formulae and using expressions similar to those used in natural languages) and flexible form, in which a low satisfaction on some criteria will be compensated by a better behaviour on other criteria.

To sum up, we propose to evaluate the success of R&D projects on the basic of the following criteria tree, applied to each project:

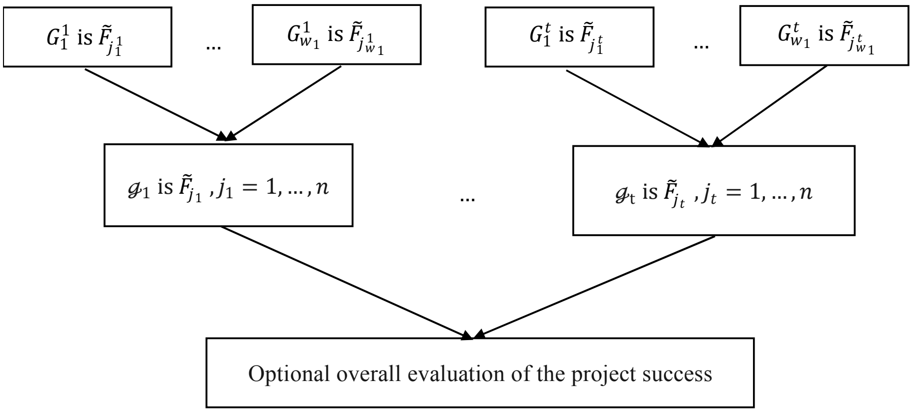


Figure 2: General scheme of R&D project evaluation

Source: Author's own elaboration.

The proposed approach will be illustrated by a case study.

6 Case study of an R&D project

The project entitled "Elaboration of a costing system for university X", of two-year duration, was funded by the Polish Centre of Research and Development. It was implemented in 2011-2012. It was managed by the present author. The objective of the project was to:

- Analyse the costing system used at university X at that time,
- Analyse managerial information provided by the existing system,
- Identify unsatisfied needs for managerial information,
- Identify necessary (existing and missing) sources of entry data for the system
- Elaborate a trial version of the system,
- Implement the system in two selected faculties of the university,
- Elaborate general indications for costing systems at universities.

Various factors contributed to the fact that most of the above objectives were satisfied only partially, to a low or even to a very low degree. Some of these factors were: concern about changes and additional work load among university administration employees, concern about revealing current financial procedures and information, disorder in the data at the university and low progress of data digitalisation. However, this project was considered by the team as a partial success, because other important objectives have been achieved:

- Identification of new research subjects,
- Identification of new opportunities for research cooperation,

- The project contributed to the achievement of PhD and ScD degrees,
- Building a well integrated research team,
- Satisfactory financial inflows for the project team members.

The decision makers of the university itself were not interested in the project at that time (mainly because of lack of time and because of numerous other challenges) and thus were rather indifferent as to its outcome. On the other hand, another, smaller university was selected as a partial substitute, where the above objectives were achieved to a higher degree. The financing institution accepted the results thanks to the continuous flow of information they were obtaining about the problems encountered in the project realisation and the introduction of a substitute university.

Thus, if we refer to the groups of success criteria described in section 3, we can say that the individual success criteria listed in section 3 were achieved to the degree given in Table 1. The evaluation in the last column was calculated using fuzzy rules, whose examples are given below in Table 3.

Table 3: Evaluation of the case study project

Criteria group	**A**	**b**	**c**	**d**	**Overall evaluation of success in the criteria group**
A	low	medium	irrelevant	low	low
B	very high	very high	high	-	very high
C	very high	medium	irrelevant	-	high
D	irrelevant	high	-	-	high
E	very high	very high	-	-	very high
F	high	very high	-	-	high
G	medium	medium	-	-	medium

Source: Author's own elaboration.

The following fuzzy rules were used here (among others):

For the criteria group A:

- If at least two of the criteria are high or more and no more than two criteria are low or less, the success in the whole group is high (this rule led to no conclusion in this case);
- If at least two criteria are low or less and at least one is medium or less, the success of the whole group is low.

For the criteria group B:

- If at least three of the criteria are very high or more and at least one criterion is high or more and no more than one criterion is medium or less, the success in the whole group is very high.

Similar rules were used for the other criteria groups. In the last step, the following rule was used to evaluate the success for the whole project:

- If at least two of the criteria are very high or more and at least one criterion is high or more and no more than three criteria are medium or less, the success of the project is high.

Thus, the experts, using their own rules, were able to judge the project success in a flexible way, expressing their opinion that the overall success of the project is high, even though the main research objectives were not achieved. Another expert group might have been of a different opinion. But it seems that in the case of research projects it is especially important to be flexible in their evaluation, because the understanding of their success is often ambiguous.

7 Conclusions

The paper proposes a flexible, easy to follow procedure for the assessment of R&D projects, which can be used for other project types as well. The procedure is based on expert opinion. The experts have to formulate rules which use language similar to natural languages. The other basis of our method are evaluation criteria of project success, which should be of various nature. This paper proposes a set of such criteria and refers to the literature where many more criteria can be found.

Setting criteria and rules is essential for the method to work. The method has to be constructed in detail and carefully tested in each specific context: in each organisation or group implementing R&D projects, possibly also in institutions funding such projects. This is because in the final analysis the success of a project has to be judged with respect to the strategy of individual institutions and groups. Thus the proposal formulated here should be regarded as a first step towards a complete and flexible system of evaluation of R&D project success.

The main method of further research should consist of case studies with active participation of various experts. In this paper we have discussed one case study, in which the role of experts was taken by the project manager and the project team. Undoubtedly, more exhaustive case studies are needed.

Acknowledgements

This research was supported by the National Science Centre (Poland), under Grant 394311: "Selected methods supporting project management, taking into consideration various stakeholder groups and using type-2 fuzzy numbers".

References

Ashley D.B., Laurie C.S., Jaselskis E.J. (1987), *Determinants of Construction Project Success*, Project Management Journal, 18(2), 69-79.

Baccarini D. (1999), *The Logical Framework Method for Defining Project Success*, Project Management Journal, 30(4), 25-32.

Baker B.N., Murphy D.C., Fisher D. (1988), *Factors Affecting Project Success*, [in:] D.I. Clenad, W.R. King (eds.), *Project Management Handbook*, Von Nostrand Reinhold, New York, NY.

Belassi W., Tukel O.I. (1996), *A New Framework for Determining Critical Success/failure in Projects*, International Journal of Project Management, 14(3), 141-151.

Camilleri E. (2011), *Project Success: Critical Factors and Behaviours*, Gower Publishing Company, Surrey, UK.

Chan A.P.C., Chan A.P.L. (2004), *Key Performance Indicators for Measuring Construction Success*, Benchmarking: An International Journal, 11(2), 203-221.

Cheng M.-Y., Tsai H.-C., Sudjono E. (2012), *Evolutionary Fuzzy Hybrid Neutral Network for Dynamic Project Assessment in Construction Industry*, Automation in Construction, 21, 46-51.

Chow T., Cao D.-B. (2008), *A Survey Study for Critical Success Factors in Agile Software Projects*, Journal of Systems and Software, 81, 961-971.

Davis K. (2014), *Different Stakeholder Groups and Their Perceptions of Project Success*, International Journal of Project Management, 32(2), 189-201.

Eilat H., Golany B., Shtub A. (2008), *R&D Project Evaluation: An Integrated DEA and Balanced Scorecard Approach*, Omega, 36(5), 895-912.

Elkadi H. (2013), *Success and Failure Factors for E-government Projects: A Case from Egypt*, Egyptian Informatics Journal, 14(2), 165-173.

Freeman M., Beale P. (1992), *Measuring Project Success*, Project Management Journal, 23 (1), 8-17.

Frese R., Sauter V. (2003), *Project Success and Failure: What is Success, what is Failure, and How Can You Improve Your Odds for Success*, [Online] cited after Camilleri (2011).

Greer M. (1999), *Handbook of Human Performance Technology*, CA: Jossey-Bass, San Francisco.

Kerzner H. (1992), *Project Management: A Systems Approach to Planning, Scheduling and Controlling*, Van Nostrand Reinhold, New York, NY.

Khan K., Turner J.R., Maqsood T. (2013), *Factors that Influence the Success of Public Sector Projects in Pakistan. Proceedings of IRNOP 2013 Conference, June 17-19, 2013*, BI Norwegian Business School, Oslo, Norway.

Klaus-Rosińska A. (2019), *Sukces projektów badawczych i badawczo-rozwojowych*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław [in print].

Korol T. (2012), *Fuzzy Logic in Financial Management*, https://www.intechopen.com/profiles/105028/tomasz-korol

Kuchta D., Betta J., Jastrzębska J., Frączkowski K., Gładysz B., Prałat E., Ptaszyńska E., Rola P., Walecka-Jankowska K., Ropuszyńska-Surma E., Skomra A., Ryńca R., Klaus-Rosińska A., Mrzygłocka-Chojnacka J. (2017), *Success and Failure Factors of R&D Projects at Universities in Poland and France*, [in:] A. Nowak, Z. Wilimowska (eds.), *Business Risk in Changing Dynamics of Global Village: Proceedings of 1st International Conference on Business Risk in Changing Dynamics of Global Village, 26-27 April 2017*, Publishing Office, University of Applied Sciences, Nysa, 253-278.

Lim C.S., Mohamed M.Z. (1999), *Criteria of Project Success*, International Journal of Project Management, 17(4), 243-248.

Mir F.A., Pinnington A.H. (2013), *Exploring the Value of Project Management: Linking Project Management Performance and Project Success*, [to be published in International Journal of Project Management].

Nahod M.-M., Vukomanovic M., Radujkovic M. (2013), *The Impact of ICB 3.0 Competences on Project Management Success. In 26th IPMA World Congress, Crete, Greece, 2012*, Procedia: Social and Behavioral Sciences, 74, 244-254.

Pinto J.K., Slevin D.P. (1988), *Critical Success Factors in Successful Project Implementation*, IEEE Transactions on Engineering Management, 34(1), 22-27.

Raz T., Shenhar A.J., Dvir D. (2002), *Risk Management*, *Project Success and Technological Uncertainty*, R&D Management, 32(2), 101-109.

Revilla E., Sarkis J., Modrego A. (2003), *Evaluating Performance of Public-private Research Collaborations: A DEA Analysis*, Journal of the Operational Research Society, 54(2), 165-174.

Shashi K.N.R., Nai T.R.G., Suma V. (2014), *SLI: A New Metric to Determine Success of a Software Project. 2014 International Conference on Electronics and Communication Systems*, ICECS 2014.

Shenhar A.J., Dvir D., Levy O. (1997), *Mapping the Dimensions of Project Success*, Project Management Journal, 28(2), 5-13.

Sutton B. (2005). *Why Projects Fail: Mastering the Monster*, cited after Camilleri (2011).

Thomas G., Fernández W. (2008), *Success in IT Projects: A Matter of Definition?*, International Journal of Project Management, 26(7), 733-742.

Thomsett R. (2002), *Project Pathology, Causes, Symptoms and Patterns of Project Failure*, cited after Camilleri (2011).

Turner J.R. (1994), *Editorial: Project Management Future Developments for the Short and Medium Term*, International Journal of Project Management, 12(1), 3-4.

Wateridge J. (1998), *How Can IS/IT Projects be Measured for Success?*, International Journal of Project Management, 16(1), 59-63.

Westerveld E. (2003), *The Project Excellence Model: Linking Success Criteria and Critical Success Factors*, International Journal of Project Management, 21, 411-418.

de Wit A. (1988), *Measurement of Project Success*, International Journal of Project Management, 6(3), 164-170.

Yourdon E. (1997), *Death March: The Complete Software Developer's Guide to Surviving Mission Impossible Projects*, NJ: Prentice Hall, Upper Saddle River.

Zadeh L. (1965), *Fuzzy Sets*, Information and Control, 8(3), 338-353.

M U L T I P L E   C R I T E R I A   D E C I S I O N   M A K I N G
Vol. 14 2019

Somdeb Lahiri*

THE NO-SPOILER CONDITION FOR CHOICE CORRESPONDENCES

DOI: 10.22367/mcdm.2019.14.04

Abstract

We show that any choice correspondence which satisfies the weak Pareto criterion and the Majority property must violate the no-spoiler condition. Subsequently we strengthen the weak Pareto criterion. We show that if the number of criteria or individuals or states of nature is odd, then there is no choice correspondence which satisfies this strengthened version of weak Pareto criteria, Majority property and no-loser spoiler condition. However if the number of criteria/individuals/states of nature is even, we need two more properties to ensure the impossibility result. The first of these two properties is top neutrality. The second property is top anonymity.

Keywords: choice correspondence, no-spoiler, no-loser spoiler, majority property.

1 Introduction

In group (or multi-criteria) choice theory, one is concerned with choosing a non-empty set of alternatives from a given set of alternatives for each profile of preferences. Each profile of preferences could either represent the preferences of individual voters (as in group decision theory) or rankings along criteria considered by a single decision maker as in multi-criteria decision making or preferences dependent on the state of nature, once again of a single individual. The multi-criteria decision making interpretation has been nicely motivated by Rubinstein (2012). Similarly, the state dependent preferences interpretation derives its relevance from the stand point of decision making under complete uncertainty (see Lahiri, 2019a, 2019c). In this paper we use the term choice correspondence to describe such procedures concerned with aggregating

* School of Petroleum Management, Pandit Deendayal Petroleum University, Gandhinagar, India, e-mail: somdeb.lahiri@gmail.com, ORCID: 0000-0002-5247-3497.

a collection of preferences into a chosen set of outcomes, whether those separate preferences represent the preferences of distinct individuals or rankings along distinct criteria or preferences which depend on the state of nature. Our framework in which profiles of preferences are mapped to chosen outcomes, and which has been discussed for instance in Denicolo (1993) is a generalization of the one normally associated with the work (and the huge literature that has grown out of it) due to Alan Gibbard and Mark Satterthwaite (Gibbard, 1973; Satterthwaite, 1975). In such a situation a choice correspondence has been traditionally referred to as a social choice correspondence.

In an interesting book by Christoph Borgers (2010), a property referred to as no-spoiler condition is introduced, which says the following. Given a preference profile and an alternative x if: (a) x is not the uniquely chosen alternative; and (b) x is removed from the set of available alternatives *then* the new set of chosen alternatives consists of those that were chosen earlier excluding x. A second interesting property introduced subsequently in Borgers (2010) is referred to as the no-loser spoiler condition and says the following. If an un-chosen alternative is removed from the set of available alternatives, then the set of chosen alternatives remains the same as before. It is easy to see from the two definitions that both require the set of available alternatives to be variable – a possibility considered by Kenneth Arrow while presenting his impossibility theorem and discussed in detail in the book by Kelly (1988). In the conventional context in which choice correspondences are discussed, the set of available alternatives is considered to be fixed and we intend to adopt the conventional approach in this paper. So, if we want to use anything like the two properties mentioned above, we will have to adjust the definitions so that they are meaningful in our context. That is precisely what we do here.

Thus, we define the no-spoiler condition as follows. Given two profiles of preferences and an alternative x if: (a) the only difference between the two profiles is that at the second profile every individual/criteria ranks x at the bottom but is otherwise the same as the ranking for the corresponding individual/along the criteria in the first profile; and (b) x is not the uniquely chosen alternative at the first profile, *then* the set of chosen alternatives at the second profile is simply the set of chosen alternatives at the first profile other than x. Similarly we define the no-loser spoiler condition in the following manner. Given two profiles of preferences and an alternative x which is un-chosen at the first profile, if the only difference between the two profiles is that at the second profile every individual/criteria ranks x at the bottom but is otherwise the same as the ranking for the corresponding individual/criteria in the first profile, *then* the set of chosen alternatives at the second profile is the same as the set of chosen alternatives at the first profile.

In this paper we show that any choice correspondence which satisfies the weak Pareto criterion and the Majority property must violate the no-spoiler condition. The Majority property is much weaker than a common assumption in the literature called Condorcet property. The Majority property says that if an alternative is ranked uniquely first by at least half of the total number of individuals/criteria and is ranked first (perhaps not uniquely) by more than half of the total number of individuals/criteria, then the alternative in question is the unique alternative that is chosen. A matter of some concern about the majority property has been voiced by Professor Prasanta Pattanaik in a private communication dated April 27th, 2018, which is being quoted below.

"While this (*interpretation of aggregating multiple attributes into a non--empty set of outcomes*: author) is formally right, the Condorcet property and the Majority property (the appeal of which depends significantly on the underlying intuition of anonymity), are far less persuasive when we deal with a single person's decision making on the basis of multiple criteria: there is no compelling reason why an individual should attach the same weight to each criterion in taking decisions on the basis of multiple criteria. This is so even when we confine ourselves exclusively to orderings in terms of each criterion without taking into account any intensity".

A partial response to Professor Pattanaik's concerns, in the form of empirical evidence in favor of majority rule in individual decision making, can be found in Zhang, Hsee and Xiao (2006).

Subsequently we strengthen the weak Pareto criterion to require that if along all but one criteria (or for all but one individual) x is ranked above y then y is not chosen. We show that if the number of individuals/criteria is odd, then there is no choice correspondence which satisfies this strengthened version of weak Pareto criteria, Majority property and no-loser spoiler condition. However if the number of individuals/criteria is even, we need two more properties to ensure the impossibility result. The first of these two properties is top neutrality. Top neutrality says that if the top two alternatives in all the rankings are x and y, then interchanging the rankings of x and y for every individual or criterion leads to a reversal of the roles of x and y from the perspective of group (or multi-criteria) choice. The second property is top anonymity. Top anonymity says that if the top two alternatives in all the rankings are x and y, then interchanging the names of the individuals/criteria, does not change the outcome of choice for x and y.

The significant use of the majority property in this paper was observed by a mathematician Professor Janez Zerovnik who suggested an alternative conceptualization which assuming his consent we refer to as Point Majority property. What Point Majority property requires is to assign a score of one to the first rank and zero to all other ranks for each ranking of the set of alternatives. If there is a tie for the first rank, then this one unit that is assigned to this rank is shared equally among the alternatives that tie for first rank. The Point Majority property states that if at a profile of rankings there is an alternative whose sum of points acquired over all criteria exceeds half of the total number of criteria, then this alternative is the uniquely chosen alternative. An easy calculation shows that given a profile of rankings, if such an alternative exists, then it must be unique. Further, we are able to show that the Point Majority property implies the Majority property, though on arbitrary domains the converse need not necessarily be true. However, for domains comprising profiles of strict rankings, the two properties − Majority Property and the Point Majority property − are equivalent.

In the rest of the paper instead of referring to "criteria" or "individual/ criteria", we will simply use the word "individual".

2 The model and some assumptions

Let X denote a non-empty finite set of alternatives containing at least three alternatives and $N = \{1,\ldots,n\}$ for some positive integer $n \geq 3$ denote the set of agents. The set of all non-empty subsets of X is denoted by $\Psi(X)$.

In what follows we shall be concerned with binary relations on X.

Given a binary relation R on X and $x,y \in X$, we will denote $(x,y) \in R$ by xRy. Given a binary relation R on X and $A \in \Psi(X)$, let $R|A$ denote the restriction of R to A, i.e. $x(R|A)y$ if and only if $x,y \in A$ and xRy.

Let $W(X)$ denote the set of weak orders on X (i.e. the set of reflexive, complete and transitive binary relations) and $L(X)$ the set of linear orders on X (anti-symmetric weak orders on X).

Given $R \in W(X)$, let $P(R)$ denote its asymmetric part and $I(R)$ denote its symmetric part. When there is no scope for confusion, we may use P instead of $P(R)$ and I instead of $I(R)$.

Further, given $R \in W(X)$ and $A \in \Psi(X)$, let $G(A,R) = \{x \in A | xRy \text{ for all } y \in A\}$. $G(A,R)$ is called the set of **best** or **greatest** elements/alternatives with respect to R in A. It is well known that for an arbitrary $R \in W(X)$, $G(A,R)$ is non-empty.

An element of $[W(X)]^N$ is called a **profile**.

Any non-empty subset Ω of $[W(X)]^N$ is called a **domain**.

Given a domain Ω a **choice correspondence** (CC) on Ω is a function $f:\Omega \rightarrow \Psi(X)$.

An CC f on Ω is said to satisfy **Weak Pareto** if for all $(R_1,\ldots,R_n) \in \Omega$ and $x,y \in X$, $[xP(R_j)y]$ for all $j = 1,\ldots,n$ implies $y \notin f(R_1,\ldots,R_n)$.

The following property which is a strengthened version of weak Pareto can be found in Pattanaik and Lahiri (2017). It was suggested to the author by Professor Prasanta Pattanaik in a private communication dated September 5th, 2016.

A CC f on Ω is said to satisfy **Strengthened Weak Pareto** if for all $(R_1,\ldots,R_n)\in\Omega$ and $x,y\in X$, $|\{k\in N|xP(R_k)y\}| \geq n-1$ implies $y\notin f(R_1,\ldots,R_n)$.

A CC f with domain Ω is said to satisfy **Condorcet Property** if for all profiles $(R_1,\ldots, R_n)\in\Omega$, the following holds: if there exists $x\in X$ such that $\#\{j\in N| xP_jy\} > \#\{j\in N| yP_jx\}$ for all $y\in X\backslash\{x\}$, then $f(R_1,\ldots,R_n) = \{x\}$.

Condorcet property says that if at a profile the number of individuals which rank an alternative x above an alternative y is greater than the number of individuals which rank y above x where y is any alternative other than x, then x is the unique alternative to be chosen by the CC at the profile.

A far weaker condition than Condorcet Property is the following which to the best of our knowledge has been introduced in a paper by Lahiri (2019b).

A CC f with domain Ω is said to satisfy **Majority Property** if for all profiles $(R_1,\ldots, R_n)\in\Omega$, the following holds: if there exists $x\in X$ such that $\#\{j\in N|G(X,R_j) = \{x\}\} \geq \frac{n}{2}$ and $\#\{j\in N|x\in G(X,R_j)\} > \frac{n}{2}$ then $f(R_1,\ldots,R_n) = \{x\}$.

Majority property says that if at a profile x is ranked uniquely first by at least half of the total number of individuals and is ranked first (perhaps not uniquely) by more than half of the total number of individuals then x is the unique alternative to be chosen by the CC at the profile.

First note if an x exists satisfying the requirements of the majority property then it has got to be unique. For if $y\in X\backslash\{x\}$, then y can be uniquely first less than $(n- \frac{n}{2}) = \frac{n}{2}$ times. Hence y cannot satisfy the requirements of the majority property.

It is easy to see that Condorcet property implies the majority property although the converse need not be true.

To see that the Condorcet property implies the Majority Property let $x\in X$ such that $\#\{j\in N|G(X,R_j) = \{x\}\} \geq \frac{n}{2}$ and $\#\{j\in N|x\in G(X,R_j)\} > \frac{n}{2}$. Let $y\in X\backslash\{x\}$. Thus, $\#\{j\in N|xR_jy\} > \frac{n}{2}$ and $\#\{j\in N|yP(R_j)x\} < n- \frac{n}{2} = \frac{n}{2}$ Hence by Condorcet property, $f(R_1,\ldots,R_n) = \{x\}$.

That the converse need not be true is shown in the following example.

Example 1

Let $X = \{x,y,z\}$ and $n = 3$. Define the CC f as follows: for all $(R_1,R_2,R_3)\in [W(X)]^N$ and $v\in X$ let $v\in f(R_1,R_2,R_3)$ if and only if for all $w\in X$, $\#\{j|v\in G(X,R_j)\} \geq \#\{j|w\in G(X,R_j)\}$. It is easy to see that f satisfies Majority property. However, f does not satisfy Condorcet property. Let $xP(R_1)yP(R_1)z$, $yP(R_2)xP(R_2)z$ and $zP(R_3)yP(R_3)x$. Thus, $f(R_1,R_3,R_n) = \{x,y,z\}$ although $\#\{j\in N| yP(R_j)w\} > \#\{j\in N| wP(R_j)y\}$ for all $w\in X\backslash\{y\}$, whence according to the Condorcet property $f(R_1,R_n, R_3)$ should have been equal to $\{y\}$.

As we shall see in the next section, there is no CC satisfying strengthened weak Pareto if n is equal to three.

The following two conditions are based on similar properties due to Borges (2010) for group choice functions.

An CC f with domain Ω is said to satisfy the **No-spoiler condition** if for all $(R_1,\ldots, R_n)$, $(T_1,\ldots,T_n)\in\Omega$ and $x\in X$ with $f(R_1,\ldots,R_n) \neq \{x\}$, the following is true: $f(T_1,\ldots,T_n) = f(R_1,\ldots,R_n)\backslash\{x\}$, where for all $k\in N$ and $y\in X\backslash\{x\}$, yP_kx and $T_k|X\backslash\{x\} = R_k|X\backslash\{x\}$.

An CC f with domain Ω is said to satisfy the **No-loser spoiler condition** if for all $(R_1,\ldots, R_n)$, $(T_1,\ldots,T_n)\in\Omega$ and $x\in X$ with $x\notin f(R_1,\ldots,R_n)$, the following is true: $f(T_1,\ldots,T_n) = f(R_1,\ldots,R_n)$, where for all $k\in N$ and $y\in X\backslash\{x\}$, yP_kx and $T_k|X\backslash\{x\} = R_k|X\backslash\{x\}$.

Note: It was pointed out to me by Professor Prasanta Pattanaik, that the no-spoiler condition (in fact the no-loser spoiler condition) was indicated in an interesting discussion immediately after Condition 3 (now better known as Arrow's Independence of Irrelevant Alternatives), in Arrow (1950,1963). The discussion pointed to the intuitive plausibility of the no-loser spoiler condition as defined by Borgers (2010) and provided an example where this condition is convincingly violated by an equally intuitively plausible aggregation rule (i.e. the Borda rule).

To establish the incompatibility of the No-loser spoiler condition with Strengthened Weak Pareto and Majority Property when the number of individuals is even and greater than two in Proposition 5, we will require these two additional properties.

An CC f with domain Ω is said to satisfy **Top Neutrality** if for $(R_1,\ldots,R_n)$, $(S_1,\ldots,S_n)\in\Omega\cap[L(X)]^N$ and $x,y\in X$ with $x \neq y$, (i) implies (ii), where:
(i) for all $j\in N$: either $[xP(R_j)yP(R_j)w$, $yP(S_j)xP(S_j)w$ for all $w\in X\backslash\{x,y\}$ and $R_j|(X\backslash\{x,y\}) = S_j|(X\backslash\{x,y\})]$ **or** $[yP(R_j)xP(R_j)w$, $xP(S_j)yP(S_j)w$ for all $w\in X\backslash\{x,y\}$ and $R_j|(X\backslash\{x,y\}) = S_j|(X\backslash\{x,y\})]$.
(ii) $\{x\} = f(R_1,\ldots,R_n)$ if and only if $\{y\}=f(S_1,\ldots,S_n)$; $\{y\}= f(R_1,\ldots,R_n)$ if and only if $\{x\}=f(S_1,\ldots,S_n)$.

An CC f with domain Ω is said to satisfy **Top Anonymity** if for $(R_1,\ldots,R_n)$, $(S_1,\ldots,S_n)\in\Omega\cap[L(X)]^N$ and $x,y\in X$ with $x \neq y$, (i) implies (ii), where:
(i) there exists a permutation π on N (i.e. one-to-one function from N to N) such that for all $j\in N$, $S_j = T_{\pi(j)}$ and for all $j\in N$:either $[xP(R_j)yP(R_j)w$ for all $w\in X\backslash\{x,y\}]$ **or** $[yP(R_j)xP(R_j)w$ for all $w\in X\backslash\{x,y\}]$.
(ii) $x\in f(R_1,\ldots,R_n)$ if and only if $x\in f(S_1,\ldots,S_n)$; $y\in f(R_1,\ldots,R_n)$ if and only if $y\in f(R_1,\ldots,R_n)$.

3 The main results of this paper

In this section we present the main results of this paper.

Proposition 1
If n is an odd integer and $[L(X)]^N \subset \Omega$, then there does not exist any CC on Ω which satisfies Weak Pareto, Majority property and no-spoiler condition.

Proof
The proof proceeds by showing that if n is an odd integer and if a CC satisfies Weak Pareto and Majority Property, then the CC must violate the no-spoiler condition.

Suppose n is an odd integer. If n is an odd integer greater than three, $\frac{n-3}{2}$ is a positive integer. Let x, y, z be three distinct elements of X and let $(R_1,\ldots,R_n) \in [L(X)]^N$ such that:
(i) $xP(R_1)yP(R_1)zP(R_1)w$ for all $w \in X\backslash\{x,y,z\}$,
(ii) $zP(R_2)xP(R_2)yP(R_2)w$ for all $w \in X\backslash\{x,y,z\}$,
(iii) $yP(R_3)zP(R_3)xP(R_3)w$ for all $w \in X\backslash\{x,y,z\}$,
(iv) $xP(R_k)yP(R_k)zP(R_k)w$ for all $w \in X\backslash\{x,y,z\}$ and $k = 4,\ldots, 3+\frac{n-3}{2}$,
(v) $zP(R_k)yP(R_k)xP(R_k)w$ for all $w \in X\backslash\{x,y,z\}$ and $k = 4+\frac{n-3}{2},\ldots,n$.

By weak Pareto for all $w \in X\backslash\{x,y,z\}$ $w \notin f(R_1,\ldots,R_n)$ and so $f(R_1,\ldots,R_n) \subset \{x, y, z\}$.

Case 1: $f(R_1,\ldots,R_n) = \{x, y, z\}$.

Let $(T_1,\ldots,T_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X\backslash\{z\}$, we have $wP(T_k)z$ and $T_k|X\backslash\{z\} = R_k|X\backslash\{z\}$.

Since at $(T_1,\ldots,T_n)$ x is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(T_1,\ldots,T_n) = \{x\} \neq f(R_1,\ldots,R_n)\backslash\{z\}$, violating the no-spoiler condition.

Case 2: $f(R_1,\ldots,R_n)$ is a two element subset of $\{x, y, z\}$.
(i) Suppose $f(R_1,\ldots,R_n) = \{x, y\}$.

Let $(S_1,\ldots,S_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X\backslash\{z\}$, we have $wP(S_k)z$ and $S_k|X\backslash\{z\} = R_k|X\backslash\{z\}$.

Since at $(S_1,\ldots,S_n)$ x is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(S_1,\ldots,S_n) = \{x\} \neq f(R_1,\ldots,R_n)\backslash\{z\}$, violating the no-spoiler condition.
(ii) Suppose $f(R_1,\ldots,R_n) = \{x, z\}$.

Let $(S_1,\ldots,S_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X\backslash\{y\}$, we have $wP(S_k)y$ and $S_k|X\backslash\{y\} = R_k|X\backslash\{y\}$.

Since at $(S_1,\ldots,S_n)$ z is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(S_1,\ldots,S_n) = \{z\}\neq f(R_1,\ldots,R_n)\backslash\{y\}$, violating the no-spoiler condition.

(iii) Suppose $f(R_1,\ldots,R_n) = \{y, z\}$.

Let $(S_1,\ldots,S_n) \in [L(X)]^N$ such that for all $k\in N$ and $w\in X\backslash\{x\}$, we have $wP(S_k)x$ and $S_k|X\backslash\{x\}= R_k|X\backslash\{x\}$.

Since at $(S_1,\ldots,S_n)$ y is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(S_1,\ldots,S_n) = \{y\}\neq f(R_1,\ldots,R_n)\backslash\{x\}$, violating the no-spoiler condition.

Case 3: $f(R_1,\ldots,R_n)$ is a singleton subset of $\{x, y, z\}$.

(i) $f(R_1,\ldots,R_n) = \{x\}$.

Let $(U_1,\ldots,U_n) \in [L(X)]^N$ such that for all $k\in N$ and $w\in X\backslash\{y\}$, we have $wP(U_k)y$ and $U_k|X\backslash\{y\}= R_k|X\backslash\{y\}$.

Since at $(U_1,\ldots,U_n)$ z is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(U_1,\ldots,U_n) = \{z\}\neq f(R_1,\ldots,R_n)\backslash\{y\}$, violating the no-spoiler condition.

(ii) $f(R_1,\ldots,R_n) = \{y\}$.

Let $(U_1,\ldots,U_n) \in [L(X)]^N$ such that for all $k\in N$ and $w\in X\backslash\{z\}$, we have $wP(U_k)z$ and $U_k|X\backslash\{z\}= R_k|X\backslash\{z\}$.

Since at $(U_1,\ldots,U_n)$ x is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(U_1,\ldots,U_n) = \{x\}\neq f(R_1,\ldots,R_n)\backslash\{z\}$, violating the no-spoiler condition.

(iii) $f(R_1,\ldots,R_n) = \{z\}$.

Let $(U_1,\ldots,U_n) \in [L(X)]^N$ such that for all $k\in N$ and $w\in X\backslash\{x\}$, we have $wP(U_k)x$ and $U_k|X\backslash\{x\}= R_k|X\backslash\{x\}$.

Since at $(U_1,\ldots,U_n)$ y is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(U_1,\ldots,U_n) = \{y\}\neq f(R_1,\ldots,R_n)\backslash\{x\}$, violating the no-spoiler condition.

Thus $f(R_1,\ldots,R_n) = \phi$, contrary to the definition of an CC.

This proves the proposition. Q.E.D.

Proposition 2

If n is even and $\Omega = [W(X)]^N$, then there does not exist any CC on Ω which satisfies Weak Pareto, Majority property and no-spoiler condition.

Proof

Suppose n is even. Then $n \geq 4$ and if n is an even integer greater than four, $\frac{n-4}{2}$ is a positive integer. Let x,y,z be three distinct elements of X and let $(R_1,\ldots,R_n)\in [L(X)]^N$ such that:

(i) $xP(R_1)yP(R_1)zP(R_1)w$ for all $w \in X \setminus \{x, y, z\}$,
(ii) $zP(R_2)xP(R_2)yP(R_2)w$ for all $w \in X \setminus \{x, y, z\}$,
(iii) $yP(R_3)zP(R_3)xP(R_3)w$ for all $w \in X \setminus \{x, y, z\}$,
(iv) $xI(R_4)yI(R_4)zP(R_4)w$ for all $w \in X \setminus \{x, y, z\}$,
(v) $xP(R_k)yP(R_k)zP(R_k)w$ for all $w \in X \setminus \{x, y, z\}$ and $k = 5, \ldots, 4+\frac{n-3}{2}$,
(vi) $zP(R_k)yP(R_k)xP(R_k)w$ for all $w \in X \setminus \{x, y, z\}$ and $k = 5+\frac{n-3}{2}, \ldots, n$.

By weak Pareto $w \in X \setminus \{x, y, z\}$ implies $w \notin f(R_1, \ldots, R_n)$, so that $f(R_1, \ldots, R_n) \subset \{x, y, z\}$.

The rest of the proof is very similar to the proof of proposition 1. Q.E.D.

Proposition 3

If $[L(X)]^N \subset \Omega$ and $n = 3$, then there does not exist any CC on Ω which satisfies Strengthened Weak Pareto.

Proof

Let x, y, z be three distinct elements of X and let $(R_1, \ldots, R_n) \in [L(X)]^N$ such that:
(i) $xP(R_1)yP(R_1)zP(R_1)w$ for all $w \in X \setminus \{x, y, z\}$,
(ii) $zP(R_2)xP(R_2)yP(R_2)w$ for all $w \in X \setminus \{x, y, z\}$,
(iii) $yP(R_3)zP(R_3)xP(R_3)w$ for all $w \in X \setminus \{x, y, z\}$.
By strengthened weak Pareto $f(R_1, \ldots, R_n) \subset \{x, y, z\}$. Further,
(i) $\#\{j \in N | \, xP(R_j)y\} = 2 = n-1$ implies $y \notin f(R_1, \ldots, R_n)$;
(ii) $\#\{j \in N | yP(R_j)z\} = 2 = n-1$ implies $z \notin f(R_1, \ldots, R_n)$;
(iii) $\#\{j \in N | zP(R_j)x\} = 2 = n-1$ implies $x \notin f(R_1, \ldots, R_n)$.

Thus $f(R_1, \ldots, R_n) = \phi$, which is not possible by the definition of an CC.

This proves the proposition. Q.E.D.

Proposition 4

If n is odd and $[L(X)]^N \subset \Omega$, then there does not exist any CC on Ω which satisfies Strengthened Weak Pareto, Majority property and No-loser spoiler condition.

Proof

Suppose n is odd. By proposition 3, this proposition is definitely satisfied if n is equal to three. Hence suppose n is greater than three. Since n is an odd integer greater than three, $\frac{n-3}{2}$ is a positive integer.

Let x, y, z be three distinct elements of X and let $(R_1, \ldots, R_n) \in [L(X)]^N$ such that:
(i) $xP(R_1)yP(R_1)zP(R_1)w$ for all $w \in X \setminus \{x, y, z\}$,
(ii) $zP(R_2)xP(R_2)yP(R_2)w$ for all $w \in X \setminus \{x, y, z\}$,
(iii) $yP(R_3)zP(R_3)xP(R_3)w$ for all $w \in X \setminus \{x, y, z\}$,
(iv) $xP(R_k)yP(R_k)zP(R_k)w$ for all $w \in X \setminus \{x, y, z\}$ and $k = 4, \ldots, 3+\frac{n-3}{2}$,
(v) $zP(R_k)xP(R_k)yP(R_k)w$ for all $w \in X \setminus \{x, y, z\}$ and $k = 4+\frac{n-3}{2}, \ldots, n$.

By strengthened weak Pareto $w \in X \backslash \{x,y,z\}$ implies $w \notin f(R_1,\ldots,R_n)$ so that $f(R_1,\ldots,R_n) \subset \{x, y, z\}$.

By strengthened weak Pareto $\#\{j \in N | \, xP(R_j)y\} = n-1$ implies $y \notin f(R_1,\ldots,R_n)$.

Thus, $f(R_1,\ldots,R_n) \subset \{x, z\}$.

Suppose $x \in f(R_1,\ldots,R_n)$ and let $(T_1,\ldots,T_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X \backslash \{y\}$, we have $wP(T_k)y$ and $T_k|X \backslash \{y\} = R_k|X \backslash \{y\}$.

Since at $(T_1,\ldots,T_n)$ z is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(T_1,\ldots,T_n) = \{z\} \neq f(R_1,\ldots,R_n) \backslash \{y\}$, violating the no-loser spoiler condition.

Thus $f(R_1,\ldots,R_n) = \{z\}$ and $x \notin f(R_1,\ldots,R_n)$.

Let $(S_1,\ldots,S_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X \backslash \{x\}$, we have $wP(S_k)x$ and $S_k|X \backslash \{x\} = R_k|X \backslash \{x\}$.

Since at $(S_1,\ldots,S_n)$ y is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(S_1,\ldots,S_n) = \{y\} \neq f(R_1,\ldots,R_n) \backslash \{x\}$, violating the no-loser spoiler condition.

Thus, $f(R_1,\ldots,R_n) \neq \{z\}$ and so $f(R_1,\ldots,R_n) = \phi$, contrary to the definition of an CC.

This proves the proposition. Q.E.D.

The next proposition shows the incompatibility of No-loser spoiler condition with Strengthened Weak Pareto, Majority property, Top Anonymity and Top Neutrality, when the number of individuals is even and greater than two.

Proposition 5

If n is an even natural number greater than two and $\Omega = [W(X)]^N$, then there does not exist any CC on Ω which satisfies Strengthened Weak Pareto, Majority property, Top Anonymity, Top Neutrality and No-loser spoiler condition.

Proof

Suppose n is an even natural number greater than 2. Thus $n \geq 4$ and $\frac{n-4}{2}$ is a non--negative integer.

Let x, y, z be three distinct elements of X and let $(R_1,\ldots,R_n) \in [W(X)]^N$ such that:

(i) $xP(R_1)yP(R_1)zP(R_1)w$ for all $w \in X \backslash \{x, y, z\}$,

(ii) $zP(R_2)xP(R_2)yP(R_2)w$ for all $w \in X \backslash \{x, y, z\}$,

(iii) $yP(R_3)zP(R_3)xP(R_3)w$ for all $w \in X \backslash \{x, y, z\}$,

(iv) $xI(R_4)zP(R_4)yP(R_4)w$ for all $w \in X \backslash \{x, y, z\}$,

and if $n > 4$ then

(v) $xP(R_k)yP(R_k)zP(R_k)w$ for all $w \in X \backslash \{x, y, z\}$ and $k = 5,\ldots, 4+\frac{n-4}{2}$,

(vi) $zP(R_k)xP(R_k)yP(R_k)w$ for all $w \in X \backslash \{x, y, z\}$ and $k = 5+\frac{n-4}{2},\ldots,n$,

(vii) $R_j|X \backslash \{x, y, z\} = R_k|(X \backslash \{x, y, z\})$ for all $j,k \in N$.

By strengthened weak Pareto for all $w \in X \backslash \{x, y, z\}$, $w \notin f(R_1,\ldots,R_n)$, so that $f(R_1,\ldots,R_n) \subset \{x, y, z\}$.

By strengthened weak Pareto we have $y \notin f(R_1,\ldots,R_n)$, so that $f(R_1,\ldots,R_n) \subset \{x,z\}$.

Suppose $x \in f(R_1,\ldots,R_n)$ and let $(T_1,\ldots,T_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X \backslash \{y\}$, we have $wP(T_k)y$ and $T_k|X\backslash\{y\} = R_k|X\backslash\{y\}$.

Since at $(T_1,\ldots,T_n)$ z is ranked uniquely first by more than half of the total number of individuals, by the Majority Property $f(T_1,\ldots,T_n) = \{z\} \neq f(R_1,\ldots,R_n)\backslash\{y\}$, violating the no-loser spoiler condition.

Thus $f(R_1,\ldots,R_n) = \{z\}$ i.e. $x \notin f(R_1,\ldots,R_n)$.

Let $(S_1,\ldots,S_n) \in [L(X)]^N$ such that for all $k \in N$ and $w \in X \backslash \{x\}$, we have $wP(S_k)x$ and $S_k|X\backslash\{x\} = R_k|X\backslash\{x\}$.

Thus,

(i) $yP(S_1)zP(S_1)wP(S_1)x$ for all $w \in X\backslash\{x, y, z\}$,

(ii) $zP(S_2)yP(S_2)wP(S_2)x$ for all $w \in X\backslash\{x, y, z\}$,

(iii) $yP(S_3)zP(S_3)wP(S_3)x$ for all $w \in X\backslash\{x, y, z\}$,

(iv) $zP(S_4)yP(S_4)wP(S_4)x$ for all $w \in X\backslash\{x, y, z\}$,

and if $n > 4$ then

(v) $yP(S_k)zP(S_k)wP(S_k)x$ for all $w \in X\backslash\{x, y, z\}$ and $k = 5,\ldots, 4+\frac{n-4}{2}$,

(vi) $zP(S_k)yP(S_k)wP(S_k)x$ for all $w \in X\backslash\{x, y, z\}$ and $k = 5+\frac{n-4}{2},\ldots,n$.

By no-loser spoiler condition, $f(S_1,\ldots,S_n) = \{z\}$.

Let $(U_1,\ldots,U_n) \in \Omega$ be such that $U_1 = S_2$, $U_2 = S_1$, $U_3 = S_4$, $U_4 = S_3$, for $k = 5,\ldots, 4+\frac{n-4}{2}$, let $U_k = S_{k+\frac{n-4}{2}}$ and for $k = 5+\frac{n-4}{2},\ldots,n$, let $U_k = S_{k-\frac{n-4}{2}}$. Thus, let π be the permutation on N such that $\pi(1) = 2$, $\pi(2) = 1$, $\pi(3) = 4$, $\pi(4) = 3$, $\pi(k) = k + \frac{n-4}{2}$ for $k = 5,\ldots, 4+\frac{n-4}{2}$ and $\pi(\mathrm{k}) = k - \frac{n-4}{2}$ for $k = 5+\frac{n-4}{2},\ldots,n$.

By Strengthened Weak Pareto, $f(U_1,\ldots,U_n) \subset \{y, z\}$.

By Top Anonymity $y \in f(R_1,\ldots,R_n)$ if and only if $y \in f(U_1,\ldots,U_n)$ and $z \in f(R_1,\ldots,R_n)$ if and only if $z \in f(U_1,\ldots,U_n)$. Thus, $f(U_1,\ldots,U_n) = \{z\}$.

However, for all $j \in N$: either $[yP(S_j)zP(S_j)w$, $zP(U_j)yP(U_j)w$ for all $w \in X\backslash\{y, z\}$ and $S_j|(X\backslash\{y, z\}) = U_j|(X\backslash\{y, z\})]$ **or** $[zP(S_j)yP(S_j)w$, $yP(U_j)zP(U_j)w$ for all $w \in X\backslash\{y, z\}$ and $S_j|(X\backslash\{y, z\}) = U_j|(X\backslash\{y, z\})]$.

Thus by Top Neutrality, $f(S_1,\ldots,S_n) = \{z\}$ implies $f(U_1,\ldots,U_n) = \{y\} \neq \{z\}$ leading to a contradiction.

This proves the proposition. Q.E.D.

4 Point Majority property and its implications

This section emerged from a very stimulating discussion with Professor Janez Zerovnik (here after referred to as JZ).

Given $R\in W(X)$ and $x\in X$, let first $(x,R) = 1$ if x is ranked first at R and first $(x,R) = 0$, otherwise. It is easy to see that for all $R\in W(X)$, $\sum_{y\in X}\text{first}\,(y,R)$ is a positive integer and if $R\in L(X)$, then $\sum_{y\in X}\text{first}\,(y,R) = 1$. For $R\in W(X)$ and $x\in X$, let point $(x,R) = \frac{\text{first}\,(x,R)}{\sum_{y\in X}\text{first}\,(y,R)}$.

If for $R\in W(X)$ and $x\in X$ it is the case that x is the only alternative to be ranked first at R, then point $(x,R) = 1$ and point $(y,R) = 0$ for all $y\in X\backslash\{x\}$. In particular if for $R\in L(X)$, x is ranked first at R, then point $(x,R) = 1$, and is equal to 0, otherwise. In general (i.e. $R\in W(X)$), point $(x,R) = \frac{1}{\text{cardinality of the set of alternatives ranked first at } R}$ if x is ranked first at R and point $(x,R) = 0$, otherwise. Further, $\sum_{x\in X}\text{point}\,(x,R) = 1$ for all $R\in W(X)$.

Given $(R_1,\ldots, R_n)\in [W(X)]^N$ if $\{x\in X|\sum_{j=1}^{n}\text{point}(x,R_j) > \frac{n}{2}\} \neq \phi$, then this set must be a singleton. This follows from the fact that if for $x, z\in X$, with $x \neq z$, it is the case that $\sum_{j=1}^{n}\text{point}(x,R_j) > \frac{n}{2}$ and $\sum_{j=1}^{n}\text{point}(z,R_j) > \frac{n}{2}$, then $n = \frac{n}{2} + \frac{n}{2} < \sum_{j=1}^{n}\text{point}(x,R_j) + \sum_{j=1}^{n}\text{point}(z,R_j) = \sum_{j=1}^{n}[\text{point}(x,R_j) + \text{point}(z,R_j)] = \sum_{j=1}^{n}[\frac{\text{first}\,(x,R_j)+\text{first}\,(z,R_j)}{\sum_{y\in X}\text{first}\,(y,R_j)}] \leq n$, leading to a contradiction.

The following concept was orally suggested to me by JZ.

A CC f with domain Ω is said to satisfy **Point Majority Property** if for all profiles $(R_1,\ldots, R_n)\in\Omega$, the following holds: if there exists $x\in X$ such that $\sum_{j=1}^{n}\text{point}(x,R_j) > \frac{n}{2}$ then $f(R_1,\ldots,R_n) = \{x\}$.

Proposition 6

(i) If a CC f on a domain Ω satisfies Point Majority Property, then it satisfies Majority Property. If $\Omega \cap [W(X)]^N \neq \phi$, then the converse need not be true.

(ii) A CC f on $[L(X)]^N$ satisfies Point Majority Property if and only if it satisfies Majority Property.

Proof

(i) Suppose f on a domain Ω satisfies Point Majority Property. Let $(R_1,\ldots,R_n)\in\Omega$ such that for some $x\in X$, $\#\{j\in N|G(X,R_j) = \{x\}\} \geq \frac{n}{2}$ and $\#\{j\in N|x\in G(X,R_j)\}> \frac{n}{2}$. Thus, we must have either (a) $\#\{j\in N|G(X,R_j) = \{x\}\} > \frac{n}{2}$, or (b) $\#\{j\in N|G(X,R_j) = \{x\}\} = \frac{n}{2}$ and $\#\{j\in N|\{x\}\subset\subset G(X,R_j)\} - \#\{j\in N|G(X,R_j) = \{x\}\} > 0$.

Then since for all $j \in N$ with $G(X,R_j) = \{x\}$ we must have point (x,R_j) and for all $j \in N$, with $\{x\} \subset\subset G(X,R_j)$, we have $0 <$ point $(x,R_j) < 1$, in either case (i.e. in both cases (a) and (b)) we have $\sum_{j=1}^{n} \text{point}(x, R_j) > \frac{n}{2}$.

Hence by Point Majority property, we get $f(R_1,\ldots,R_n) = \{x\}$ and so f satisfies Majority Property.

In order to show that if $\Omega \cap [W(X)]^N \neq \phi$, then the converse need not be true, let $X = \{x,y,z\}$ with $n = 3$. Suppose that f is a CC on Ω such that for all $(R_1, R_2, R_3) \in \Omega$, if for some $w \in X$ it is the case that $\#\{j \in N | G(X,R_j) = \{w\}\} \geq 2$, then $f(R_1,R_2,R_3) = \{w\}$; otherwise $f(R_1, R_2, R_3) = G(X,R_1)$. Clearly f satisfies the Majority property.

Let $(R_1, R_2, R_3) \in \Omega$ with $G(X, R_1) = \{x,y\}$, $G(X,R_2) = \{x,z\}$ and $G(X,R_3) = \{x\}$. Then by the definition of f, we get $f(R_1, R_2, R_3) = \{x,y\}$, although point (x, R_1) + point (x, R_2) + point $(x,R_3) = \frac{1}{2} + \frac{1}{2} + 1 = 2 > \frac{3}{2}$, leading to a violation of the Point Majority property.

This counterexample can be generalized to arbitrary n and arbitrary non--empty finite X containing at least two alternatives. Let $x \in X$ and suppose there exists a preference profile $(S_1,\ldots,S_n) \in \Omega$, such that $G(X,S_j) = \{x,w_j\}$ for $j = 1,\ldots,n-1$ and $G(X,S_n) = \{x\}$, where $w_1,\ldots,w_n \in X \backslash \{x\}$, not necessarily distinct. Let f be the CC on Ω such that if there exists $y \in X$ satisfying $\#\{j \in N | G(X,R_j) = \{y\}\} \geq \frac{n}{2}$ and $\#\{j \in N | y \in G(X,R_j)\} > \frac{n}{2}$ then $f(R_1,\ldots,R_n) = \{y\}$. Clearly, f satisfies Majority property and $f(S_1,\ldots,S_n) = \{x,w_1\} \neq \{x\}$. However $\sum_{j=1}^{n} \text{point}(x, S_j) > \frac{n}{2}$ and thus $f(S_1,\ldots,S_n) \neq \{x\}$ implies a violation of the point majority property.

(ii) In order to show that for $\Omega = [L(X)]^N$, the two properties are equivalent, in view of (i) we need to show that for $\Omega = [L(X)]^N$, Majority property implies the Point Majority property. Thus let f be a CC on $[L(X)]^N$ satisfying the Majority Property and suppose that for some $(R_1, R_2,\ldots,R_n) \in [L(X)]^N$ it is the case that $\sum_{j=1}^{n} \text{point}(x, R_j) > \frac{n}{2}$. Since $(R_1, R_2,\ldots,R_n) \in [L(X)]^N$ for all $y \in X$ and $j \in N$, point $(y,R_j) \in \{0,1\}$. Hence it must be the case that $\#\{j \in N | G(X,R_j) = \{x\}\} > \frac{n}{2}$. By Majority property we get $f(R_1, R_2,\ldots,R_n) = \{x\}$ and so f satisfies the Point Majority property.

This proves the above proposition. Q.E.D.

In view of the above proposition, we have the following corollaries to propositions 1, 2, 4 and 5, whose proofs being very simple are being stated without proofs.

Corollary of proposition 1

If n is an odd integer and $[L(X)]^N \subset \Omega$, then there does not exist any CC on Ω which satisfies Weak Pareto, Point Majority property and No-spoiler condition.

Corollary of proposition 2

If n is even and $\Omega = [W(X)]^N$, then there does not exist any CC on Ω which satisfies Weak Pareto, Point Majority property and No-spoiler condition.

Corollary of proposition 4

If n is odd and $[L(X)]^N \subset \Omega$, then there does not exist any CC on Ω which satisfies Strengthened Weak Pareto, Point Majority property and No-loser spoiler condition.

Corollary of proposition 5

If n is an even natural number greater than two and $\Omega = [W(X)]^N$, then there does not exist any CC on Ω which satisfies Strengthened Weak Pareto, Point Majority property, Top Anonymity, Top Neutrality and No-loser spoiler condition.

Acknowledgments

I would like to thank two anonymous referees and all concerned with this journal for comments.

References

Arrow K.J. (1950), *A Difficulty in the Concept of Social Welfare*, The Journal of Political Economy, 58(4), 328-346.

Arrow K.J. (1963), *Social Choice and Individual Values*, Cowles Foundation for Research in Economics at Yale University, Monograph 12, John Wiley & Sons, New York.

Borgers Ch. (2010), *Mathematics of Social Choice: Voting, Compensation, and Division*, Society for Industrial and Applied Mathematics, Philadelphia.

Denicolo V. (1993), *Fixed Agenda Social Choice Theory: Correspondence and Impossibility Theorems for Social Choice Correspondences and Social Decision Functions*, Journal of Economic Theory, 59, 324-332.

Gibbard A. (1973), *Manipulation of Voting Schemes: A General Result*, Econometrica, 41(4), 587-601.

Kelly J.S. (1988), *Social Choice Theory: An Introduction*, Springer Verlag, Berlin, Heidelberg.

Lahiri S. (2019a), *Re-interpreting the Arrow-Sen Framework of Social Choice*, https://www.academia.edu/40224623/Re-interpreting_the_Arrow-Sen_Framework_of_Social_Choice

Lahiri S. (2019b), *On a Theorem Due to Alan D. Taylor about Aggregation of Preferences*, Arthaniti: Journal of Economic Theory and Practice, 18(1), 17-31.

Lahiri S. (2019c), *State-Salient Decision Rules for Choice under Uncertainty*, AIMS-International Journal of Management, 13(3), 191-195.

Pattanaik P.K., Lahiri S. (2017), *A Note on A.D. Taylor's Property of Independence of Irrelevant Alternatives for Voting Rules*, Studies in Microeconomics, 5(2), 99-104.

Rubinstein A. (2012), *Lecture Notes in Microeconomic Theory: The Economic Agent*, Princeton University Press, Princeton,Oxford.

Satterthwaite M.A. (1975), *Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions*, Journal of Economic Theory, 10(10), 187-217.

Zhang J.H., Christopher K., Xiao Z. (2006), *The Majority Rule in Individual Decision Making*, Organizational Behavior and Human Decision Process, 29(1), 102-111.

MULTIPLE CRITERIA DECISION MAKING
Vol. 14 2019

David M. Ramsey[*]

OPTIMAL SELECTION FROM A SET OF OFFERS USING A SHORT LIST

DOI: 10.22367/mcdm.2019.14.05

Abstract

The rise of the Internet has led to a huge amount of information being available at the touch of a button. This article presents a model of searching for a valuable good, e.g. a new flat, using the Internet to gain initial information about the offers available. This information is used to create a short list of offers to be observed more closely before making a final decision. Although there has been a lot of recent work on the use of short lists in decision making procedures, there has been very little work on how the length of a short list should depend on the parameters of the search problem. This article addresses this problem and gives results on the optimal length of a short list when a searcher is to choose one of n offers and the search costs are convex in the length of the short list. Several examples are considered.

Keywords: secretary problem, incomplete information, payment for additional information, agglomeration of multiple traits.

1 Introduction

This paper considers procedures of searching for a valuable good using short lists of promising offers to be inspected more closely. For example, suppose an individual living in a relatively large city is looking for a new flat. Nowadays, it is easy to obtain basic information on a large number of flats, e.g. price, floor space and location. Purchasing a flat simply on the basis of information from the Internet would be highly risky (without closer, physical inspection it would be difficult to assess how appropriate an offer is). However, viewing all the flats that might be suitable based on size, price and location might well be prohibitive in terms of the search costs involved. Hence, a strategy based on the concept of

[*] Wrocław University of Science and Technology, Faculty of Computer Science and Management, Wrocław, Poland, e-mail: david.ramsey@pwr.edu.pl, ORCID: 0000-0002-7186-1436.

a short list, i.e a relatively small set of seemingly attractive offers that are inspected more closely before a purchasing decision is made, is a natural heuristic to use in such scenarios.

Heuristics are useful tools due to the cognitive limitations of decision makers (DMs). Successful heuristics should be adapted to the abilities of DMs and the structure of the information gained during search (see Simon, 1955, 1956; Todd and Gigerenzer, 2000, as well Bobadilla-Suarez and Love, 2018). There has been a lot of recent research on the concept of short lists and the situations in which they are useful in decision making. Short lists are useful when the amount of information available exceeds the cognitive abilities of DMs or the costs of exhaustive search are too high (see Masatlioglu et al., 2012; Lleras et al., 2017). For example, short lists of possible holiday destinations can be constructed using information from friends and colleagues (see Bora and Kops, 2019). The use of short lists may also be useful when offers can be categorized (Armouti-Hansen and Kops, 2018). When using the Internet to search for offers, filters may be used to search a list of offers with multiple attributes by arranging offers according to traits that are seen to be the most important in the decision process (see Rubinstein and Salant, 2006; Mandler et al., 2012; Kimya, 2018). Mandler et al. (2012) argue that by considering the traits of an offer in order of decreasing importance, a DM acts in a similar manner to a utility maximizer. Kimya (2018) considers a similar problem where offers that are assessed the least positively according to a given trait are eliminated, starting with the most important trait. Such a procedure can be thought of as a strategy based on constructing short lists of decreasing size until a final decision is made.

This paper considers a model in which information from the Internet (or other information-rich source) plays an instrumental role in searching for a valuable resource, such as a second-hand car or a flat, but is supplemented by closer inspection of offers, which is assumed to occur offline. The cost of physically visiting and inspecting an offer is assumed to be much more costly than finding an offer via the Internet and inspecting the information contained there. For example, Internet search for a flat could begin by eliminating offers that do not satisfy the hard constraints defined by the DM based on price, floor space, number of rooms and location. If a large number of offers still remain (which is likely in large towns or cities), then the DM can make a short list of flats to view based on his/her preferences. The model assumes that the market is large enough to ensure that the best of the offers observed from this short list is acceptable to the DM.

Although much research has been carried out recently on theoretical aspects of using short lists, little work has been carried out on the question of how long short lists should be depending on the parameters of the search problem and the structure of the information. For example, when searching for a specialist employee, an employer often interviews a short list of four or five promising candidates based on a set of written applications. One obvious question regards the scenarios in which using short lists of such moderate length are optimal or near-optimal. The model presented in this paper is a step towards answering such questions.

The approach used in this paper is broadly based on theory regarding choice from a set of offers observed in parallel. Even when offers can be, in theory, considered in parallel, Saad and Russo (1996) and Bobadilla-Suarez and Love (2018) note that the order in which offers are observed may be important. For example, since the physical appearance of two flats cannot be compared in parallel, it is difficult to assess the first flat to be visited, since it cannot even be compared to a mental picture of other offers. Similarly, an average offer might be regarded as attractive if it appears after one or two unattractive offers. Hence, the theory of sequential search is also applicable to such decision problems.

In problems involving the choice of one offer from a set presented in parallel, one might suppose that a search procedure should satisfy the Weak Axiom of Revealed Preferences (WARP, see Lleras et al., 2017). In words, this criterion states that when offers x and y belong to both of the sets S and T and x is chosen from the set S, then y is never chosen from the set T. In mathematical terminology,

$$[c(S) = x, y \in S, x \in T] \Rightarrow [c(T) \neq y]. \tag{1}$$

The following example shows that a search procedure based on forming a short list of fixed length does not satisfy the WARP criterion. Suppose x and y are two offers such that initial inspection indicates that offer y is more attractive than offer x, but closer inspection reveals that offer x is better than offer y. One may define sets of offers S and T, such that both x and y would be on the short list of offers from set S and x is chosen, but x is not on the short list of offers from set T and the offer y is chosen.

Stigler (1961) presents a model of costly search for a good. The strategy of the DM is the number of offers to investigate, k. After observing these offers, the DM accepts the best offer (i.e. offers are essentially observed in parallel). The optimal number of offers to observe, k^* is the smallest value k such that the expected gain from observing an additional offer is less than the cost of observing an offer. In order to derive this optimal strategy, the DM needs to know the distribution of the values of offers and the cost of observing an offer.

MacQueen (1964) derives the form of the optimal strategy for a sequential search problem of a similar form to the one presented here. The search costs are assumed to be linear in both the number of offers seen and the number of offers inspected closely. A DM first decides whether to inspect an offer more closely based on an initial signal. After closer investigation, the DM either accepts the offer or continues searching. No recall of previous offers is possible. Under the optimal strategy, the DM closely investigates an offer only when the initial indicator of the value of an offer exceeds a given threshold. An offer is only accepted when its value according to the two signals exceeds a given threshold. Ramsey (2015) gives theoretical and numerical results for such problems when the two signals come from a joint normal distribution. To realise a strategy of this form, the precise values of the signals must be observed and to derive the optimal strategy, the DM should know the joint distribution of the signals.

Simon (1955) introduces the concept of satisficing, which can be adapted to both parallel and sequential search. Suppose two signals indicate the value of an offer. In sequential search, the DM observes the second signal only if the initial signal exceeds a given threshold. Given the second signal is observed, an offer is accepted if the second signal (considered on its own) exceeds an appropriate threshold). Assuming that the traits are observed in decreasing order of importance, such a strategy is normally near optimal (see Bearden and Connolly, 2007, 2008; Chun 2015).

Hogarth and Karelaia (2005) consider a similar model of parallel search based on deterministic elimination by aspects. Traits are assessed in decreasing order of importance. Offers that do not satisfy constraints based on successive traits are eliminated until either only one offer is left (which is then chosen), none of the offers satisfy the current constraint or all traits have been considered. In the final two cases, choice is made at random from the final non-empty set of offers.

Analytis et al. (2014) presents a similar model to the one presented here. There are two rounds of inspection. In the first round (parallel search), offers are ranked on the basis of an initial signal. In the second round (sequential search), the DM closely observes offers starting with the highest ranked and stops when the value of an offer exceeds the expected reward from future search. To realise such a strategy, the values of the offers must be observed and deriving the optimal strategy requires knowledge of the distribution of the value of an offer given the signal observed in the first round.

The model considered here assumes that search is parallel in both rounds. In the first round, a fixed number, n, of offers are ranked according to an initial signal. The k most highly ranked offers are then investigated more closely, after which the currently highest ranked offer is accepted. The payoff obtained is assumed to be a function of the values of the signals (or the rank according to this function) minus the search costs incurred. To realise a strategy of this form, the DM must be able to rank the observations according to the signals observed. Derivation of the optimal strategy requires knowledge of the joint distribution of the signals. Future research will investigate how robust a given strategy is to changes in the parameters (distribution of the signals, search costs). The search costs are assumed to be convex in the length of the short list. This reflects the fact that both the time spent searching and the cognitive effort involved are increasing in the length of the short list.

Section 2 considers the model. Section 3 presents a general result regarding the optimal length of a short list. Numerical results for three examples are presented in section 4. Section 5 gives a discussion of these results and section 6 presents some conclusions and directions for future research.

2 A model of search using short lists

A decision maker (DM) must choose one of n offers. The DM first observes in parallel a signal of the value of each offer. Assume that a linear ranking (from 1 to n) can be assigned to these initial signals (the rank i corresponds to the i-th best offer). This ranking will be called the initial ranking. The strategy of the DM is defined by the length of the short list, k, where $1 \leq k \leq n$. When $1 < k < n$, then in the second round the DM observes another signal of the value of the k best offers from the initial ranking. If $k = 1$, then the DM automatically chooses the best offer according to the initial ranking without observing the additional signal. If $k = n$, then the DM observes all of the offers closely. It is assumed that given the DM closely observes all the offers, then he/she can assign a linear ranking to these offers based on these signals combined. This ranking will be called the overall ranking. Naturally, after closer inspection, the DM can only rank the k offers on the short list with respect to each other (the DM's partial ranking). Assume that this partial ranking is in accordance with the overall ranking, i.e. offer i is ranked more highly than offer j in a partial ranking if and only if i is ranked more highly than offer j in the overall ranking. For convenience, it is assumed that if a DM is using a short list of length k, then the partial ranking of any offer that is not included in the second round of inspection is $k + 1$.

Note that any method for ranking based on multiple criteria (e.g. TOPSIS, see Yoon and Hwang, 1995) may be used to rank offers according to the initial signal and then rank the offers on the short list according to both signals. Such an approach will be adopted in the future.

Denote the set of permutations of (1, 2,…, n) by S. Let $\pi = (a_1, a_2, \dots, a_n) \in S$. Such a permutation can be used to denote the initial ranking of the offers according to their overall rankings such that the offer ranked i overall has ranking a_i in the initial ranking. Let $p(\pi)$ be the probability that the initial ranking is given by π and define $p_i(j)$ to be the probability that the overall rank of an offer is i given that its initial rank is j. It follows that

$$p_i(j) = \sum_{\pi:\, a_i=j} p(\pi). \tag{2}$$

Let R_j denote the random variable describing the overall rank of the offer with initial rank j. Assume that when $i < j$, then R_i is stochastically dominated by R_j, i.e. if object i has a better rank than object j based on the initial signal, then offer i is expected to have a better rank than offer j overall.

For the purposes of this article, it will be assumed that each offer can be described by a pair of random variables, X_1 and X_2, which come from a continuous joint distribution. The random variable X_1 describes the signal observed on initial inspection (the initial ranking is based on this signal) and the random variable X_2 describes the signal received after closer inspection. Note that X_1 and X_2 may be correlated with each other, but the pairs of numbers describing each offer are independent realizations from this joint distribution. The overall ranking of the object is based on U, where $U = X_1 + X_2$. It is assumed that U is stochastically increasing in the value of X_1 (this is automatically satisfied when X_1 and X_2 are independent).

We consider two models. According to Model A, the value of an offer to the DM is a non-increasing function of its overall rank. Under Model B, the value of an offer to the DM is given by U. It should be noted that the DM does not observe the numerical values of these signals, and hence does not observe the value of an offer, but can rank all the offers according to the initial signal and rank the offers that are closely inspected according to their value (the partial ranking). In other words, it is assumed that the DM can make perfect pairwise comparisons between objects, e.g. if $i < j$, then an offer of partial rank i has a greater value to the DM than an offer of partial rank j. Define W_i to be the value of the offer with initial rank i. Note that when $i < j$, then under both models R_i and W_j are stochastically dominated by R_j and W_i, respectively.

The DM's goal is to maximize the expected reward from search, which is defined to be the value of the offer accepted minus the search costs. Under Model A, the expected value of the offer accepted should be calculated with respect to the distribution of the overall rank of an offer given that its partial rank is equal to one. Under Model B, the expected value of the offer accepted should be calculated with respect to the distribution of U given that the partial rank of the offer is equal to one. The following two problems appear with regard to these assumptions. Firstly, calculation of the expected value of an offer given that its partial ranking is equal to one can be highly complex. Although in the particular cases of $k = 1$ and $k = n$, such calculations may be tractable, the results presented here are derived from simulations. Secondly, although the DM aims to maximize the expected reward from search, the DM cannot actually measure this reward. In order to solve this problem, it is assumed that the reward from search is a measure of the utility of an individual from the search procedure. Additionally, such rules are used by a population in which individuals copy/learn search procedures that are seen to be successful and/or the reproductive success of individuals depends on their success in such search procedures. Under such assumptions, the search rules that evolve should be near optimal in terms of the expected reward from search given the limits on the perceiptive abilities of the searchers.

It is intuitively clear that a short list of length k should consist of the k highest ranked offers from the initial observations. This results from the fact that the distribution of the reward obtained by selecting from these offers stochastically dominates the reward obtained by selecting from any other set of k offers given the initial ranking. The search costs are split into the costs of initial inspection and costs of closer inspection. The costs of initial inspection, given by $f_1(k, n)$, are strictly increasing in both the number of offers and the length of the short list, n and k respectively. These costs cover the effort needed for initial inspection of the offers and maintainence of the short list. Also, assume that f_1 is convex in k, i.e. $f_1(k, n) - f_1(k-1, n)$ is non-decreasing in k. This reflects the cognitive effort required to control short lists of long length. This is a simplification, since when $k = n$ the DM automatically inspects all the offers closely and hence, in this case, the search costs should not include the costs of controlling a short list.

The costs of closer inspection of the offers on the short list, given by $f_2(k)$ are assumed to be increasing and convex in length of the short list. Note that it may be natural to assume that the costs of closer inspection are linear in k (when $k \geq 2$, then each successive offer on the short list is inspected and compared to the best ranked of previously inspected offers). Let $f(k, n) = f_1(k, n) + f_2(k)$ denote the total search costs and $C_k = f(k, n) - f(k-1, n)$ denote the marginal costs of increasing the length of the short list from $k - 1$ to k.

3 Form of the optimal strategy

Let V_k be the value of the offer accepted when the length of the short list is k, i.e. $V_k = \max_{1 \leq i \leq k} W_k$. It follows that when $i < j$, then V_i is stochastically dominated by V_j. Denote by M_k the marginal increase in the expected value of the offer accepted when the length of the short list is increased from $k - 1$ to k, i.e. $M_k = E[V_k - V_{k-1}]$.

The criterion determining the optimal length of the short list is based on the following theorem:

Theorem 1
The marginal increase in the expected value of the offer accepted, M_k, is non--increasing in k.

Proof
By definition
$M_k = E[\max \{0, W_k - V_{k-1}\}]$; $M_{k+1} = E[\max \{0, W_{k+1} - V_k\}]$.

The fact that $M_k \geq M_{k+1}$ follows directly from the facts that W_k stochastically dominates W_{k+1} and V_k stochastically dominates V_{k-1}. ■

Corollary
The optimal length of the short list, k^*, satisfies the following condition:
i) when $M_2 \leq C_2$, then the optimal strategy is to automatically accept the object with initial ranking 1, i.e. $k^* = 1$ and none of the objects are inspected closely.
ii) when the above condition does not hold, then k^* is the largest integer k, such that $k \leq n$ and $M_k > C_k$.

This corollary follows from the fact that M_k is non-increasing in k and C_k is non-decreasing in k. Hence, when the first condition holds, then the marginal costs of increasing the length of the short list always exceed the marginal gain. When the first condition does not hold, then for all $k \leq k^*$, it follows that $M_k > C_k$ and for $k > k^*$, then $M_k \leq C_k$. Hence, the DM always gains by increasing the length of the short list when $k < k^*$, but when $k \geq k^*$the costs of increasing the length of the short list always outweigh the gains. Thus the optimal length of the short list is given by k^*. Note that when $M_k = C_k$, then the DM is indifferent between using a short list of length $k-1$ and using a short list of length k. The condition given above assumes that when the optimal length of the short list is not unique, then the smallest length from the set of optimal lengths is used.

4 Examples

In this section, we consider three examples.

4.1 Example A: Independent signals

Suppose X_1 and X_2 are independent random variables from the N(0, 1) and the N(0, σ^2) normal distributions, respectively, where the first parameter denotes the mean and the second parameter the variance. The value of an offer is given by

$U = X_1 + X_2$ and thus has a normal distribution with mean 0 and variance $1 + \sigma^2$. The DM makes an initial ranking of the offers based on X_1. The best k offers according to this initial ranking are then observed more closely and the DM then accepts the offer with the highest value of U from these offers. The search costs associated with initial inspection and maintanence of the short list are $f_1(k, n) = 0.0001(n + k^2)$ and the costs of closer inspection are $f_2(k) = c\sigma$, where c is a constant. The choice of the cost parameters should reflect the logic that strategies based on short lists should be successful when the costs of initial observation are low relative to the costs of closer inspection. The costs of close inspection are assumed to be proportional to the standard deviation of the second signal, since under sequential search based purely on the second signal the expected number of offers that are seen when c is fixed is independent of σ (see Ramsey, 2015). Hence, any changes in the optimal length of the short list when σ increases and c is fixed result from the amount of information contained in the second signal relative to the information contained in the first signal (as σ increases, the importance of the second signal compared to the first signal increases).

4.2 Example B: A best choice problem

The goal of the DM is to choose one of the best two offers overall. The initial rankings are based on X_1 (as defined in Example A) and the overall rankings of the offers are based on $U = X_1 + X_2$. The best offer overall is accepted if and only if it appears on the short list. The second best offer overall is accepted if and only if this offer appears on the short list and the best offer overall does not appear. The value of the second best object overall relative to the best object is r, where $0 \leq r \leq 1$. No reward is obtained by accepting any other object. Note that when $r = 0$, this problem is similar to the classical secretary problem (see Gilbert and Mosteller, 1966).

The value of the best object overall is defined such that the variance of the values of the offers in the second example is equal to the variance of U in the corresponding version of the first example. Hence, if the reward obtained by accepting the best offer overall is s, then

$$s^2(1 + r^2)\left(\frac{1}{n} - \frac{1}{n^2}\right) = 1 + \sigma^2. \tag{3}$$

This scaling is done in order to make the values of the offers and the search costs comparable for corresponding realizations of Example A and Example B. This enables comparison of the length of the optimal short list when the distribution of the values of the offers are non-skewed (Example A) and very highly skewed (Example B).

Obtaining analytical results for either of these problems is very difficult. In the first example, calculation of the marginal gains from increasing the length of the short list requires knowledge of the distribution of $W_k - V_{k-1}$. In the second example, this calculation requires knowledge of the distributions of the overall rank given the initial rank. Hence, the optimal length of the short list was obtained empirically on the basis of 100 000 simulations written in the R language of the search procedure when the number of offers $n \in \{20, 50, 100, 200\}$. These empirically obtained optimal lengths are described in Tables 1 and 2 (for relatively low and relatively high costs of close inspection, $c = 0.02$ and $c = 0.1$, respectively).

Several phenomena are visible from these results. Firstly, as the proportion of the information about the value of an offer contained in the second signal increases, the optimal length of the short list increases. Secondly, given the variance of the value of offers, the more skewed this distribution, then the greater the optimal length of the short list (it is shortest when the value comes from a normal distribution and greatest when a reward is obtained only when the best offer overall is accepted). Thirdly, as the costs of closer inspection relative to the amount of information contained by the second signal increase, then the optimal length of the short list decreases (as would be expected). Fourthly, when the values of offers come from a normal distribution, then the optimal length of the short list is almost independent of the total number of offers, particularly when the costs of closer inspection are relatively large.

In the best choice problems (Example B), the optimal length of the short list increases as the total number of offers increases. Due to the scaling of the value of the best offer according to the number of offers available, as n increases then the value of the best offer increases and it is often worthwile to carry out exhaustive search of all the offers in order to ensure obtaining the best of all the offers, especially when the costs of closer inspection are relatively low and the second signal contains a large proportion of the information regarding the ranking of an object. One interesting aspect is that when the costs of closer inspection are high and the total number of offers is large, then the optimal length of the short list can actually fall as the amount of information contained in the second signal increases. In such cases, it seems likely that the search costs required in order to ensure obtaining the best offer with a large probability become prohibitive and thus the intensity of search falls (the optimal length of the search list falls). This conclusion seems to be confirmed by the behaviour of the optimal length of the short list when the value of the best offer is not scaled according to the number of offers, i.e. is always equal to 1. In this case, as the amount of information in the second signal increases, the optimal length of the short list initially increases before falling to 1 and the optimal expected reward from search becomes negative. This indicates that when the number of offers becomes large and the amount of information contained in the first signal is small, then it does not pay to search for the best offer, since the search costs become too large.

Table 1: Optimal lengths of short lists for relatively low costs of close inspection ($c = 0.02$). The numbers in each cell give the optimal lengths of the short list for Example A, Example B with $r = 1$ and Example B with $r = 0$, respectively

σ	n=20	n=50	n=100	n=200
1/5	(2, 3, 5)	(3, 4, 7)	(3, 5, 9)	(3, 6, 10)
1/4	(3, 4, 6)	(3, 5, 8)	(3, 6, 10)	(3, 7, 12)
1/3	(3, 4, 7)	(4, 6, 10)	(4, 7, 14)	(5, 9, 16)
1/2	(5, 6, 9)	(5, 8, 15)	(6, 10, 20)	(5, 14, 25)
1	(6, 9, 15)	(7, 15, 27)	(9, 22, 39)	(9, 30, 56)
2	(10, 13, 20)	(11, 25, 42)	(14, 41, 67)	(14, 60, 106)
3	(12, 15, 20)	(15, 30, 49)	(14, 52, 89)	(15, 85, 148)
4	(14, 16, 20)	(15, 36, 50)	(16, 58, 96)	(18, 92, 168)
5	(14, 17, 20)	(16, 38, 50)	(17, 64, 99)	(19, 103, 187)

Source: Based on simulations written by the author in R.

Table 2: Optimal lengths of short lists for relatively high costs of close inspection ($c = 0.1$). The numbers in each cell give the optimal lengths of the short list for Example A, Example B with $r = 1$ and Example B with $r = 0$, respectively

σ	$n = 20$	$n = 50$	$n = 100$	$n = 200$
1/5	(2, 3, 4)	(2, 3, 5)	(2, 4, 7)	(2, 5, 8)
1/4	(2, 3, 5)	(2, 4, 6)	(2, 4, 7)	(2, 5, 9)
1/3	(2, 3, 5)	(2, 4, 7)	(2, 5, 9)	(2, 6, 11)
1/2	(2, 5, 7)	(3, 6, 10)	(3, 7, 12)	(3, 9, 17)
1	(3, 6, 10)	(3, 9, 16)	(4, 12, 23)	(4, 18, 32)
2	(4, 8, 14)	(4, 14, 22)	(4, 19, 29)	(5, 27, 45)
3	(4, 9, 17)	(5, 17, 29)	(5, 23, 38)	(5, 32, 43)
4	(5, 10, 19)	(5, 17, 31)	(5, 26, 38)	(5, 30, 41)
5	(5, 11, 20)	(5, 18, 36)	(5, 23, 42)	(5, 35, 35)

Source: Based on simulations written by the author in R.

Note that due to the nature of the simulations, the empirically derived optimal lengths of short lists show some fluctuation around what would be the theoretically optimal length. This is particularly true when the optimal length of the short list is large (in such cases, slightly changing the length of the short list from the optimal length has very little impact on the expected reward). The expected reward from search in Example A as a function of the length of the short list is presented in Figure 1 for the case $n = 200, c = 0.02, \sigma = 1$.

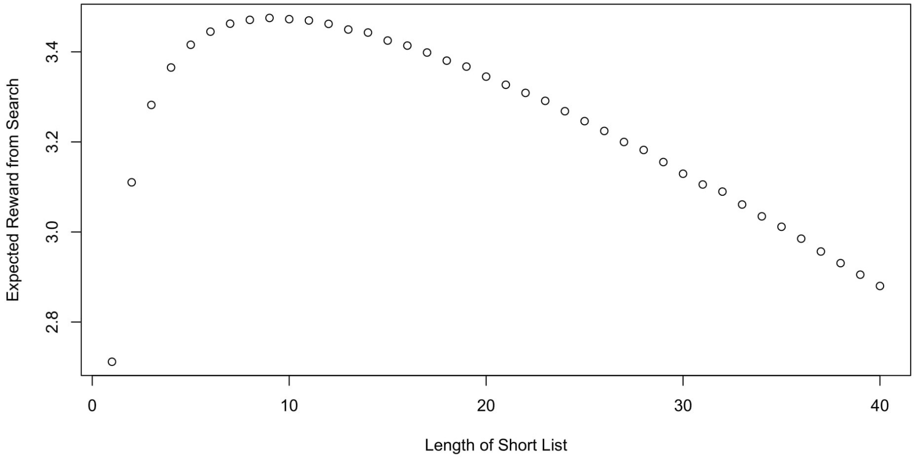


Figure 1: Expected reward as a function of the length of the short list for Example A when the number of offers is large, the costs of closer inspection relatively small and the initial and second signal contain the same amount of information ($n = 200, c = 0.02, \sigma = 1$)

4.3 Example C: Correlated signals

Since the signals indicating the value of an offer may be correlated, it is assumed that the value of an offer is based on two signals (X_1, X_2) from a bivariate normal distribution. X_1 comes from a standard normal distribution, the coefficient of correlation between X_1 and X_2 is ρ and the residual variance of X_2, i.e. the variance in X_2 that is not explained by X_1, is σ^2. Hence, given X_1, X_2 has a normal distribution with mean ρX_1 and variance σ^2. The overall variance of the signal X_2 is $\frac{\sigma^2}{1-\rho^2}$. As in Example A, the value of an offer is $U = X_1 + X_2$. From these assumptions, $E(U) = 0$ and

$$\begin{aligned} Var(U) &= Var(X_1) + Var(X_2) + 2\rho\sqrt{Var(X_1)Var(X_2)} \\ &= 1 + \frac{\sigma^2}{1-\rho^2} + \frac{2\rho\sigma}{\sqrt{1-\rho^2}}. \end{aligned} \tag{4}$$

It can be shown by differentiation that this variance is increasing in ρ for $\rho > 0$.

The search costs are defined as in Example A, i.e. the costs of closer inspection are proportional to the residual variance of X_2. On one hand, the increase in the overall variance of the offer favours more intense search (a longer short list). On the other hand, observing X_1 gives us information about the value of the second signal (hence, all other things being equal, as ρ increases X_1 contains relatively more information about the value of an offer). This effect favours short lists with fewer items. The optimal lengths of short lists were derived empirically for $\rho \in \{0.2, 0.4, 0.6, 0.8\}$ on the basis of 100 000 simulations, written in R, of the search process. Table 3 presents the empirically obtained optimal lengths of short lists when the relative costs of closer inspection are $c = 0.1$.

Table 3: Optimal lengths of short lists for relatively high costs of close inspection ($c = 0.1$) and correlated signals (Example C). The numbers in each cell give the optimal thresholds for $\rho = 0, 0.2, 0.4, 0.6$ and 0.8, sequentially

σ	$n = 20$	$n = 50$	$n = 100$	$n = 200$
1/5	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)
1/4	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)
1/3	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)	(2, 2, 2, 2, 2)
1/2	(2, 2, 2, 2, 2)	(3, 3, 2, 2, 2)	(3, 3, 2, 2, 2)	(3, 3, 2, 2, 2)
1	(3, 3, 3, 2, 2)	(3, 3, 3, 3, 2)	(4, 3, 3, 3, 3)	(4, 3, 3, 3, 3)
2	(4, 4, 3, 3, 2)	(4, 4, 4, 3, 3)	(4, 4, 4, 3, 3)	(5, 4, 4, 3, 3)
3	(4, 4, 3, 3, 3)	(5, 4, 4, 3, 3)	(5, 4, 4, 3, 3)	(5, 5, 4, 4, 3)
4	(5, 4, 4, 3, 3)	(5, 4, 4, 3, 3)	(5, 5, 4, 4, 3)	(5, 4, 4, 4, 3)
5	(5, 4, 4, 3, 3)	(5, 5, 4, 3, 3)	(5, 4, 4, 4, 3)	(5, 5, 4, 4, 3)

Source: Based on simulations written by the author in R.

Table 3 indicates that when the residual variance of the second signal is fixed, there is a tendency for the optimal length of the short list to fall, especially when the residual variance of the second signal is relatively large. Hence, the increase in the overall variance in the value of an offer is outweighed by the fact that X_1 explains an increasing proportion of the variance in the value of an offer. Note that if the (overall) variances of the two signals were fixed, then this negative relationship of the coefficient of correlation with the optimal length of the short list would be stronger.

5 Discussion of the results from the simulations

Although a large number of articles have been recently published on the concept of short lists in decision making, there has been little work on models which indicate what the optimal length of a short list should be depending on the parameters of a search problem. Such an approach to decision making seems very fruitful in the Internet age, since when searching for a valuable resource in a large market, basic information about offers can be found at very little cost via the Internet. However, DMs should investigate promising offers more closely, before making a final decision. The model presented here illustrates the qualitative properties of optimal strategies based on short lists.

In practical terms, if there are a large number of candidates for a given position, it seems more reasonable to assume that the reward obtained by a real life DM will depend on the intrinsic value of an offer, rather than whether the offer accepted is the best offer or not. Also, the signals associated with an offer may be correlated. Hence, for practical purposes the results from Example C, where the two signals observed are correlated, are the most instructive. Of course, independent signals can be treated as a particular case within this framework.

The results indicate that the optimal length of the short list is only very weakly dependent on the total number of offers available. On the other hand, in such problems the optimal length of the short list is increasing in the importance of the second signal relative to the first signal. Although it may be difficult for a DM to estimate the total number of offers before the search procedure is realized, in general the DM will know the type of information gained in the two stages of the search procedure and be able to estimate the weight of the information gained at both stages. For example, procedures applied by employers to search for a specialist employee might be one interesting practical example of the use of short lists. The short list is made on the basis of written applications. The cost of collecting these applications and creating a short list are assumed to be small compared to the value of the employee. The members of this short list are invited for interview. Since the employer often covers the travel costs of the interviewees and interviews occupy a significant amount of time, the relative costs of closer inspection are high. In such problems, the rule of thumb "invite four or five candidates for interview and then offer the position to the best of these candidates" is often used. The results from this paper indicate that such rules are optimal or close to optimal when a) the values of employees come from a non-skewed distribution, b) the DM obtains a similar amount of novel information from the initial signal (written application) and closer inspection (interview), c) the costs of closer inspection are high compared to the costs of initial inspection. For example, when the two signals contain the same amount of information and the number of offers is large ($n = 200$) and $c = 0.1$ (high costs of closer inspection), the optimal length of the short list is 4. When $c = 0.02$ (low costs of closer inspection), the optimal length of the short list is 9. However, using a short list of length 4 ensures an expected payoff that is not much lower than the optimal payoff (see Figure 1). Hence, using short lists of moderate length ensures at least near-optimal rewards over a wide range of parameter sets.

Fixing the marginal variances of the signal, the optimal length of the short list tends to decrease as the coefficient of correlation ρ, between the signals becomes larger, particularly when the overall variance of the second signal is large. The overall variance of the value of an offer is increasing in ρ, which might lead to an increase in the overall search effort (i.e. a longer short list). However, this effect is more than counteracted by the fact that the weight of the initial signal as an estimate of the overall value of an offer becomes greater. In terms of searching for a skilled employee, e.g. if the verbal skills of a candidate are highly correlated with the quality of the candidate as observed in a written application, then it is only necessary to interview a small number of candidates.

It should be noted that these results also indicate that the greater the skew in the distribution of the values of offers, then the greater is the optimal length of a short list. However, this should be investigated in more detail by considering signals with different distributions.

6 Conclusions and directions for future research

There are obvious weaknesses in this approach, as outlined below.

Firstly, it is assumed that a linear ranking of offers can be defined according to a) initial information, and also on b) initial information combined with information from closer inspection. This implicitly assumes that any pairwise comparison of offers always results in one offer being preferred to another and the results of such comparisons are always correct. In practice, comparison of offers based on the initial information may be difficult since it embraces multiple traits, e.g. when searching for a new flat, Internet sites will give quantitative information about price, floor space, number of rooms and location (e.g. distance from city centre). Hence, future work will concentrate on how such multivariate information can be used to create a short list. Approaches to multicriteria decision making, such as TOPSIS or AHP, could be used to create a short list (see Yoon and Hwang, 1995). Due to the ever increasing availability of information and the speed with which computers can analyze data, it seems that methods of automating at least the initial stage of such a search procedure would be beneficial. For example, the database of available flats in a city may be too large for efficient manual handling. Hence, it seems natural that an algorithm could be developed to suggest a short list of flats to view based on the stated preferences of the DM regarding price, floor space etc. This would decrease the search costs in the first round. However, there may exist a conflict of interest between the goals of the writer or implementer of the algorithm (a seller or intermediary) and the DM (e.g. Skulimowski, 2017). In this case, it would be natural to examine game theoretic models of such search procedures. Future research will model such algorithms and investigate their effectiveness in various types of problem (e.g. computer aided search procedures for selecting a short list of flats or second-hand cars to view).

Also, future models should investigate problems related to bounds on the ability of a DM to compare offers (the phenomena of inconsistent comparisons and the inability to state a preference between two offers). Similar phenomena also appear when comparing the offers on the short list after the round of closer inspection.

The assumption that the costs of maintaining the short list are convex in the length of the short list seem reasonable, due to the increased cognitive effort required to control longer short lists. However, the costs of initial observation and maintaining the short list should be modelled more formally. This might be done by defining the costs of initial observation according to the number of pairwise comparisons required. Such an approach will be used in future research. Also, when the length of the short list is large in comparison to the total number of offers, it is cognitively simpler to carry out an exhaustive search of all the offers as this only requires remembering the best offer seen so far on the basis of closer inspection. Future research should compare the expected reward under the optimal short list policy with the expected reward obtained using exhaustive search based on an appropriate model.

The optimal short list policy should also be compared with search based purely on the initial signal. Finally, comparison with the optimal reward obtained in the corresponding sequential decision problem where the search costs are linear in both the number of offers observed and the number of offers inspected closely (see Ramsey, 2015) will give a measure of the efficiency of short list rules.

The model does not take into account various factors that might be relevant in real-life decision problems. For example, a DM might feel, after closely inspecting the offers on the short list, that none of the offers are suitable and thus the search procedure should be repeated. Also, the price might be negotiable or the offer chosen after close inspection might become unavailable before the DM can make a decision. How should such factors be modelled? Also, the decision to purchase a valuable good is often made by households rather than individuals. Hence, models of group decision making or game theoretic models should be considered.

In practice, there is great uncertainty regarding the parameters of a problem. In order to determine the optimal length of the short list, the model presented here assumes that the DM should know a) the distribution of the value of the offers, b) the proportion of the value of an offer described by the first signal, c) the total number of offers, d) the search costs. There will be uncertainty regarding these parameters, which might be modelled using probability theory or fuzzy theory.

In conclusion, the frequency with which procedures based on short lists are applied when humans make decisions indicate that such heuristic policies are successful. Future research should investigate how the perceptive abilities of searchers can be supported using automatic systems designed, for example, to search a large database of offers in order to suggest a short list of offers appropriate for close inspection.

Acknowledgements

The author is grateful for the financial support of the Polish National Science Centre via grant no. 2018/29/B/HS4/02857, "Logistics, Trade and Consumer Decisions in the Age of the Internet".

References

Analytis P.P., Kothiyal A., Katsikopoulos K. (2014), *Multi-attribute Utility Models as Cognitive Search Engines*, Judgment and Decision Making, 95, 403-419.

Armouti-Hansen J., Kops C. (2018), *This or That? Sequential Rationalization of Indecisive Choice Behavior*, Theory and Decision, 84(4), 507-524.

Bearden J.N., Connolly T. (2007), *Multi-attribute Sequential Search*, Organizational Behavior and Human Decision Processes, 103(1), 147-158.

Bearden J.N., Connolly T. (2008), *On Optimal Satisficing: How Simple Policies can Achieve Excellent Results*, [in:] Y.-J. Son, J.C. Smith, T. Connolly (eds.), *Decision Modeling and Behavior in Complex and Uncertain Environments*, Springer, New York, NY, 79-97.

Bobadilla-Suarez S., Love B.C. (2018), *Fast or Frugal, but not Both: Decision Heuristics under Time Pressure*, Journal of Experimental Psychology: Learning, Memory, and Cognition, 44(1), 24.

Borah A., Kops C. (2019), *Rational Choices: An Ecological Approach*. Theory and Decision, https://doi.org/10.1007/s11238-019-09689-5.

Chun Y.H. (2015), *Multi-attribute Sequential Decision Problem with Optimizing and Satisficing Attributes*, European Journal of Operational Research, 243(1), 224-232.

Gilbert J.P., Mosteller F. (1966), *Recognizing the Maximum of a Sequence*, Journal of the American Statistical Association, 61(313), 35-73.

Hogarth R.M., Karelaia N. (2005), *Simple Models for Multiattribute Choice with Many Alternatives: When it does and does not Pay to Face Trade-offs with Binary Attributes*, Management Science, 51(12), 1860-1872.

Kimya M. (2018), *Choice, Consideration Sets, and Attribute Filters*, American Economic Journal: Microeconomics, 10(4), 223-247.

Lleras J.S., Masatlioglu Y., Nakajima D., Ozbay E.Y. (2017), *When More is Less: Limited Consideration*, Journal of Economic Theory, 170, 70-85.

MacQueen J.B. (1964), *Optimal Policies for a Class of Search and Evaluation Problems*, Management Science, 10(4), 746-759.

Mandler M., Manzini P., Mariotti M. (2012), *A Million Answers to Twenty Questions: Choosing by Checklist*, Journal of Economic Theory, 147(1), 71-92.

Masatlioglu Y., Nakajima D., Ozbay E.Y. (2012), *Revealed Attention*, American Economic Review, 102(5), 2183-2205.

Ramsey D.M. (2015), *On a Sequential Decision Process Where Offers Are Described by Two Traits*, Multiple Criteria Decision Making, 10, 141-154.

Rubinstein A., Salant Y. (2006), *A Model of Choice from Lists*, Theoretical Economics, 1(1), 3-17.

Saad G., Russo J.E. (1996), *Stopping Criteria in Sequential Choice*, Organizational Behavior and Human Decision Processes, 67(3), 258-270.

Simon H.A. (1955), *A Behavioral Model of Rational Choice*, The Quarterly Journal of Economics, 69(1), 99-118.

Simon H.A. (1956), *Rational Choice and the Structure of the Environment*, Psychological Review, 63(2), 129.

Skulimowski A.M.J. (2017), *Cognitive Content Recommendation in Digital Knowledge Repositories – A Survey of Recent Trends*, [in:] L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. Zadeh, J. Zurada (eds.), *Artificial Intelligence and Soft Computing*, ICAISC 2017, Lecture Notes in Computer Science, 10246, 574-588, Springer, Cham.

Stigler G.J. (1961), *The Economics of Information*, The Journal of Political Economy, 69(3), 213-225.

Todd P.M., Gigerenzer G. (2000), *Précis of Simple Heuristics that Make us Smart*, Behavioral and Brain Sciences, 23(5), 727-741.

Yoon K.P., Hwang C.L. (1995), *Multiple Attribute Decision Making: An Introduction*, Sage Publications, Thousand Oaks, CA.

MULTIPLE CRITERIA DECISION MAKING

Vol. 14 2019

Honorata Sosnowska*
Paweł Zawiślak**

DIFFERENCES BETWEEN JURORS IN CLASSICAL MUSIC COMPETITIONS: THE MCDM AND NETWORK THEORY APPROACHES

DOI: 10.22367/mcdm.2019.14.06

Abstract

This paper analyses the voting in two of the major international classical music competitions, which were held recently, viz. the International Henryk Wieniawski Violin Competition and the International Chopin Piano Competition, as well as the hypothesis, raised in some media reports, that there were juror cliques in the Wieniawski Competition. Network theory is used to compare the rankings of the two Chopin competitions. Jurors are nodes and they are linked if the correlation between the ordered list of competitors, as measured by the Kendall rank correlation coefficient, exceeds a given threshold value. The obtained networks were found linked in the case of the Chopin Competition, but disconnected in the case of the Wieniawski Competition. The results indicate that there may have been cliques in the Wieniawski Competition, but not in the Chopin Competition.

The problem can be descibed in MCDM terminology by labelling the contestants 'variants' and the jurors (or, more precisely, their musical preferences) – 'criteria'. The similarity of any two criteria is measured by correlating the orders of the alternatives (i.e. variants) that

* SGH Warsaw School of Economics, Department of Mathematics and Mathematical Economics, Warsaw, Poland, e-mail: honorata@sgh.waw.pl, ORCID: 0000-0001-8249-2368.

** SGH Warsaw School of Economics, Department of Mathematics and Mathematical Economics, Warsaw, Poland, e-mail: pzawis@sgh.waw.pl, ORCID: 0000-0002-5297-7644.

result from applying them. The problem of juror cliques is thereby transformed into one of finding groups of criteria that are similar in the case of these variants.

Keywords: correlation network, minimal spanning tree, voting method.

1 Introduction

Two of the worlds' major classical music competitions are held in Poland every five years, viz. the International Chopin Piano Competition and the International Henryk Wieniawski Violin Competition. This paper analyses the most recent competitions, viz. the 16^{th} (2010) and 17^{th} (2015) Chopin Competitions and the 2016 Wieniawski Competition. The voting methodology is analysed using social choice methods in Sosnowska (2013, 2017). Each competition has its own specific voting methodology. Analyses of classical music competitions are not all that common, especially when compared with e.g. sports competitions. There are a few papers that connect music, sport and art in their assessment of expert competitions, e.g.: Rzążewski et al. (2014), Gambarelli (2008), Gambarelli et al. (2012), Przybysz (2000) and Flores and Ginsburgh (1996). These papers also briefly analyse ski-jumping, figure skating, wrestling, the Triennale Grafiki, the Prime Minister's Prize, and European grants competitions. Voting systems in classical music competitions, sports competitions and grant competitions are similarly constructed and can be considered a special case of MCDM.

In a multi-stage classical music competition many voting methods are used prior to the decisive final stage, where some kind of ranking is usually employed. There were 10 contestants and 12 jurors in the final of the 16^{th} Chopin Competition (the voting methodology is described in Sosnowska (2013)) and 10 contestants and 17 jurors in the final of the 17^{th} Chopin Competition. The ranking was based on a distribution of 55 points on a scale 1-10, so that only one contestant could get 10 points. In the most recent Wieniawski Competition there were 7 contestants and 13 jurors in the final. A reverse Borda count was used, i.e. the winner received 1 point. In each competition exceptions were made to the rules if a contestant was a student of a juror.

Sport competitions such as ski-jumping and figure skating are judged using a similar methodology. It therefore follows that any competition judged by experts can be analysed using a methodology similar to that applied in classical music competitions.

Both "Gazeta Wyborcza", one of Poland's most popular daily newspapers (Dębowska, 2016) and "Ruch Muzyczny" (Januszkiewicz and Chorościak, 2016), the country's leading music journal, raised the possibility that the jurors of the most recent Wieniawski Competition formed cliques. This hypothesis is tested using network theory (see Jackson, 2006; Newman et al., 2010, 2006).

The paper is structured as follows: section 2 transforms the voting in classical music competitions into an MCDM problem, section 3 includes basic information about networks, section 4 applies network theory to analyse jury homogenity and section 5 presents the conclusions.

2 Voting in classical music competitions as an MCDM problem

The variants (also called alternatives) and the decision makers, as well as their characterizations and their decision-making methodologies have to be established first. The variants are the contestants and there are very few of them in the final of a competition. The variants are ranked according to certain criteria. A criterion is the musical preference of a particular juror. There are as many criteria as there are jurors. The variants can be weakly ordered according to each criterion. The criteria are aggregated into the final decision according to the voting method chosen.

One of the applications of MCDM is to detect and eliminate manipulation (Kontek and Sosnowska, 2018). The final result is the arithmetic mean of the jurors' votes (i.e. the scores given according to the criteria laid down). As mean values are very sensible to outliers, the mean is treated as a reference point and the distances between it and the individual results are measured. For this purpose, the Manhattan distance was selected from the many metrics available. This is applied in Kontek and Sosnowska (2018). The 20% of results that are most distant from the mean are then removed, i.e. the criteria farthest from the reference point are not considered, and the mean is recalculated. It has been shown both theoretically and empirically that this method is less prone to manipulation than the original method.

3 Basic network theory

This section briefly discusses the network concepts used in this paper. It begins by reviewing basic concepts, such as nodes and links, before shifting focus to weighted networks and correlation networks. Finally, the concept of minimal spanning trees is revisited. All these concepts are examinated in more detail in Newman et al. (2010, 2006), Horvath (2011) and West (2001).

3.1 Networks

This paper considers mainly simply undirected networks. A *(simple undirected) network* is a pair $N = (N(N), L(N))$ consisting of a (usually finite) set $N(N)$ of *nodes* and a set $L(N)$ of *links*, where every link $l \in L(N)$ is a subset of $N(N)$ and consists of two (different) elements. Networks are often called *graphs* in the literature whereas nodes and links – *vertices* and *edges*, *sites* and *bonds*, or *actors* and *ties*, respectively.

The *degree* of a node is the number of links for which it is an endpoint.

A *path* in a network is a finite sequence $(i_0, i_1, \ldots, i_k)$ of nodes, such that every two consecutive nodes form a link. A network is *connected* if there is a path $(i = i_0, i_1, \ldots, i_k = j)$ connecting every two nodes i and j. A *connected component* of a network is its maximal (in the sense of inclusion) connected subnetwork. A *tree* of n nodes is a connected network with $n-1$ links.

3.2 Weighted networks and correlation networks

Let N be a network, and assume that there exists a map $w : L(N) \to \mathbb{R}$. The triple $(N(N), L(N), w)$ is called a *weighted network*.

Consider a family $\{X_s : s \in S\}$ of variables and a threshold coefficient ρ. A *correlation network* N can be constructed as follows. S is the set $N(N)$ of nodes. The set $L(N)$ consists of the subsets $\{s, t\} \subset S$, such that $s \neq t$ and $|\mathrm{corr}(X_s, X_t)| \geq \rho$. This network is a weighted network whose weights are given by $w(l) = \mathrm{corr}(X_s, X_t)$ for all links $l = \{s, t\} \in L(N)$.

Weighted networks are used in biology (Horvath, 2011; Newman et al., 2006). Correlation networks are used in stock markets analysis (Górski et al., 2006; Cherifi et al., 2017) and in studies on the structural and functional organization of the human brain (Park and Friston, 2013).

3.3 Minimal spanning trees

Let N be a connected network and let $w : L(N) \to \mathbb{R}$ be a weighting function. Assume that $w(l) \geq 0$ for every link $l \in L(N)$. A *minimal spanning tree* (MST) for N is a subtree of N containing all the nodes of N such that the sum of the weights of all links is minimal.

MSTs were first used in Mantegna (1999) to indicate the most important (in some sense) currencies.

4 Voting in Polish classical music competitions – a network approach

This section analyses the correlation networks of jurors' votes in final stages of the 15$^{\text{th}}$ International Henryk Wieniawski Violin Competition and the 16$^{\text{th}}$ and 17$^{\text{th}}$ International Chopin Piano Competitions.

4.1 15$^{\text{th}}$ International Henryk Wieniawski Violin Competition

This section analyses the final results of the 15$^{\text{th}}$ International Henryk Wieniawski Violin Competition. There were seven participants and eleven jurors, who had been using the inverse *Borda count* method (see Hołubiec and Mercik, 1994; Nurmi, 1987; Ordeshook, 1986 for a description of this method). The results are illustrated in Table 1.

These results were used to create a weighted network W^{2016} as follows. The node set $N(W^{2016})$ corresponds to the jurors and the link set $L(W^{2016})$ comprises all links $\{J_s, J_t : s \neq t\}$ betweem them. For a link l_{st} connecting nodes J_s and J_t, weight $w(l_{st}) = w_{st}$ is assigned, where $w_{st} = \tau_{st}$ is *Kendall's τ coefficient* (Abdi, 2007; Kendall, 1938, 1948) of the voting results of jurors J_s and J_t for $s, t = 1, 2, \ldots, 11$.

As this is a complete network on 11 nodes, it can be made more readable by creating networks W_p^{2016} for $p = 0.1, 0.2, 0.3, 0.4, 0.5$ such that links l_{st} with weights satisfying the condition $|w(l_{st})| \leq p$ are removed. These networks are presented in Figure 1.

Note that $W_{0.5}^{2016}$ has three groups of jurors ($\{J_1, J_4, J_8\}$, $\{J_2, J_3, J_6, J_7\}$ and $\{J_5, J_9, J_{11}\}$) such that the jurors voted coherently inside these groups, whereas jurors from the first group voted incoherently with the members of the second group (see Figure 4a).

4.2 16$^{\text{th}}$ and 17$^{\text{th}}$ International Chopin Piano Competitions

This section examines the final results of the 16$^{\text{th}}$ and 17$^{\text{th}}$ International Chopin Piano Competitions. The voting methods employed are described in Sosnowska (2013, 2017). These results are illustrated in Table 2 and Table 3, respectively.

In these cases the networks C^{2010}, C_p^{2010} for $p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$, C^{2015} and C_p^{2015} for $p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ are studied. These networks are constructed in a similar way to those analysed earlier. The Kendall τ coefficients are calculated by omiting unavailable data. The networks are presented in Figures 2 and 3, respectively.

Note that although $C^{2010}_{0.5}$ and $C^{2015}_{0.5}$ are disconnected, they still have one main connected component. This means that the votes of individual jurors were not strongly mutually correlated. Networks C^{2010}_{p} and C^{2015}_{p} split for $p = 0.6$ and $p = 0.7$, respectively (Figures 2g and 3h), but they still have the main connected components - core groups of jurors voting very coherently.

4.3 Minimal spanning trees

This section analyses the MSTs of the networks described above. Let N be a correlation network. The weights of the links in this network are given by $w(l) = \mathrm{corr}(X_s, X_t)$ for $l = \{s, t\}$. For stock market networks, this correlation is measured by a *Pearson correlation coefficient* (Boddy and Smith, 2009).

As the correlation coefficient takes values from the interval $[-1, 1]$, to determine the MST for N, the correlation coefficients need to be transformed into a metric given by $d_{st} = \sqrt{2(1 - \rho_{st})}$, where ρ_{st} is the Pearson correlation coefficient of the variables X_s and X_t (see Mantegna, 1999).

As Kendall's τ coefficient is used here, the *Kendall distance* (Abdi, 2007; Kendall, 1938, 1948) is used instead.

The minimal spanning trees for networks W^{2016}, C^{2010} and C^{2015} are presented in Figure 5.

Note that the nodes with the highest degree in MST of W^{2016} are in the three main connected components in $W^{2016}_{0.5}$, whereas those with the highest degree in MSTs of C^{2010} and C^{2015} are in the same main connected component in $C^{2010}_{0.6}$ and $C^{2015}_{0.8}$, respectively (see Figures 4 and 5).

4.4 Conclusions from the comparison of the networks

A cursory inspection of the W^{2016} and C^{2010} and C^{2015} networks is sufficient to conclude that they differ (see Figures 1a, 2a and 3a). The individual votes in the two Chopin Competitions were coherent (there were only a few pairs of jurors whose votes were negatively correlated), whereas those in the Wieniawski Competition were much more inconsistent.

We now focus on analysing the connectedness of the networks for different threshold values p. Only those links with positive weights (marked solid) are considered. This is because a positively weighted link indicates coherent voting on the part of the jurors who correspond to the nodes connected by it.

For $p = 0.2$ networks $C^{2010}_{0.2}$ and $C^{2015}_{0.2}$ are connected (see Figures 2c and 3c). For $p = 0.3$ and $p = 0.4$ networks $C^{2010}_{0.3}$ and $C^{2010}_{0.4}$ have two connected components (see Figures 2d and 2e), although one of them consists of a single node (the juror MK, who voted the most inconsistently), and networks $C_{0.3}^{2015}$ and $C_{0.4}^{2015}$ are connected (Figures 3d and 3e). For $p = 0.5$, networks $C_{0.5}^{2010}$ and $C_{0.5}^{2015}$ have three and two connected components respectively (Figures 2f and 3f). Even though $C_{0.5}^{2010}$ and $C_{0.5}^{2015}$ are disconnected, they still have one main connected component, which implies that the votes of the individual jurors were not strongly correlated with the votes of others.

Networks C_p^{2010} and C_p^{2015} split for $p = 0.6$ and $p = 0.7$, respectively (Figures 2g and 3h), but they still have the main connected components, i.e. core groups of jurors voting very coherently.

It should be noted that for $p > 0.1$, networks C_p^{2010} and C_p^{2015} have no negatively weighted links. This implies that there were no strongly negative correlated votes amongst jurors during the Chopin Competitions in 2010 and 2015.

On the other hand, even though for $p < 0.4$ networks W_p^{2016} are connected and for $p = 0.4$ network $W_{0.4}^{2016}$ has one main connected component (Figure 1), they have negatively weighted links. This interesting phenomenon can be observed for the threshold value $p = 0.5$. $W_{0.5}^{2016}$ has three groups of jurors ($\{J_1, J_4, J_8\}$, $\{J_2, J_3, J_6, J_7\}$ and $\{J_5, J_9, J_{11}\}$) such that amongst these groups jurors voted coherently, whereas jurors from the first group voted contrarily to the ones from the second group (see Figure 4a).

The analysis of the MSTs of networks W^{2016}, C^{2010} and C^{2015} concentrates on the degrees of nodes. The construction of the MST leads to the conclusion that a node (i.e. a juror) with a high degree in MST has many neighbours (other jurors) in the initial network that voted similarly. Therefore nodes with high degrees in MST can indicate the most influential jurors. In the present case, these influential jurors are J_1, J_7 and J_9 in W^{2016}, KK and DST in C^{2010}, and TD in C^{2015} (see Figures 4 and 5). In the first network, jurors J_1, J_7 and J_9 are in the three groups mentioned above (Figure 4a), whereas jurors KK and DST are in the same connected component: the second one (Figure 4b).

The nature of impact of the most influential jurors in W^{2016} differs from that of the most influential jurors in C^{2010} and C^{2016}. In the first case, the most influential jurors affect only jurors from their own groups, whereas in the second case, the most influential jurors affect almost all the others.

5 Conclusions and recommendations for further research

The voting results described above contain a great deal of information about the preferences of voters and their structure. The problem of voting can be described as an MCDM problem. Network theory can be applied to highlight the properties of networks constructed on the basis of jurors' votes. The obtained networks can be used to describe the homogeneity or heterogeneity of those votes.

Acknowledgments

While working on this paper, the second author was partially supported by the following SGH funds: KAE/S16/03/16, KAE/S17/BD, KAE/S18/BD and KAE/S18/08/18.

All calculations and figures were prepared with R 3.4.4 (R Core Team, 2019).

References

Abdi H. (2007), *The Kendall Rank Correlation*, [in:] N.J. Salkind (ed.), *Encyclopedia of Measurement and Statistics*, SAGE Publications, Thousand Oaks (CA).

Boddy R., Smith G. (2009), *Statistical Methods in Practice: For Scientists and Technologists*, Wiley, Chichester.

Cherifi H., Gaito S., Quattrociocchi W., Sala A. (eds.) (2017), *Complex Networks & Their Applications V*, Proceedings of the 5th International Workshop on Complex Networks and their Applications, Springer, Cham.

Dębowska A.S. (2016), *Konkurs Wieniawskiego 2016: Wewnętrzna wojna jurorów. Zwyciężczyni ledwo weszła do finału* [*2016 Wieniawski Competition. Internal Jury War. Winner Barely Makes Final*], http://wyborcza.pl/7,113768,20925157,konkurs-wieniawskiego-2016-wewnetrzna-wojna-jurorow.html?disableRedirects=true (16 September 2019).

Flores R., Ginsburgh V. (1996), *The Queen Elisabeth Musical Competition: How fair is the final ranking*, Journal of the Royal Statistical Society D-Sta, 45(1), 97-104.

Gambarelli G. (2008), *The „Coherent Majority Average" for Juries' Evaluation Processes*, Journal of Sports Sciences, 26, 1091-1095.

Gambarelli G., Iaquinta G., Piazza M. (2012), *Anti-collusion Indices and Averages for Evaluation of Performances and Judges*, Journal of Sports Sciences, 30, 411-417.

Górski A.Z., Dróżdż S., Kwapień J., Oświęcimka P. (2006), *Complexity characteristics of currency networks*, Acta Physica Polonica B, 37(11), 2987-2995.

Hołubiec J., Mercik J.W. (1994), *Inside Voting Procedures*, Accedo Verlagsgesellschaft, Munich.

Horvath S. (2011), *Weighted Network Analysis: Applications in Genomics and Systems Biology*, Springer Science & Business Media, New York.

Jackson M. (2006), *The Economics of Social Networks*, [in:] R. Blundell, W. Newey, T. Persson (eds.), *Advances in Economics & Econometrics: Theory & Applications, Ninth World Congress of the Econometric Society*, Cambridge University Press, Cambridge, 1, 1-56.

Januszkiewicz M., Chorościak E. (2016), *Wielka gra [Great game]*, Ruch Muzyczny, 11, 50-55.

Kendall M.G. (1938), *A New Measure of Rank Correlation*, Biometrika, 30(1-2), 81-93.

Kendall M.G. (1948), *Rank Correlation Methods*, Charles Griffin & Company Limited, London.

Kontek K., Sosnowska H. (2018), *Specific Tastes or Cliques of Jurors?*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3297252, working paper.
Mantegna R.N. (1999), *Hierarchical Structure in Financial Markets*, The European Physical Journal B, 11(1), 193-197.
Newman M.E.J. (2010), *Networks: An Introduction*, Oxford University Press, Oxford.
Newman M.E.J., Barabási A.-L., Watts D. (2006), *The Structure and Dynamics of Networks*, Princeton University Press, Princeton.
Nurmi H. (1987), *Comparing Voting Systems*, D. Reidel Publishing Company, Dordrecht.
Ordeshook P. (1986), *Game Theory and Political Theory*, Cambridge University Press, Cambridge.
Park H.J., Friston K. (2013), *Structural and Functional Brain Networks: From Connections to Cognition*, Science, 342(6158), 1238411, 1-8.
Przybysz D. (2000), *Agregowanie ocen sędziów sportowych jako przykład zbiorowego podejmowania decyzji [Aggregation of Sports Judges' Scores as an Example of Collective Decision Making]*, Studia Socjologiczne, 1-2, 105-136.
R Core Team (2019), *R: A Language and Environment for Statistical Computing*, Home page: http://www.R-project.org/.
Rzążewski K., Słomczyński W., Życzkowski K. (2014), *Każdy głos się liczy [Every vote counts]*, Wydawnictwo Sejmowe, Warsaw.
Sosnowska H. (2013), *The Rules for the Jury of the Fryderyk Chopin Piano Competition as a non Standard Voting Rule*, Roczniki Kolegium Analiz Ekonomicznych, 32, 1-9.
Sosnowska H. (2017), *Comparison of Voting Methods Used in Some Classical Music Competitions*, [in:] N.T. Nguyen, R. Kowalczyk, J. Mercik (eds.), *Transactions on Computational Collective Intelligence XXVII*, LNCS (Lecture Notes in Computer Science 10480), Springer, 108-117.
West D.B. (2001), *Introduction to Graph Theory*, Second Edition, Prentice Hall, Upper Saddle River.

6 Tables and Figures

Table 1: Final results of the 15$^{\text{th}}$ International Henryk Wieniawski Violin Competition

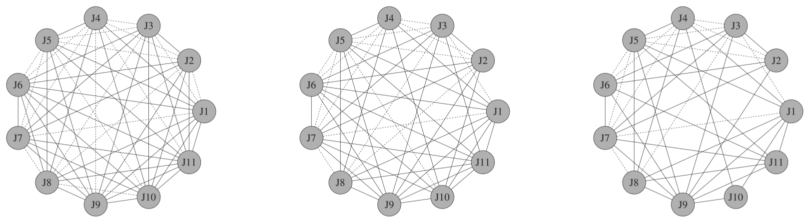
	J1	**J2**	**J3**	**J4**	**J5**	**J6**	**J7**	**J8**	**J9**	**J10**	**J11**
A	7	3	2	7	7	4	3	7	7	7	7
B	4	7	7	2	2	7	7	2	5	6	5
C	5	5	5	3	6	6	5	5	6	1	6
D	3	6	4	5	1	5	4	4	3	5	1
E	1	4	6	1	3	3	6	3	4	3	4
F	6	2	1	6	4	2	1	6	1	2	2
G	2	1	3	4	5	1	2	1	2	4	3

Table 2: Final results of the 16$^{\text{th}}$ International Chopin Piano Competition
(s denotes that the participant was a student of a juror
and was therefore not rated by that juror)

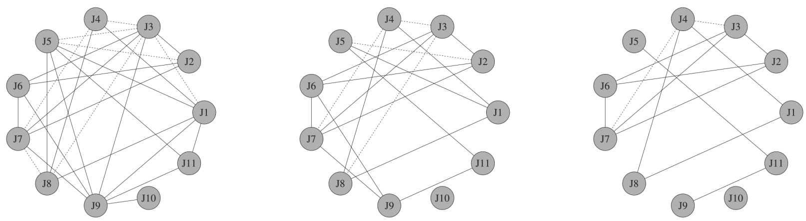
	MA	**DTS**	**BD**	**PE**	**FT**	**NF**	**AH**	**AJ**	**KK**	**MK**	**PP**	**KP**
A	1	1	1	1	1	2	4	4	1	4	1	2
B	2	2	3	9	4	2	6	7	2	10	6	1
C	9	10	5	3	2	8	2	5	7	9	5	5
D	2	3	2	4	3	3	3	2	1	1	2	4
E	6	5	6	7	9	5	9	10	8	4	4	7
F	3	4	7	8	6	6	10	6	7	6	7	6
G	1	3	9	2	8	1	7	3	2	5	4	3
H	10	10	10	5	10	9	8	9	9	8	10	9
I	5	7	4	6	7	7	5	8	5	3	8	8
J	4	1	1	1	3	4	s	2	3	6	1	1

Table 3: Final results of the 17$^{\text{th}}$ International Chopin Piano Competition
(s as in Table 2)

	DA	**MA**	**TD**	**AE**	**PE**	**NG**	**AH**	**AJ**	**GO**	**JO**	**PP**	**EP**	**KP**	**JR**	**WS**	**DY**	**Y**
A	10	9	8	9	1	9	6	9	9	10	9	9	9	9	9	9	9
B	2	4	2	3	3	2	6	3	6	1	1	1	1	5	1	5	4
C	1	6	7	5	2	5	2	8	1	5	4	6	6	3	5	3	5
D	7	4	s	3	5	5	9	10	9	9	10	8	9	8	10	8	6
E	9	5	s	4	8	8	3	6	4	8	8	8	8	4	7	7	6
F	4	4	1	3	4	2	2	5	2	6	5	2	4	2	5	4	3
G	3	4	4	5	4	2	5	2	5	3	2	2	1	7	4	6	2
H	8	9	8	9	8	10	7	6	8	7	6	7	10	9	8	9	9
I	6	5	5	5	7	6	7	3	7	2	3	5	1	6	1	2	4
J	5	5	s	8	6	6	2	2	3	4	7	6	6	1	5	2	5

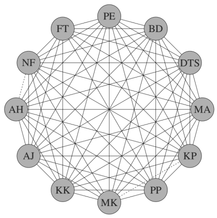


(a) Network W^{2016} (b) Network $W^{2016}_{0.1}$ (c) Network $W^{2016}_{0.2}$

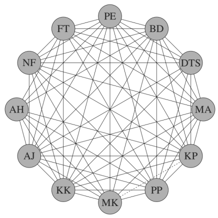


(d) Network $W^{2016}_{0.3}$ (e) Network $W^{2016}_{0.4}$ (f) Network $W^{2016}_{0.5}$

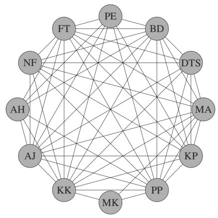
Figure 1: Vote correlation networks constructed from the final results of the 15$^{\text{th}}$ International Henryk Wieniawski Violin Competition (solid – positive correlation, dashed – negative correlation)



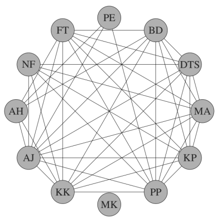
(a) Network C^{2010}



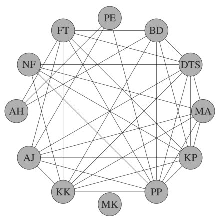
(b) Network $C^{2010}_{0.1}$



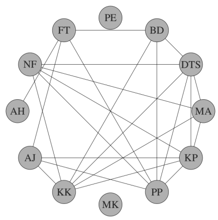
(c) Network $C^{2010}_{0.2}$



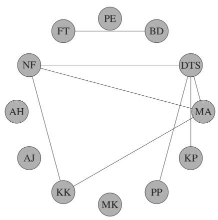
(d) Network $C^{2010}_{0.3}$



(e) Network $C^{2010}_{0.4}$

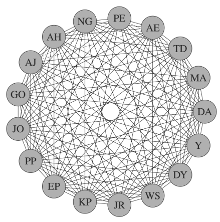


(f) Network $C^{2010}_{0.5}$

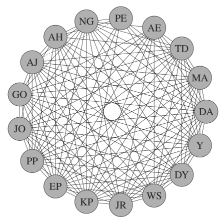


(g) Network $C^{2010}_{0.6}$

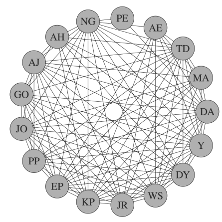
Figure 2: Vote correlation networks constructed from the final results of the 16^{th} International Chopin Piano Competition



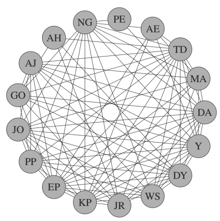
(a) Network C^{2015}



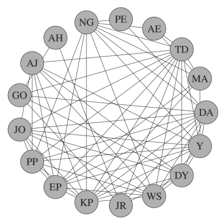
(b) Network $C_{0.1}^{2015}$



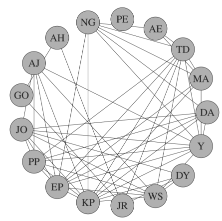
(c) Network $C_{0.2}^{2015}$



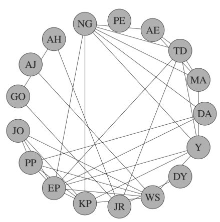
(d) Network $C_{0.3}^{2015}$



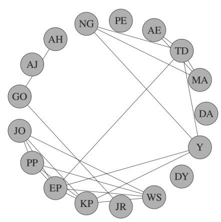
(e) Network $C_{0.4}^{2015}$



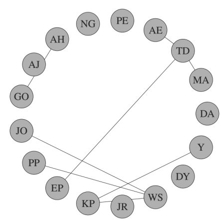
(f) Network $C_{0.5}^{2015}$



(g) Network $C_{0.6}^{2015}$

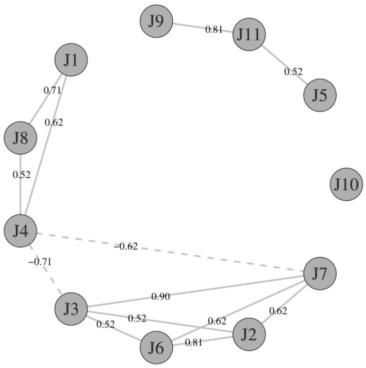


(h) Network $C_{0.7}^{2015}$

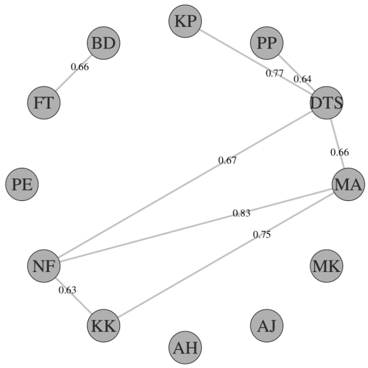


(i) Network $C_{0.8}^{2015}$

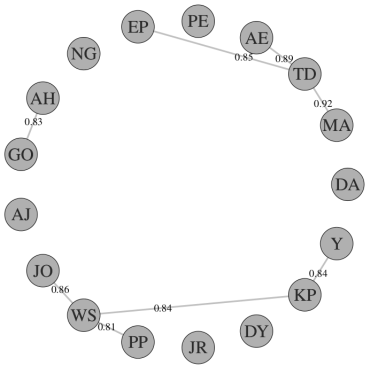
Figure 3: Vote correlation networks constructed from the final results of the 17$^{\text{th}}$ International Chopin Piano Competition



(a) Network $W_{0.5}^{2016}$

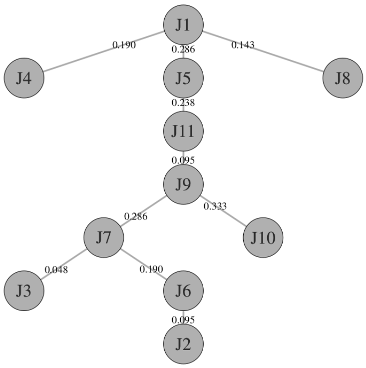


(b) Network $C_{0.6}^{2010}$

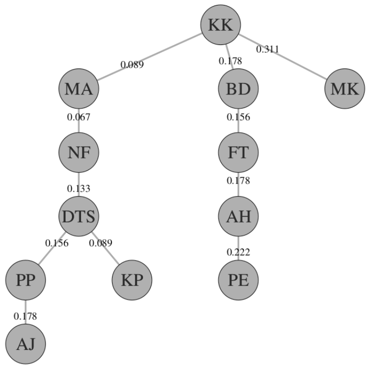


(c) Network $C_{0.8}^{2015}$

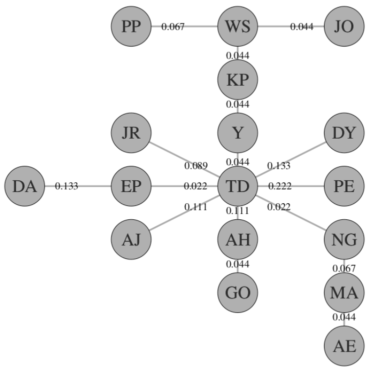
Figure 4: Networks of strongly correlated votes



(a) MST of W^{2016}



(b) MST of C^{2010}



(c) MST of C^{2015}

Figure 5: Minimal spanning trees

Faith-Michael E. Uzoka*
Boluwaji A. Akinnuwesi**

DEVELOPMENT AND EVALUATION OF AN AHP MODEL FOR SOFTWARE SYSTEMS SELECTION

DOI: 10.22367/mcdm.2019.14.07

Abstract

Decision-making in the field of information systems has become more complex due to larger number of alternatives, multiple and sometimes conflicting goals, and an increasingly uncertain environment. Software systems play unique roles in the translation of corporate strategic and tactical plans into actions. We present the results of a study designed to develop and evaluate an Analytical Hierarchy Process (AHP) model to support decision making in the selection of appropriate software system to meet organizational needs. Our results show the viability of the AHP methodology in software system/project selection, and points to the importance of functionality (35.26%), quality (22.00%) and usability (19.34%) criteria in the overall decision process. Cost and vendor service did not seem to exert significant weight in the decision matrix.

Keywords: analytical hierarchy process, multi-criteria decision, software project, selection factors, functionality, cost.

1 Introduction

Software projects are complex and dynamic, comprising of a number of unstructured tasks that are affected by internal, external, and social factors (Meso et al., 2006). Software solutions to organizational needs are achieved through a systematic process of analysis, involving defining alternatives and selecting the best option in terms of software or software project (Hoffer, George and Valacich, 2016). A wrong software system/project selection could adversely

* Mount Royal University, Department of Mathematics and Computing, Calgary, Canada, e-mail: fuzoka@mtroyal.ca.
** University of Eswatini, Department of Computer Science, Kwaluseni M201, Eswatini (Swaziland), e-mail: moboluwaji@gmail.com.

affect the organization’s ability to function effectively and accomplish its strategic and tactical goals (Rouyendegh and Erkan, 2011; Uzoka et al., 2016). Software evaluation and selection is an intense activity, which could take months and several personnel in planning and deciding on critical concomitants that should go into the decision matrix. According to Uzoka et al. (2008) software evaluation and selection is a technology adoption decision, which revolves around product and organizational characteristics.

Software system selection decision involves multiple, sometimes conflicting objectives, and a blend of qualitative and quantitative criteria (Hwang and Yoon, 2012). The process of software selection is made difficult by the multiplicity of products, variation in product performance, and uncertainties of users’ needs. The selection of inappropriate packages may compromise business processes, impact negatively on the functioning of the organization, and could jeopardize the very existence of the organization (Uzoka et al., 2016; Verville et al., 2002). Software products from different backgrounds are likely to exhibit different strengths and weaknesses; therefore, it is essential to employ methodical means for evaluating and selecting appropriate software that is cost effective and suits the business process needs, structure, culture, and environment of the organization. The existing structured methodologies for IS project selection range from single-criteria cost/benefit analysis (Hares et al., 1994) to multiple criteria scoring models (Melone et al., 1984), and ranking methods (Buss, 1983).

In this paper, we built on our previous works (Uzoka et al., 2016; Akinnuwesi and Uzoka, 2016) that identified variables that could be referenced by management in the evaluation of software project proposals. We recognize that the process of evaluating and selecting appropriate software project proposal for an organization is multi-criteria oriented and hence the use of AHP to prioritize and rank the proposals submitted for evaluation based on judgmental evaluation through peer ratings. The rest of the paper is organized as follows: in section 2, we present the AHP methodology and the results of our exploratory factor analysis, which helped us reduce the variables into manageable factors; in section 3, we present the model evaluation and results, while in sections 4 and 5, we present the limitations or the study and conclusion respectively.

The rest of the paper is organized as follows: in section 2, we present the AHP methodology and the results of our exploratory factor analysis, which helped us reduce the variables into manageable factors; in section 3, we present the model evaluation and results, while in sections 4 and 5, we present the limitations or the study and conclusion respectively.

2 Research methodology

This study adopted the classical AHP methodology (Saaty, 1977) in the development of a model for the evaluation software project, with the intent of selecting the best vendors/products. We developed three sets of questionnaires to obtain data for: 1) Identifying software project evaluation variable and carrying out factor analysis with the aim of reducing the variables to manageable factors; 2) Developing the AHP model; 3) Evaluating some software projects, based on the model. Our research methodology is depicted visually in Figure 1.

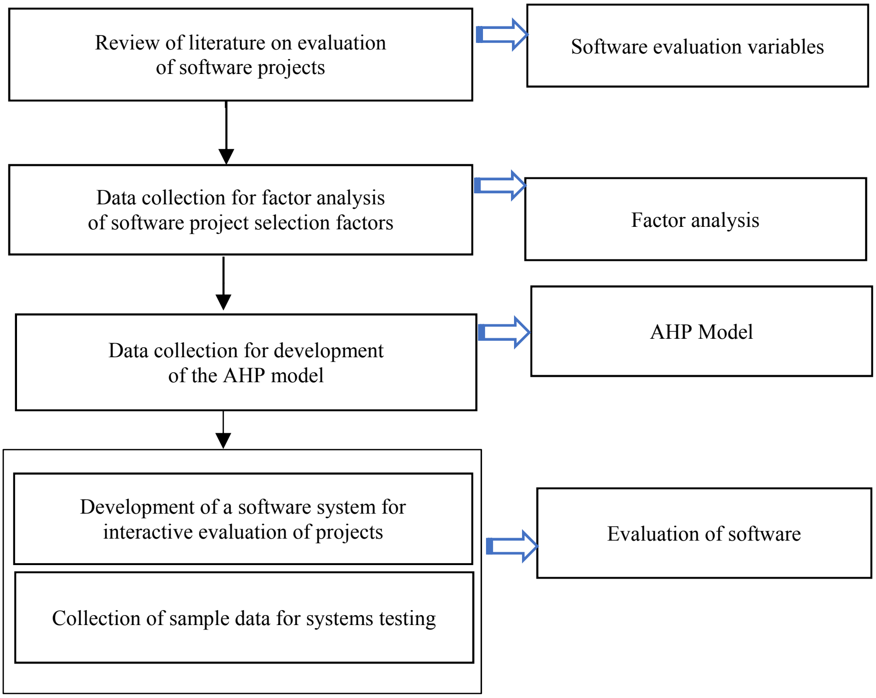


Figure 1: Research methodology

2.1 Factor analysis

The review of relevant literature on software and IT project evaluation produced 83 variables, which were utilized in the initial questionnaire, leading to factor analysis. The questionnaire had two sections (A and B). Section A consisted of respondent's demographic information, while section B evaluated the relevance of each of the 83 factors in software project evaluation. The variables were measured using a 5-point Likert-Type scale, ranging from 1-5. A total of 200 questionnaires were distributed physically and electronically via emails to individuals, who were directly or indirectly involved in IT/software projects in Nigeria. A total of 160 questionnaires (80%) were correctly filled/returned and used for the factor analysis in SPSS. Our respondents consisted mainly of users (88.8%) who had long years (5 year and above) experience in the use of software packages (90.6%), mainly from ICT, communications, audit, and insurance (67%).

The exploratory factor analysis employed maximum likelihood extraction method to reduce the evaluation variables to a set of significant variables in the evaluation and selection of software. The KMO measure (0.534) and the Barlett's test of sphericity (4749.152, p = 0.00) point to the adequacy of data for factor analysis. Fourteen factors were extracted in more than 25 iterations with convergence = 0.072. Applying the social science rule on the initial factor matrix generated, this did not give a meaningful factor loading. To obtain a meaningful factor loading, the initial matrix was rotated by orthogonal transformation by Quartimax with Kaiser Normalization. The rotated factor matrix provided a clear pattern of loading and was more meaningful for interpretation and therefore, used for the analysis. The rotated factor matrix produced fourteen factors: *Module Content, User's Experience, Vendors Technical Know-how, Ease of Customization, Vendor Experience, System Adaptability, User Interest, Interoperability and Completeness*, *Reliability, Organizational Budget*, *Ease of Use, Integration, Cost of Implementation, System Efficiency.* These factors and the variables that loaded on them, were utilized in the development of the AHP questionnaire for developing the AHP model.

2.2 AHP model

The AHP (Saaty, 1977) helps the decision maker in understanding the structuring of decision variables to determine their relative importance in the decision process. A major advantage of the AHP methodology is the ability to convert qualitative constructs into numerical values and allows diverse variables to be compared with one another in a rational and consistent way. The AHP process can be summarized as follows:

Step one – decomposition phase:

a. ***Identify all decision alternatives***: For this research, the decision alternatives are the software choices.
b. ***Identify all the criteria for evaluation***: the criteria are the evaluation variables. The evaluation variables used in the proposed framework were sourced from various literature (e.g. Al-Harbi, 2001; Chau, 1995; Davis et al., 1994; Jadhav et al., 2009; Khaddaj et al., 2004; Liberatore et al., 2003; Maidamisa et al., 2012; Nandi et al., 2011; Rouyendegh et al., 2011; Saaty, 2008; Uzoka et al., 2008, 2009, 20013, 2016; Vargas et al., 2010; Verner et al., 2009; Verville et al., 2003; Wei et al., 2004; Wei et al., 2005; Zielsdorff et al., 2010). The variables were reduced to manageable factors, using factor analysis. This made it easier to develop a hierarchy of criteria.

c. *Develop the hierarchy of criteria for prioritizations*:

i. Identify the overall goal/objective of the selection
ii. Identify appropriate criteria to satisfy a goal
iii. Identify where appropriate, sub-criteria under each criterion. This is represented in Table 1.

The factor analysis produced 44 variables, which loaded on 14 different constructs. We further grouped the related constructs into six major criteria, namely: cost, functionality, system flexibility, usability, quality and vendor service.

Table 1: IT Project evaluation criteria hierarchy

Goal: Software Project Evaluation	Cost	ORGANIZATIONAL BUDGET (BUDG)	A1: Service speed of the system
			B1: Defined organization policies relating to systems and vendors
			C1: Project budget
		COST OF IMPLEMENTATION (COST)	A2: Installation and implementation of the software/hardware
			B2: License cost
			C2: Cost of hardware
	Functionality	MODULE CONTENT (MODL)	A3: Number of modules on distributed tiers of the s/w
			B3: Number of modules on separate server
			C3: Number of independently modules
			D3: Number of workstation provided
			E3: Provision of reference site by vendor
		EASE OF CUSTOMIZATION (CUST)	A4: Customizable fields in modules of the s/w
			B4: Customizable report produced by the s/w
			C4: User Interface type
			D4: Communication standards provided by the system
		INTEROPERABILITY & COMPLETENESS (NTRP)	A5: Availability of modules in the s/w
			B5: Completeness of the software
			C5: Interoperability of the system with other systems
	System Flexibility	SYSTEM ADAPTABILITY (ADPT)	A6: Openness of the software
			B6: Parameter in the settings of the settings
			C6: Adaptability in the system to the organization's environment
		SYSTEM INTEGRATION (INTG)	A7: Platform Independence
			B7: upgradability of the system
			C7: Ease of integration with other IS
	Usability	USER INTEREST (USINT)	A8: User interest in s/w
			B8: Willingness of the user to use the system
		USERS EXPERIENCE (EXPR)	A9: User experience in the problem area of the s/w system
			B9: Professional qualification of the users of the system
			C9: familiarity of user with the IT tools provided by the system
			D9: Length of experience of user of the system
		EASE OF USE (EASE)	A10: Ease of use of graphical interface
			B10: Ease of operation of s/w and hardware

Table 1 cont.

	Quality	RELIABILITY (RELB)	A11: Stability of both s/w and h/w
			B11: Recovery ability in case of failure
		SYSTEM EFFICIENCY (EFFCY)	A12: Main storage constraint of the system
			B12: Service execution time of the system
			C12: Strength of communication devices
	Vendor Service	VENDOR EXPERIENCE (VDEX)	A13: Length of experience of vendor
			B13: Warranty provided by the vendor
			C13: Past business experience of vendor
		VENDOR TECHNICAL KNOWHOW (VDTK)	A14: Ease of implementation of the system
			B14: Good implementation service
			C14: Technical business skills of vendor/developer
			D14: Internal technical knowledge of the vendor/developer

Step two – analysis phase

Establish a priority model by identifying the relative importance of criteria through pairwise comparison. The pairwise comparison is done from the top level of the hierarchy to the bottom level to establish the overall priority index. Measurement of preferences involves a pairwise comparison of evaluation variables, which are verbal statements about the strength of importance of a variable over another, represented numerically on an absolute scale. The comparison is done from the top level of the hierarchy to the bottom level in order to establish the overall priority index.

Let $P(i, j)$ be a pairwise comparison of two elements i and j;
where $\{i, j\} \in n_k$ (n_k = node k of the AHP tree).

The larger the value of $P(i, j)$, the more i is preferred to j in the priority rating.

The following rules govern the entries in the PWC.

***Rule 1*:** $P(j, i) = [P(i, j)]^{-1} \quad (1 \leq P \leq 9)$ (1)

***Rule 2*:** If element i is judged to be of equal importance with element j , then $P(i, j) = P(j, i) = 1$; in particular, $P(i, i) = 1$ for all i. (2)

The pairwise comparison (PWC) matrices for levels 2 and 3 criteria are shown in Table 2.

Table 2: Levels 2 and 3 PWC matrix

		Cost (CO)		**Functionality (FU)**			**System Flexibility (SF)**		**Usability (US)**			**Quality (QA)**		**Vendor Service (VS)**	
		BUDG	COST	MODL	CUST	NTRP	ADPT	INTG	USINT	EXPR	EASE	RELB	EFFCY	VDEX	VDTK
CO	BUDG	1.000	1.210	0.140			0.215		0.180			0.172		0.390	
	COST	0.827	1.000												
FU	MODL	7.148		1.000	0.383	0.291	2.896		2.037			1.898		5.278	
	CUST			2.614	1.000	1.122									
	NTRP			3.432	0.891	1.000									
SF	ADPT	4.643		0.345			1.000	1.099	0.719			0.574		2.687	
	INTG						0.910	1.000							
US	USINT	5.560		0.491			1.392		1.000	4.648	1.000	0.917		3.552	
	EXPR								0.215	1.000	0.221				
	EASE								1.000	4.526	1.000				
QA	RELB	5.811		0.527			1.742		1.090			1.000	0.763	3.955	
	EFFCY											1.311	1.000		
VS	VNDEX	2.566		0.189			0.372		0.282			0.253		1.00	0.518
	VNDTK													1.929	1.00
Consistency		0.0049		0.0175			0.0012		0.0093			0.0038		0.0025	
Consensus		80.4%		70.4%			88.1%		79.2%			84.6%		86.2%	

The single-lined cells show the pairwise comparisons for the level 2 factors. For example, a comparison of cost (CO) and Functionality (FU) shows a value of 0.140, but functionality against cost shows a value of 7.148. This implies that functionality is considered more important than cost by a factor of 7.148 out of 9. The diagonal (double-lined) boxes are the level 3 pairwise comparison matrices. The first is the PWC matrix for the cost factor, while the second diagonal box is the PWC matrix for the functionality factor. Within the functionality factor, ease of customization (CUST) is almost equally valued as interoperability and completeness (NTRP) – 1.122 and 0.891 respectively.

Data for the pairwise comparison matrices were obtained from twenty domain experts, who were involved in software projects, either as members of the project management team, or involved in the software acquisition decision. The data were analyzed using Expert Choice software, and AHP templates obtained from http://bpmsg.com. It was important to determine the level of group consensus among the raters. The last row of Table 2 shows group consensus of 70% and above, which validates the utility of the evaluations. One other measure that is important, is the consistency ratios of the pairwise comparison matrices. The ability of AHP to test for consistency is one of the method's greatest strengths. The AHP view of consistency is based on the idea of cardinal transitivity. For example, if criteria A is twice as important as criteria B, and criteria B is three times as important as criteria C, then it would imply (by perfect cardinal consistency) that criteria A be considered six times more important than criteria C. If the domain experts (participants) judge criteria A to be less important than criteria C, it implies that a judgmental error exists, and the prioritization matrix is inconsistent. Our results show consistency ratios which meet the Saaty (1977) threshold of 0.1.

Step three – synthesis

This involves the computation of eigenvalues and the eigenvector. Synthesis yields the percentage of relative priorities, which is expressed in a linear form to give the eigenvector. The implication of the eigenvector is that it expresses the relative importance of an attribute over another in the minds of the decision maker. The eigenvalues and eigenvector provide a means of obtaining linear relationships among the evaluation variables. Expert Choice was used to synthesize the pairwise comparison judgments. It involves the computation of the eigenvector, which presents linear relationships among the evaluation variables; thus, establishing the priority model.

The level 2 software evaluation criteria give an eigenvector, λ_1, while the level 3 criteria produce the eigenvector, λ_2 for each factor, and the level 4 criteria produce the eigenvector, λ_3 for each sub factor (variable). λ_1, combines with the column vector of level 2 factors to give the project evaluation factor index for level 2 criteria ($PEFI_1$), while λ_2, combines with the column vector of the level 3 sub factors to give the project evaluation factor index for level 3 criteria ($PEFI_2$) and λ_3 combines with the column vector of the level 4 variables to give the evaluation factor index for level 4 criteria ($PEFI_3$) as shown in (3), (4), (5). Thus, it is possible to evaluate the software project at various levels of factor abstractions. At a higher level, the evaluation could involve just the level 2 criteria (cost, functionality, flexibility, usability, quality, and vendor services). Alternatively, an organization may decide to evaluate their software project in terms of the level 3 factors (Module Content, User's Experience, Vendors Technical Know-how, Ease of Customization, Vendor Experience, System Adaptability, User Interest, Interoperability and Completeness, Reliability, Organizational Budget, Ease of Use, Integration, Cost of Implementation, System Efficiency), or in terms of the 44 level 4 sub-factors/variables.

$$PEFI_1 = \lambda_1 A_1 \tag{3}$$

where A_1 is the column vector of the level 2 criteria. This gives the $PEFI_1$ as:

$$PEFI_{1(GOAL)} = 0.035(COS) + 0.353(FXN) + 0.141(FLX) + 0.194(USA) + 0.22(QUA) + 0.06(VES) \tag{4}$$

$$PEFI_2 = \lambda_2 A_2 \tag{5}$$

where A_2 is the column vector of the level 3 criteria. This gives the $PEFI_2$ as:

$$\left.\begin{array}{l} PEFI_{2(CO)} = 0.541(BUDG)+0.453(COST) \\ PEFI_{2(FU)} = 0.143(MODL)+0.425(CUST)+0.431(NTRP) \\ PEFI_{2(SF)} = 0.524(ADPT)+0.476(NTRP) \\ PEFI_{2(US)} = 0.45(USINT)+0.098(EXPR)+0.449(EASE) \\ PEFI_{2(QA)} = 0.433(RELB)+0.567(EFFCY) \\ PEFI_{2(VS)} = 0.341(VDEX)+0.659(VDTK) \end{array}\right\} \tag{6}$$

If the evaluator decided to utilize the level 3 factors as the unit of evaluation, then the evaluation weights for the factors would be as shown in Table 3, which is a linear relationship that takes in the qualitative evaluation of the software project on a numeric scale, to produce an overall evaluation index. The summation of the priority weights is 1.

Table 3: Level 3 Factors Priority Weights

Level 1: Goal	Level 2: Criteria	Level 3: Factors	Priority Weight	
IT Project Evaluation	Cost	Organizational Budget (BUDG)	0.0189	0.0348
		Cost of Implementation (COST)	0.0159	
	Functionality	Module Content (MODL)	0.0505	0.3526
		Ease of Customization (CUST)	0.1500	
		Interoperability and Openness (NTRP)	0.1521	
	System Flexibility	System Adaptability (ADPT)	0.0739	0.141
		System Integration (INTG)	0.0671	
	Usability	User Interest (USINT)	0.0873	0.1934
		User Experience (EXPR)	0.0190	
		Ease of Use (EASE)	0.0871	
	Quality	Reliability (RELB)	0.0953	0.22
		System Efficiency (EFFCY)	0.1247	
	Vendor Service	Vendor Experience (VDEX)	0.0205	0.06
		Vendor Technical Know-How (VDTK)	0.0395	

It is evident from Table 3 that Functionality had the highest local priority weight of 35.29%, while Cost and Vendor Services had the least priority weights of 3.47% and 6.26% respectively The Pareto graph (Figure 2) shows the contributions of each factor to the overall evaluation, and clearly points to the importance of Functionality and Quality factors, which in total, contribute to over 50% of the factor weightings in the software evaluation and acquisition decision mix.

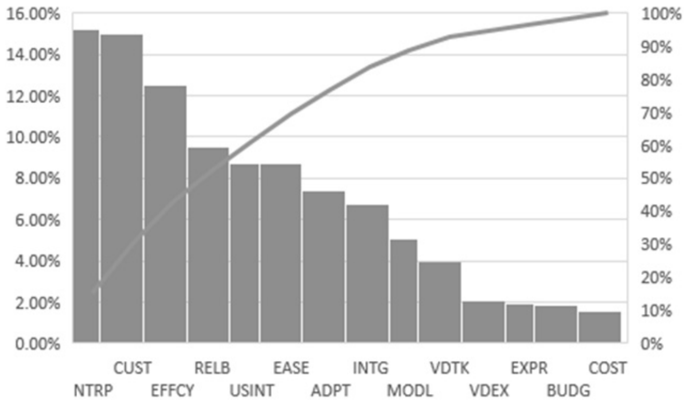


Figure 2: Pareto chart of level 3 factors

Usually, it should be possible to evaluate any software project, based on the level 3 factors; however, organizations that are very process heavy, may decide to further granulate the evaluation into the level 4 sub-criteria. In that case, the project evaluation factor index is given as:

$$PEFI_3 = \lambda_3 A_3 \tag{7}$$

where A_3 is the column vector of the level 4 criteria. This gives the $PEFI_3$ as:

$$\left.\begin{aligned}
PEFI_{3(BUDG)} &= 0.576(SES)+0.229(DEP)+0.196(PRB)\\
PEFI_{3(COST)} &= 0.67(INI)+0.11(LIC)+0.22(COH)\\
PEFI_{3(MODL)} &= 0.038(NMD)+0.411(NMS)+0.152(NIM)+0.184(NOW)+0.215(POR)\\
PEFI_{3(CUST)} &= 0.052(CUF)+0.487(CUR)+0.197(INT)+0.264(CST)\\
PEFI_{3(NTRP)} &= 0.090(AOM)+0.595(COT)+0.315(INO)\\
PEFI_{3(ADPT)} &= 0.088(OPE)+0.436(PIS)+0.476(AIS)\\
PEFI_{3(INTG)} &= 0.312(PLI)+0.197(UPG)+0.491(EIT)\\
PEFI_{3(USINT)} &= 0.481(UIN)+0.519(WTU)\\
PEFI_{3(EXPR)} &= 0.052(UEP)+0.407(PQU)+0.393(FRU)+0.147(LOU)\\
PEFI_{3(EASE)} &= 0.552(EGI)+0.448(EOP)\\
PEFI_{3(RELB)} &= 0.50(STS)+0.50(RAB)\\
PEFI_{3(EFFCY)} &= 0.077(MAS)+0.644(SET)+0.279(SOC)\\
PEFI_{3(VNDEX)} &= 0.114(LOE)+0.184(WPV)+0.701(PBE)\\
PEFI_{3(VNDTK)} &= 0.055(EOI)+0.159(GIS)+0.437(TBS)+0.349(ITK)
\end{aligned}\right\} \tag{8}$$

Combining (4), (6) and (8), we produce the following Aggregate Project Evaluation Factor Index (APEFI) , which serves as a linear equation for the evaluation of any given software project:

APEFI = 0.011(SES) + 0.004(DEP) + 0.004(PRB) + 0.011(INI) + 0.002(LIC) + 0.003(COH) + 0.002(NMD) + 0.021(NMS) + 0.008(NIM) + 0.009(NOW) + 0.011(POR) + 0.008(CUF) + 0.073(CUR) + 0.030(INT) + 0.04(CST) + 0.014(AOM) + 0.091(COT) + 0.048(INO) + 0.007(OPE) + 0.032(PIS) + 0.035(AIS) + 0.021(PLI) +0.013(UPG) + 0.033(EIT) + 0.043(UIN)+0.045(WTU) + 0.001(UEP) + 0.008(PQU) + 0.007(FRU) + 0.003(LOU) + 0.048(EGI) + 0.039(EOP)+0.048(STS) + 0.048(RAB) + 0.01(MAS) + 0.080(SET) + 0.0348(SOC) + 0.002(LOV) + 0.004(WPV) + 0.014(PBE) + 0.002(EOI) + 0.006(GIS) + 0.017(TBS) + 0.014(ITK) (9)

3 Model evaluation and results

To test the evaluation system, we visited ten organizations to identify the individual(s) who had the competence to make decisions on software projects based on: 1) their positions in the respective organizations, and 2) involvement in information systems projects. In some organizations, we identified more than one person who had the competence. In such situations, an individual was requested to coordinate the group decision process in arriving at one group evaluation for their existing major software system or a software project being proposed by the organizations. Our goal was to collect simple data that would be useful in testing the utility of the AHP model for evaluation of software projects. A survey was administered to each organization that agreed to participate in the evaluation exercise. The survey utilized the 44 level 4 evaluation variables and provided the evaluators with a five-point linguistic Likert scale (poor, fair, good, very good, excellent). The snap shot of results of the final evaluation are shown in Appendix 1.

The APEFI was applied to data in Appendix 1 to obtain the final evaluation of each organization’s software system. This was done through the following:

$$S_i = \sum_{k=1}^{n} R_{ik} X_k \tag{10}$$

where: S_i is the evaluation score of organization i,
R_{ik} is the rating of organization i on variable k,
X_k is the APEFI value of variable k,
n is he number of level 4 evaluation variables.

An illustration using organization 1 is given below:

APEFI = 0.011(4) + 0.004(3) + 0.004(4) + 0.011(3) + 0.002(4) + 0.003(4) + 0.002(4) + 0.021(5) + 0.008(3) + 0.009(4) + 0.011(3) + 0.008(3) + 0.073(3) + 0.030(3) + 0.04(4) + 0.014(5) + 0.091(4) + 0.048(4) + 0.007(4) + 0.032(2) + 0.035(3) + 0.021(2) +0.013(5) + 0.033(4) + 0.043(4)+0.045(3) + 0.001(4) + 0.008(5) + 0.007(3) + 0.003(5) + 0.048(4) + 0.039(4)+0.048(4) + 0.048(4) + 0.01(4) + 0.080(3) + 0.0348(4) + 0.002(4) + 0.004(4) + 0.014(3) + 0.002(3) + 0.006(4) + 0.017(3) + 0.014(3) = 3.444.

The summary of the evaluation results, before and after the application of the AHP model is presented in Table 4.

Table 4: Evaluation summaries

Organization	Pre-AHP Evaluation		Post-AHP Evaluation	
	total score	rank	total score	rank
1	165	3	3.47	5
2	167	2	3.58	4
3	141	8	2.91	9
4	140	9	3.17	6
5	201	1	4.43	1
6	164	4	3.94	2
7	126	10	2.81	10
8	143	7	3.00	8
9	145	6	3.06	7
10	159	5	3.69	3

The results show that the application of AHP refined the initial evaluations by taking into cognizance, the priorities attached to the evaluation variables. After the application of AHP, organization 4, 6, and 10 had improvements in the rankings of their information systems, while organizations 1, 2, 3, 8 and 9 saw a drop in the ranking of their information systems; organizations 5 and 7 did not see any change in the ranking of their system; organization 5 being the best and organization 7 being the worst. While the primary aim of this paper was not to rank individual organization's software systems, we used this to demonstrate the utility of the system in ranking various vendors' proposals. Organizations could also apply the AHP model in the evaluation of their existing information systems.

The AHP model also reveals the relative importance of various factors in the software/vendor evaluation process. The results show that functionality, quality and usability are very critical in the software evaluation decision, while cost and vendor service rank low in decision process. We briefly discuss the findings, relating to the level 2 evaluation criteria.

Cost

Cost had a priority weight of 3.48%, the least weight among the evaluation factors, which implies that the cost of a software project proposal is not considered very critical in the evaluation and selection of software projects presented by vendors. This aligns with the research presented in Jain et al. (2008) and in Khan et al. (2011) where cost was not found to be the driving factor in the process of selecting vendors for software development. Also, in Lai et al. (1999) cost was found to be relatively unimportant in software selection process. Client organizations need to lay more emphasis on the other evaluation factors with the view to selecting an appropriate software project proposal that best satisfy the needs of the client organization (i.e. end users), and improves the quality of work they perform (Lai et al., 1999). It is worthy to note that organizations are more willing to accept the cost of software project in so far as the functionality and quality meet their requirements (Khan et al., 2011). According to Stefanou (2001), client organization works within its budget; however, if a software solution is found to provide the organization with best service, there will be the need to strike a cost balance in order not to play down on other software factors such as quality, functionality, effectiveness, efficiency etc.

Functionality

Functionality had the highest priority weight of 35.26%, which makes it a key factor in choosing software solutions for organizations. Functionality relates with the functional requirements of the client organization; thus, if a software project proposal, based on the client's judgments, has a high functionality rating, there is the tendency that such software would likely meet most of the organization's functional requirements. Khan et al. (2011) and Lai et al. (1999) considered excellent functional behavior of a given software solution to be key determinants in the software selection process. In this study, the following sub-factors were considered under functionality: module content, ease of customization, interoperability and openness. The highest level 3 criteria priority weights were recorded by ease of customization (15.0%) and interoperability and openness (15.2%). Many organizations are moving toward enterprise resource planning (ERP) software systems, and consider software interoperability to be a very crucial attribute in software evaluation (Bertram et al., 2016; Keil and Tiwana, 2006).

System Flexibility

The priority weight for system flexibility was 14.1%. The elements considered under flexibility were *system adaptability* (ADPT) (0.0739) and *system integration* (INTG) (0.0671). These are non-functional requirements that have some level of significance in the selection of software solutions for company services. The importance of system flexibility, especially in ERP environments has been emphasized in Atal et al. (2016), Khan et al. (2011), Uzoka et al. (2008), Uzoka (2009) and aligns with the need to have systems that could easily be adapted to meet the dynamic needs of organizations.

Usability

This is a non-functional requirement having priority value of 19.34%. The priority attached to usability underscores the need for user friendly systems, which has been severally emphasized in literature such as Abdelaziz et al. (2016), Lewis (2014), Uzoka (2009). In emphasizing the importance of usability of software systems in the overall performance of the organization, the authors in Engelbrecht et al. (2017) noted that business managers often underestimate its impact on processes and people; and further suggested that organizations embrace the culture of usability testing and training, especially with enterprise systems. It was recommended in Lewis (2014) that software practitioners should emphasize iterative formative (rather than summative) usability testing, using one of the available standard usability instruments, as a means of improving objective and perceived usability. The usability factors identified in our study [User Interest (0.0873), User Experience (0.0190) and Ease of Use (0.0871)] have been severally recognized in technology adoption as having significant impacts on the potential user's intention to adopt technology (including software).

Quality

Software quality has one of the high priority weights (22%) in the AHP model, pointing to its importance in the software evaluation process. Since organizations are prepared to invest in information systems, they would obviously expect high value from such investments, especially in a global software landscape that is characterized by many vendors and products. Finding a product that is suitable for the organization's needs is a key challenge in the software selection process. Our study further emphasizes the quality factors of reliability and efficiency, which account for 9.53% and 12.47% respectively of the level 3 evaluation factor weights. The importance of quality factors in the software evaluation process is emphasized in ALMohiza et al. (2016), while Uzoka et al. (2008) found quality to be of utmost importance in the selection of ERP systems.

Vendor Service

The experience of vendors vis-à-vis their technical know-how cannot be underestimated while considering vendors to choose for a given project (Al-Harbi, 2001; Jain et al., 2008). The priority weight of vendor service attribute is the second lowest compared to other factors (priority weight = 0.06). This presents an interesting result, which is at variance with previous studies that emphasized the importance of vendor's support and technical know-how in the software selection process (Jain et al., 2008; Uzoka et al., 2008). Similar to cost, vendor service seems not to matter so much, especially because many medium to large organizations have in-house technical competence to manage and maintain their software system.

4 Limitations

Data sample used for this study was small, which could impact on the generalization of our model. Moreover, we did not consider implementation process for selected software project. A wrong implementation process by the client organization could be responsible for failed software solution in an organization. Therefore, a future research could incorporate an implementation framework that focuses on the cause and effect relationship that software project selection process activities/results have on the implementation process. We also note that the AHP model was developed using pairwise comparison information provided by experts in software project management. This model could be better generalized with a larger number of domain experts in various project environments, with a good diversity of constraints.

5 Conclusion

This study provides organizations with valuable knowledge that would prompt them to make significant changes in the manner in which they currently proceed with the selection of any software project proposal, which in turn, could result in substantial savings in terms of economics (actual costs, time, and improved administrative procedures). The proposed system will enable: 1) active involvement of the users in the client organization in the software selection and development process; 2) the software vendor to have intimate relationship with the entire management of the client organization, thus minimizing potential challenges during users' requirements gathering; 3) users to easily accept, adopt and understand and use the software when deplored. This study has helped to provide significant criteria that management of organizations could utilize to evaluate the IT solutions proposed by IT vendors and these criteria align with the terms in the IT procurement policy presented in IT Procurement Policy (2005). It helps to enrich the knowledge of client organizations on the theoretical and practical principles of the selection process for valuable IT application package with the ultimate goal of end-user satisfaction. AHP through its structured hierarchy of decision levels and pair wise comparison of elements for value judgment is more effective than utility models and scoring charts in working with semi-quantitative data as realistic inputs to the priority-setting agenda. They help to overcome in a significant way, the fuzzy nature of quantitative information related to deliverable, logistics, and outcome. In the resource constrained situation of the developing countries, AHP provides a vital tool to select and rank projects based on judgmental evaluation through peer ratings. AHP provides a comprehensive and rational framework for structuring a decision problem, for representing and quantifying its elements, for relating those elements to overall goals, and for evaluating alternative solutions. It also considers a set of evaluation criteria, and a set of alternative scenarios among which the best decision is to be made. It generates a weight for each evaluation criterion and scenario according to the information provided by the decision maker. AHP is effective in dealing with complex decision making because it reduces complex decisions to a series of pairwise comparisons and reduces the bias in the decision-making process because it also checks the consistency of the decision maker's evaluations. The system proposed in this study could be scaled with more data, to a generalizable level that could serve as a standard model for software system evaluation and selection.

References

Abdelaziz T.M., Maatuk A.M., Rajab F. (2016), *An Approach to Improvement the Usability in Software Products*, International Journal of Software Engineering & Applications (IJSEA), 7(2), 11-18.

Akinnuwesi B.A., Uzoka F.M.E. (2017), *Assessment of Software Project Proposal using Analytical Hierarchy Process: A Framework*, Journal of Research and Review in Science, 4(1), Publication of Faculty of Science, Lagos State University, Lagos, Nigeria, 44-55.

Akinnuwesi B.A., Uzoka F.M., Olabiyisi S.O., Omidiora E.O. (2012), *A Framework for User--centric Model for Evaluating the Performance of Distributed Software System Architecture*, Expert Systems with Applications, 39(10), 9323-9339.

Akinnuwesi B.A., Uzoka F.M., Olabiyisi S.O., Omidiora E.O., Fiddi P. (2013), *An Empirical Analysis of End-user Participation in Software Development Projects in a Developing Country Context*, The Electronic Journal of Information Systems in Developing Countries, 58(6), 1-25.

Akinnuwesi B.A., Uzoka F.M., Osamiluyi A.O. (2013c), *Neuro-Fuzzy Expert System for Evaluating the Performance of Distributed Software System Architecture*, Expert Systems with Applications, 40(9), 3313-3327.

Al-Harbi K.M.A.-S. (2001), *Application of the AHP in Project Management*, International Journal of Project Management, 19(1), 19-27.

ALMohiza M.A., Sparto P.J., Marchetti G.F., Delitto A., Furman J.M., Miller D.L., Whitney S.L. (2016), *A Quality Improvement Project in Balance and Vestibular Rehabilitation and Its Effect on Clinical Outcomes*, Journal of Neurologic Physical Therapy, 40(2), 90-99.

Atal V., Bar T., Gordon S. (2016), *Project Selection: Commitment and Competition*, Games and Economic Behavior, (96), 30-48.

Bertram V., Manhart P., Plotnikov D., Rumpe B., Schulze C., Wenckstern M. von (2016), *Infrastructure to Use OCL for Runtime Structural Compatibility Checks of Simulink Models*, Modellierung, 109-116.

Bertram V., Roth A., Rumpe B., Wenckstern M. von (2016), *Extendable Toolchain for Automatic Compatibility Checks*.

Bredillet C.N. (2010), *Blowing Hot and Cold on Project Management*, Project Management Journal, 41(3), 4-20.

Buss M.D.J. (1983), *How to Rank Computer Projects*, Harvard Business Review, 61(1), 118-125.

Chau P.Y. (1995), *Factors Used in the Selection of Packaged Software in Small Businesses: Views of Owners and Managers*, Information & Management, 29(2), 71-78.

Davis L., Williams G. (1994), *Evaluating and Selecting Simulation Software Using the Analytic Hierarchy Process*, Integrated Manufacturing Systems, 5(1), 23-32.

Engelbrecht J., Johnston K.A., Hooper V. (2017), *The Influence of Business Managers' IT Competence on IT Project Success*, International Journal of Project Management, 35(6), 994-1005.

Hares J., Royle D. (1994), *Measuring the Value of Information Technology*, John Wiley and Sons, Inc., New York.

He J., Butler B.S., King W.R. (2007), *Team Cognition: Development and Evolution in Software Project Teams*, Journal of Management Information Systems, 24(2), 261-292.

Hoffer J., George J., Valacich J. (2016). *Modern Systems Analysis and Design*, (8th Edition), Pearson, Essex.

Hwang C.-L., Yoon K. (2012), *Multiple Attribute Decision Making: Methods and Applications a State-of-the-art Survey*, Springer Verlag, New York.

Ika L.A. (2009), *Project Success as a Topic in Project Management Journals*, Project Management Journal, 40(4), 6-19.

IT Procurement Policy 2005, Information Technology (IT) Procurement Policy.

Jadhav A.S., Sonar R.M. (2009), *Evaluating and Selecting Software Packages: A Review*, Information and Software Technology, 51(3), 555-563.

Jadhav A.S., Sonar R.M. (2011), *Framework for Evaluation and Selection of the Software Packages: A Hybrid Knowledge Based System Approach*, Journal of Systems and Software, 84(8), 1394-1407.

Jain V., Benyoucef L., Bennett D., Deep A., Guttridge P., Dani S., Burns N. (2008), *Investigating Factors Affecting ERP Selection in Made-to-order SME Sector*, Journal of Manufacturing Technology Management, 19(4), 430-446.

Kahraman C., Büyüközkan G., Ruan D. (2004), *Measuring Software Development Value Using Fuzzy Logic*, [in:] D. Ruan, X. Zeng (eds.), *Intelligent Sensory Evaluation: Methodologies and Applications*, Springer, 285-308.

Keil M., Tiwana A. (2006), *Relative Importance of Evaluation Criteria for Enterprise Systems: A Conjoint Study*, Information Systems Journal, 16(3), 237-262.

Khaddaj S., Horgan G. (2004), *The Evaluation of Software Quality Factors in Very Large Information Systems*, Electronic Journal of Information Systems Evaluation, 7(1), 43-48.

Khan S.U., Niazi M., Ahmad R. (2011), *Factors Influencing Clients in the Selection of Offshore Software Outsourcing Vendors: An Exploratory Study Using a Systematic Literature Review*, Journal of Systems and Software, 84(4), 686-699.

Lai V.S., Trueblood R.P., Wong B.K. (1999), *Software Selection: A Case Study of the Application of the Analytical Hierarchical Process to the Selection of a Multimedia Authoring System*, Information & Management, 36(4), 221-232.

Lewis J.R. (2014), *Usability: Lessons Learned… and Yet to be Learned*, International Journal of Human-Computer Interaction, 30(9), 663-684.

Liberatore M.J., Pollack-Johnson B. (2003), *Factors Influencing the Usage and Selection of Project Management Software*, IEEE Transactions on Engineering Management, 50(2), 164-174.

Maidamisa A.A., Ahmad R., Aziz M.I.A. (2012), *Selecting the Appropriate Decision Alternatives Using SM and AHP*, International Journal of Computer Applications, 37(10), 14-18.

Melone N., Wharton T. (1984), *Strategies for MIS Project Selection*, Journal of Systems Management, 35(2), 26-33.

Meso P., Jain R. (2006), *Agile Software Development: Adaptive Systems Principles and Best Practices*, Information Systems Management 23(3), 19-30.

Molla A., Licker P.S. (2005), *eCommerce Adoption in Developing Countries: A Model and Instrument*, Information & Management, 42(6), 877-899.

Muralidhar K., Santhanam R., Wilson R.L. (1990), *Using the Analytic Hierarchy Process for Information System Project Selection*, Information & Management, 18(2), 87-95.

Nandi S., Paul S., Phadtare M. (2011), *An AHP-based Construction Project Selection Method*, Decision, 38(1), 91-118.

Niessink F., Van Vliet H. (1999), *Measurements Should Generate Value, Rather than Data* [Software Metrics], Software Metrics Symposium, 1999, Proceedings. Sixth International, IEEE, 31-38.

Oppermann R., Reiterer H. (1997), *Software Evaluation Using the 9241*, Behaviour and Information Technology, 16(4/5), 232-245.

Patanakul P., Iewwongcharoen B., Milosevic D. (2010), *An Empirical Study on the Use of Project Management Tools and Techniques Across Project Life-cycle and Their Impact on Project Success*, Journal of General Management, 35(3), 41-65.

Rivard S., Dupré R. (2009), *Information Systems Project Management in PMJ: A Brief History*, Project Management Journal, 40(4), 20-30.

Rouyendegh B.D., Erkan T.E. (2011), *ERP System Selection by AHP Method: Case Study from Turkey*, International Journal of Business and Management Studies, 3(1), 39-48.

Saaty T.L. (1977), *A Scaling Method for Priorities in Hierarchical Structures*, Journal of Mathematical Psychology, 15(3), 234-281.

Saaty T.L. (1990), *How to Make a Decision: The Analytic Hierarchy Process*, European Journal of Operational Research, 48(1), 9-26.

Saaty T.L. (2008), *Decision Making with the Analytic Hierarchy Process*, International Journal of Services Sciences, 1(1), 83-98.

Seong Leem C., Yoon Y. (2004), *A Maturity Model and an Evaluation System of Software Customer Satisfaction: The Case of Software Companies in Korea*, Industrial Management & Data Systems, 104(4), 347-354.

Sherer S.A. (2004), *IS Project Selection: The Role of Strategic Vision and IT Governance*, System Sciences, Proceedings of the 37th Annual Hawaii International Conference on, IEEE.

Stefanou C.J. (2001), *A Framework for the Ex-ante Evaluation of ERP Software*, European Journal of Information Systems, 10(4), 204-215.

Uzoka F.M., Abiola R.O., Nyangeresi R. (2008), *Influence of Product and Organizational Constructs on ERP Acquisition Using an Extended Technology Acceptance Model*, International Journal of Enterprise Information Systems, 4(2), 67-70.

Uzoka F.M., Akinnuwesi B.A., Oluwole N., Adekoya, A.F., Egbekunle O.Y. (2016), *Identifying Factors for Evaluating Software Project Proposals*, International Journal of Quality Engineering and Technology, 6(1/2), 93-114.

Uzoka F.M., Fedoruk A., Osakwe C., Osuji J., Gibb K. (2013), *A Multi-criteria Framework for Assessing Scholarship Based on Boyer's Scholarship Model*, Information Knowledge Systems Management, 12(1), 25-51.

Uzoka F.M.E. (2009), *Examining the Effects of TAM Constructs on Organizational Software Acquisition Decision*, Information Resources Management Journal (IRMJ), 22(3), 40-58.

Vargas R.V. (2010), *Using the Analytic Hierarchy Process (AHP) to Select and Prioritize Projects in a Portfolio*, PMI Global Congress 2010 North America, Available: http://www.ricardo-vargas.com/wp-content/uploads/downloads/articles/ricardo_vargas_ahp_project_selection_en.pdf, Access Date: July 30, 2016, 2010.

Verner J., Sampson J., Cerpa N., Bleistein S. (2009), *What Factors Lead to Software Project Failure and Whose Fault Was It?*, Ingénierie des Systèmes d'Information, 14(4), 55-75.

Verville J., Halingten A. (2002), *An Investigation of the Decision Process for Selecting an ERP Software: The Case of ESC*, Management Decision, 40(3), 206-216.

Verville J., Halingten A. (2003), *A Six Stage Model of the Buying Process for ERP Software*, Industrial Marketing Management, (32), 585-594.

Wei C.-C., Chien C.F., Wang M.-J.J. (2005), *An AHP-based Approach to ERP System Selection*, International Journal of Production Economics, 96(1), 47-62.

Wei C.-C., Wang M.-J.J. (2004), *A Comprehensive Framework for Selecting an ERP System*, International Journal of Project Management, 22(2), 161-169.

Williams J.B., Gibbon M., First M.B., Spitzer R.L., Davies M., Borus J., Howes M.J., Kane J., Pope H.G., Rounsaville B. (1992), *The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite Test-retest Reliability*, Archives of General Psychiatry, 49(8), 630-636.

Yoon K.P., Hwang C.-L. (1995), *Multiple Attribute Decision Making: An Introduction*, SAGE Publications, London.

Zielsdorff C., McGinnis C., Daim T., Basoglu N. (2010), *Enterprise Resource Planning (ERP) Selection for a Medical Devices Manufacturing Company*, International Journal of Business Information Systems, 6(2), 265-293.

Appendix 1

Raw Evaluation of Organizational Software Systems

Variables	GPW	DM1	DM2	DM3	DM4	DM5	DM6	DM7	DM8	DM9	DM10
A1	0.011	4	4	4	3	3	3	3	4	3	4
B1	0.004	3	3	3	3	4	3	3	4	2	3
C1	0.004	4	4	4	3	5	3	4	4	4	4
A2	0.011	3	4	4	3	5	3	4	4	3	3
B2	0.002	4	4	4	2	5	3	4	3	3	4
C2	0.003	4	3	3	2	5	3	4	3	4	3
A3	0.002	4	4	5	3	5	3	4	5	5	4
B3	0.021	5	5	4	3	5	3	4	5	5	4
C3	0.008	3	3	3	3	4	3	3	3	2	3
D3	0.009	4	4	4	4	4	3	3	5	3	4
E3	0.011	3	4	3	4	4	3	3	3	3	4
A4	0.008	3	4	3	4	3	5	3	3	3	4
B4	0.073	3	3	2	3	3	5	3	3	3	3
C4	0.030	3	3	3	3	5	5	3	3	3	3
D4	0.040	4	4	2	3	5	5	3	4	3	4
A5	0.014	5	5	3	4	5	5	3	4	4	5
B5	0.091	4	4	3	3	5	5	3	3	3	4
C5	0.048	4	4	2	4	5	4	3	4	2	4
A6	0.007	4	5	3	3	4	4	3	5	4	4
B6	0.032	2	3	2	4	3	4	2	2	3	2
C6	0.035	3	3	2	4	4	4	2	2	3	3
A7	0.021	2	3	2	4	4	3	2	2	2	3
B7	0.013	5	4	3	2	4	3	2	3	3	4
C7	0.033	4	3	3	2	5	3	3	3	4	4
A8	0.043	4	2	3	2	5	3	3	4	3	4
B8	0.045	3	3	3	2	5	3	2	3	1	5
A9	0.001	4	4	3	3	5	3	3	4	3	3
B9	0.008	5	5	4	4	5	4	4	5	4	4
C9	0.007	3	4	3	3	4	4	3	3	3	3
D9	0.003	5	5	4	3	4	5	3	4	5	5
A10	0.048	4	5	3	3	4	4	3	4	4	5
B10	0.039	4	5	4	3	5	5	3	3	4	5
A11	0.048	4	4	5	4	5	5	3	3	3	4
B11	0.048	4	5	4	4	5	5	3	3	4	4
A12	0.010	4	5	4	4	5	4	3	3	4	4
B12	0.080	3	3	2	4	5	4	3	2	3	4
C12	0.035	4	4	3	4	5	4	3	3	4	4
A13	0.002	4	4	2	3	5	5	3	2	3	3
B13	0.004	4	3	3	3	5	3	3	3	3	2
C13	0.014	3	4	4	3	5	3	3	3	3	3
A14	0.002	3	2	2	3	5	3	2	2	3	2
B14	0.006	4	4	4	3	5	3	3	3	4	3
C14	0.017	3	3	3	3	5	3	3	2	3	2
D14	0.014	3	3	4	3	5	3	3	3	4	3
TOTAL	1.005	165	167	141	140	201	164	126	143	145	159
RANK		3	2	8	9	1	4	10	7	6	5

Key: GPW = Global Priority Weight, DM1…DM10 represent the system evaluations by the ten decision makers (DM) in the sampled organizations.

MULTIPLE CRITERIA DECISION MAKING
Vol. 14 2019

Tea Vizinger[*]
Janez Žerovnik[**]

ROBUST OPTIMISATION METAHEURISTICS FOR THE INVENTORY-ALLOCATION PROBLEM

DOI: 10.22367/mcdm.2019.14.08

Abstract

As an example of a successful application of a relatively simple metaheuristics for a stochastic version of a multiple criteria optimisation problem, the inventory-allocation problem is discussed. Stochastic programming is introduced to deal with the demand of end consumers. It has been shown before that simple metaheuristics, i.e., local search may be a very competitive choice for solving computationally hard optimisation problems. In this paper, robust optimisation approach is applied to select more promising initial solutions which results in a significant improvement of time complexity of the optimisation algorithms. Furthermore, it allows more flexibility in choosing the final solution that need not always be minimising the sum of costs.

Keywords: robust optimisation, local search, stochastic programming, distribution.

1 Introduction

New metaheuristics paradigms are introduced and are getting popular in recent years and in recent decades because NP-hard problems provide challenging optimisation tasks (Talbi, 2009; Aarts and Lenstra, 1997; Sorensen et al., 2016). Despite of the dramatic increase of available computational power, the need for heuristic methods remains because also the size of practical problems to solve increases. While it is widely accepted that the most successful heuristics are those that use the very properties of a particular problem and the domain, it is

[*] University of Maribor, Faculty of Logistics, Celje, Slovenia, e-mail: tea.vizinger@um.si, ORCID: 0000-0003-1539-9600.
[**] University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia, e-mail: janez.zerovnik@fs.uni-lj.si, and Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia, e-mail: janez.zerovnik@imfm.si, ORCID: 0000-0002-6041-1106.

not so commonly believed that simple metaheuristics are expected to be overperfoming the more complicated ones (for further discussion, see Žerovnik, 2015, 2003). In the talk given by one of the authors at the 8th International Workshop on Multiple Criteria Decision Making, arguments and examples were provided supporting the claim. A very general theoretical argument (Ferreira and Žerovnik, 1993) is that any local search asymptotically outperforms the often used heuristic simulated annealing (Kirkpatrick et al., 1983), on any problem (!). We mention here three examples. The first example of a simple local search type heuristic is the "remove and reinsert" heuristic that has been applied to the traveling salesman problem (Brest and Žerovnik, 1999), the probabilistic traveling salesman problem (Žerovnik, 1995), the resource-constrained project scheduling problem (Pesek et al., 2007) and the job shop scheduling problem (Zupan et al., 2016). The second example is the Petford-Welsh algorithm (Petford and Welsh, 1989), a heuristic for graph 3-colouring based on the antivoter model (Donelly and Welsh, 1983), that has later been applied to various generalised colouring problems including the k-colouring (Žerovnik, 1994), frequency assignment (Ubeda and Žerovnik, 1997), and very recently to the clustering problem (Ikica et al, 2019). For details of the close relation of the Petford-Welsh algorithm to the Boltzman machine and the simulated annealing algorithm, see Žerovnik (2000). The last example that will be elaborated in more detail in this paper is the application of local search heuristics to the inventory--allocation problem.

The rest of the paper is organised as follows. Inventory allocation in a supply chain is introduced in section 2. In section 3, the formal definition of the problem is given and the robust optimisation approach that extends our previous heuristics is described. The new approach allows a sizeable improvement in computation speed, as shown by the results of a computational experiment on a realistic example described in section 4. Conclusions are given in section 5.

2 Inventory allocation in a supply chain

A typical retail supply chain consists of one or several warehouses that distribute products to several stores, which have to deal with stochastic demand patterns. The idea is to align the decisions, reflecting the ordering policies that in retail companies are usually taken independently by several decision makers. On the one hand, we are dealing with warehouse managers, whose orders are naturally based on the price and availability of a product. On the other hand, we have store managers, whose orders are usually based on the actual requirements of the merchant. The ordered quantities from the external suppliers depend therefore on the stock market prices and not on the actual requirements. Of course there are many situations with higher or lower stock levels, causing overstocking effects or lost sales (state of stock-out with possible lower sale realisation).

In our previous research paper (Vizinger and Žerovnik, 2019) we have presented the idea of an on-going optimisation approach. In this approach we first find a tactical plan and then (re)define some strategic and operational decisions. Tactical planning for a chosen period (month, season, etc.) defines the appropriate inventory levels in warehouses and stores, and consequently the allocation of resources among retail facilities. The replenished quantities are defined on the operational level using the difference between the actual inventory and the pre-defined maximum level of a certain store inventory. Since the tactical plan already determines the necessary stock levels, the demand of the warehouses becomes more or less deterministic. With precise tactical planning, retailers may be able to contract constant supply quantities, which may result in a lower unit price and the corresponding higher profit. At the stores' side of the supply chain, precisely pre-defined inventory levels prevents stock accumulation, which may result in increased product quality and a related customer service level.

The model for product flow coordination in a retail supply chain that was introduced in Vizinger and Žerovnik (2019) considers optimisation of three criteria (distribution costs, overstocking effects and lost sales). While the costs are estimated on the basis of the expected demand, only the distribution cost can be calculated for the quantities predicted. On the other hand, both types of supply risk (overstocking effects and lost sales) are unknown a priori. Although the last two cannot appear at the same time (they are mutually exclusive), it is reasonable to deal with each of the three costs separately. On the one hand, we consider a single-item lot sizing problem and, on the other hand, a resource allocation problem. The stochastic model introduced in Vizinger and Žerovnik (2019) is the first that tackles the coordination problem at the tactical level of planning. The inventory-allocation problem may be seen as a generalized material flow problem, where the goal is to minimise the distribution costs of goods delivered from some supply points to a number of destination points (Anholcer, 2016). Below we refer to the combined inventory-routing problem, as there exist similarities to our inventory-allocation problem.

For the inventory-routing problem, the literature introduces mostly the use of mixed integer optimisation, multi-objective optimisation and stochastic programming (e.g., Liu and Papageorgiou, 2013; Grossman and Guillén--Gosálbez, 2010). The idea is to find an appropriate policy with minimal costs of distribution, minimal overstocking effects and maximal customer service level.

Liu and Papageorgiou (2013) interpret customer service level as the percentage of customer demand satisfied on time. Lower customer service level therefore causes lost sales or lost customers, and this results in profit loss of the supply chain.

Most of the applied stochastic models are two-stage programs, and are used to deal with demand uncertainties when assigning probability distributions (Franca et al., 2010). A stochastic transportation problem may be transformed into a deterministic one by removing the demand constraints, which are used to introduce a new cost function related to the expected extra cost (resulting in a difference between the delivered amount and the actual demand). Although risk is measured in our paper with the cost metric, the direct and indirect costs should not be summed up, because they have a totally different origin. As opposed to the well-known lot-sizing models or the Newsvendor approach (for coordination of the supply chain flow), we are not limited to use only a simple or a weighted sum of the criteria considered.

Beside stochastic programming and heuristics solution procedures, Grossman and Guillén-Gosálbez (2010) introduce robust optimisation and probabilistic programming. In many cases we are not able to identify the underlying probability distributions or such a stochastic description may simply not exist (Sarimveis et al., 2008). In such a situation it is reasonable to fit a suitable probability distribution for each parameter based on an expert's subjective knowledge derived from past experiences and feelings. Uncertain data are therefore unknown but bounded quantities, while constraints are satisfied for all realisations of the uncertain parameters. In robust programming, not every scenario represents a feasible solution. Once an uncertainty is realised, the solution obtained from robust optimisation ensures that constraints are satisfied with a certain probability.

The optimisation problem that arises from the model is a computationally hard problem. For time prohibitive stochastic programs, the use of heuristic approaches (which provides good feasible solutions) have become very popular. Several versions of the local search heuristics were adopted in Vizinger and Žerovnik (2019), including iterative improvement (a basic form of local search), tabu search and threshold accepting. The best performance was shown by the tabu search heuristic that proved to provide very good solutions on the instances tested. However, the computational time for a single product instances of moderate size is considerable. Even though these calculations are to be performed only occasionally, it is important to have a faster method if possible.

3 Formal definition of the problem

We represent the coordination problem for a retail supply chain product flow as a multi-objective discrete optimisation problem. A typical retail supply chain consists of one or several warehouses $i \in (1, \ldots, I)$ who deliver products to a number of stores $j \in (1, \ldots, J)$, where dc_{ij} is the distribution cost, and x_{ij} is the quantity of the product distributed. There are two types of vertices: a_i represents a given fixed supply available at each origin or warehouse, and C_j represents a fixed inventory holding capacities of stores. In addition, we are given the demand of the stores as random variables b_j. In other words, we model the customers' shopping habits with random variable b_j with some probability distribution that is not known a priori. Here we consider discrete distributions and assume that we are given hypothetical distributions based on past experience (managers' knowledge, information from the system) and/or intuition.

A feasible solution X is given by the matrix

$$X = \left[x_{ij}\right]_{i \in I, j \in J}, \tag{1}$$

where x_{ij} is the amount transferred from warehouse i to store j. A solution X is feasible if it satisfies the inventory holding capacity of stores C_j, and complies with the supply available at each origin or warehouse a_i.

A possible sale realisation is represented by the scenario, described by vector $L = \left[l_j\right]$, where l is the fixed scenario realised at store j.

3.1 Optimisation criteria

Given a scenario L, the cost of overstocking effects OS is calculated as:

$$OS(X, L) = \sum_j (\sum_i x_{ij} - l_j) \cdot c_{OS} \tag{2}$$

and the cost of lost sales LS is defined as:

$$LS(X, L) = \sum_j (l_j - \sum_i x_{ij}) \cdot c_{LS}. \tag{3}$$

In (2) and (3), c_{OS} is the cost of overstocking effects for a unit of product at store j, and c_{LS} is the cost of lost sales for a unit of product at store j. Here we can optimise only the expected values because the optimisation criteria depend on the a priori unknown values of the future sales. The expected cost of overstocking effects is represented as the weighted sum of the costs over all scenarios:

$$E(OS(X)) = \sum_L p(L) \cdot OS(X, L), \tag{4}$$

where $p(L)$ is the probability of scenario L. Similarly, for the expected cost of lost sales:

$$E(LS(X)) = \sum_L p(L) \cdot LS(X, L). \tag{5}$$

As indicated in Introduction, the relationship between the decision criteria is often represented by a (weighted or simple) sum of criteria. However, when defining the general mathematical model, we wish to consider the multi--objective optimisation problem in a more general and somewhat more natural way. The stochastic model and experimental study are described in detail in Vizinger and Žerovnik (2019). Note that the goal function to be minimised in the local search procedure was defined at first as a sum of the criteria. Simple local search heuristics have been shown to provide near optimal solutions of very good quality in a reasonable time (Vizinger and Žerovnik, 2019, 2018). Nevertheless, when larger instances and, in particular, when more products are considered, the computational time may be large, therefore we adopted the robust optimisation approach in order to restrict attention to a subset of promising feasible solutions. In short and roughly speaking robust optimisation here means that we attempt to speed up the optimisation procedure by focusing first on the two criteria modelling the risk and considering the third criterion only in the case when the first two are within reasonable bounds. In this way, costly optimisation of the distribution cost that involves linear programming is avoided. The optimisation criteria remain the same, but the set of feasible solutions and thus potential Pareto optimal solutions is reduced to those which have bounded risk costs. A preliminary report on this research, the robust optimisation approach for tactical planning of a retail supply chain product flow was announced in an extended abstract by Vizinger, Kokolj and Žerovnik (2017). Here we outline the entire solution procedure and test the model on a real-life instance.

3.2 Robust optimisation

A robust optimisation approach is performed in four consecutive steps (see Figure 1), where we first generate an initial solution (having at most some percent of lost sales or overstocking effects realisation). After examination of a limited number of testing scenarios, we iteratively seek solutions with minimal supply risks. For the set of best solutions (with minimal supply risk) we evaluate the distribution costs. This approach is closely related to robust programming, at least as regards the generation of scenarios. Moreover, the objective function to be minimised is no longer in the form of a sum of all three costs (as is the case in the exact solution approaches and previously used heuristics), but the criteria are rather placed in the hierarchy way. This allows us to exclude time prohibitive linear programming from the iterative improvement, which greatly speeds up the heuristic solution procedures.

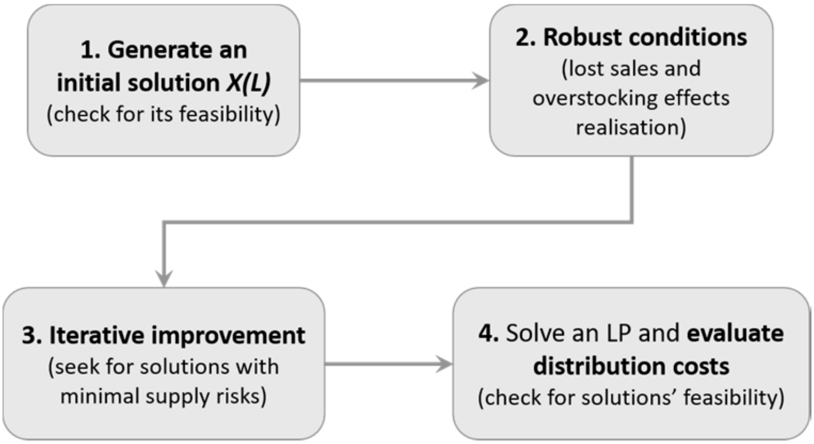


Figure 1: Four steps of a robust optimisation approach

3.2.1 Generating an initial solution

At first, an initial solution (or, it may be known from past experience) is generated at random. Let us assume that we know the future sales (in reality, actual demand is known a posteriori), given in a vector $l_j = [0, \dots, l_j(m)]$, where $l_j(m)$ represents the last possible (maximal) sold quantity at store j. Each l_j is assumed to take a limited number of values, and so may take the maximal possible value as does the maximal past sale (defined with the random variable b_j). Recall that the possible sale realisation is given by a scenario described in vector $L = [l_j]$, where l is the quantity needed at store j.

As indicated above, we first generate a scenario L and use $X(L)$ as the initial solution. Here $X(L)$ is the optimal solution of the linear problem solving the deterministic transportation problem corresponding to scenario L. A solution X represents the distribution plan given in a matrix, presented in expression (1).

3.2.2 Robust conditions

In contrast to the previous approach (see Vizinger and Žerovnik, 2019), the initial solution X is further checked for coincidence with the specified robust criteria (maximum allowable supply risks). In particular, here this means that the initial solution must allow at most 25% of overstocking effects and/or with at most e.g. 30% of lost sales realisation. If the solution generated does not fit, a new solution is generated at random, until a feasible fitting solution is obtained. Note that if we assume that the probability distributions of L are given by (independent) random realisations of the random variables l_j (defined by b_j), the probability of a scenario L is clearly:

$$p(L) = \prod_j p(j, l). \tag{6}$$

We calculate the expected overstocking effects for the solution generated initially as well as for the last possible solution (maximal distributed quantities). If the ratio of $E\big(OS(X)\big)$ for the solution tested to $E\big(OS(X)\big)$ for the last possible solution is less than some pre-defined percentage (for example 25%), we may, with reasonable confidence, accept the solution generated initially. Furthermore, we may search for feasible solutions with at most e.g. 30% of lost sales, where we consider the ratio of $E\big(LS(X)\big)$ for the tested solution to $E\big(LS(X)\big)$ for the first possible solution (having minimal distributed quantities). Note that solutions with higher probability for sales realisation are more likely to be generated.

3.2.3 Tabu search

The iterative improvement phase proceeds along the lines of our previous experiments, i.e. the tabu search heuristic is applied because this procedure had the best performance when several local search based heuristics were tested (see Vizinger and Žerovnik, 2019, 2018). Tabu search generates a random neighbour (random selection of a store and a random change of the amount to be delivered), and moves to the new solution based on the difference in the goal function. The goal function is improved if it is minimised compared to the objective function value of the previous solution. The objective function to be minimised is represented here as a sum of supply risks $E\big(OS(X)\big) + E\big(LS(X)\big)$. The best known solutions are reported as the final results.

3.2.4 Evaluation of the solutions

For the solutions with minimal supply risks we then solve a linear program and evaluate the distribution costs $DC(X)$. The preferred solution may be the one that appears most often in the set of solutions (with minimal supply risks):

$$\min_{X \in \chi} \Big(E\big(OS(X)\big) + E\big(LS(X)\big)\Big), \tag{7}$$

or the one with minimal total costs (trade-off between the direct and indirect costs):

$$\min_{X \in \chi} \Big(E\big(DC(X)\big), E\big(OS(X)\big) + E\big(LS(X)\big)\Big). \tag{8}$$

The choice of the best solution depends on the decision-maker's requests and preferences. If we are selling a product of higher value, the retailer would naturally like to minimise the supply risks costs and he might choose the solution obtained with equation 7. On the other hand, if we are selling a product of lower value, we would like for the solution to have minimal total costs (so we may sum up the costs in equation 8). If a highly demanded product with a rather low value is under investigation, we want the solution with minimal different types of costs (thus we observe separate direct distribution costs and indirect supply risks as presented in equation 8). We wish to stress that we have more options and the right one should be chosen on the basis of the decision maker's preferences.

Finally, in the robust optimisation approach we check the best solutions whether they satisfy the inventory holding capacities of stores C_j, and whether they do not exceed the supplies available at each origin or warehouse a_i. If none of the solution is feasible, we check the next set of the best solutions from the tabu search solution procedure.

4 Numerical example

The numerical example deals with the distribution of a non-substitutable perishable product from the fruit and vegetable program, i.e., bananas. Retailers usually sells products through stores of multiple formats; in our analysis we focus on the largest store format: megamarket. We assume that megamarkets have the most complete and well maintained databases regarding stocks, orders, etc.

The idea of this analysis is to set up a tactical plan for the selected sub-season of the chosen summer season. Actual sales data were statistically analysed and we found out that there are eight selling seasons (for the banana sales) and each of these we may further divide into at least three sub-seasons. In the selected summer season (July-August) we distinguish four sub-seasons (Monday--Wednesday, Thursday, Friday-Saturday, Sunday). In our example we set up a tactical plan for Fridays and Saturdays of the selected summer season.

The retailer distributes bananas between two warehouses and several hundreds of stores (we focus on 18 megamarkets and disregard other store formats). The chosen product (bananas) is packed into basic units (packages), each weighting 18 kg. We assume that transportation between warehouses and stores is provided once per day, and that the stores may order only a whole number of packages, as a package is a basic transportation unit. For a unit of product (package) we use cost estimates (in €) for daily distribution (transport, warehouse) and supply risks (overstocking effects, lost sales). Costs are estimated on the basis of the interviews with practitioners from the company: $c = €0.02/\text{package}/\text{km}$, $h_w = €0.25/\text{package}/\text{day}$, $h_s = €0.5/\text{package}/\text{day}$, $c_{OS} = €5/\text{package}/\text{day}$, and $c_{LS} = €6/\text{package}/\text{day}$. We also estimate entries d_{ij} of the distance matrix (in km) which are used to calculate the transportation costs $\left(c_{ij} = c \cdot d_{ij}\right)$. Note that the distribution cost (storing and transportation) from warehouse i to store j per unit is computed to be: $dc_{ij} = c_{ij} + h_w + h_s$.

Sales are recorded in kilograms of product sold. Because only whole-numbers of packages can be distributed, kilograms into packages have to be converted first. Since one package of bananas weights approximately 18 kg, we cannot fill the distribution classes with integers only, but need to divide them, e.g., into quarter, half, three-quarter and an entire package. For the case of megamarket 13 the sales distribution for Fridays and Saturdays of the summer season is shown in Figure 2. As we can see, megamarket 13 will sell up to ten packages of bananas in the chosen sub-season, and most probably it will sell between six and eight packages per day. Similar results hold true for other stores. When defining demand distributions we found out that all the stores considered have 20 to 75 sales possibilities (demand classes), and there are $8.6 \cdot 10^{27}$ possible scenarios or sales realisations in total.

For the stochastic model we have first tested the basic local search solution procedures (iterative improvement, tabu search, threshold accepting and a combination of all three) and showed that they are very efficient when addressing the inventory-allocation optimisation problem (Vizinger and Žerovnik, 2019, 2018). The convergence curve of the tabu search heuristic is shown in Figure 3 (note that here the objective function is represented by the sum of all three criteria). Since the tabu search turns out to be the most reliable among all the heuristics tested, we integrated this solution procedure into the robust optimisation approach. Instead of optimisation of the sum of criteria, we optimise the goal function that is defined as follows. First we optimise the cost of risk, which is the sum of two criteria: the expected lost sale cost and the expected overstock cost. Only feasible solutions with low cost of risk are then considered and their transportation costs are computed. Of course, we might optimise the criteria in some other way (hierarchy); for some systems it is perhaps important to minimise only the overstocks in the first stage of the optimisation procedure. Therefore, we may argue that the final decision about the importance of the criteria should be made by the decision makers and that their interactive involvement is definitely desirable.

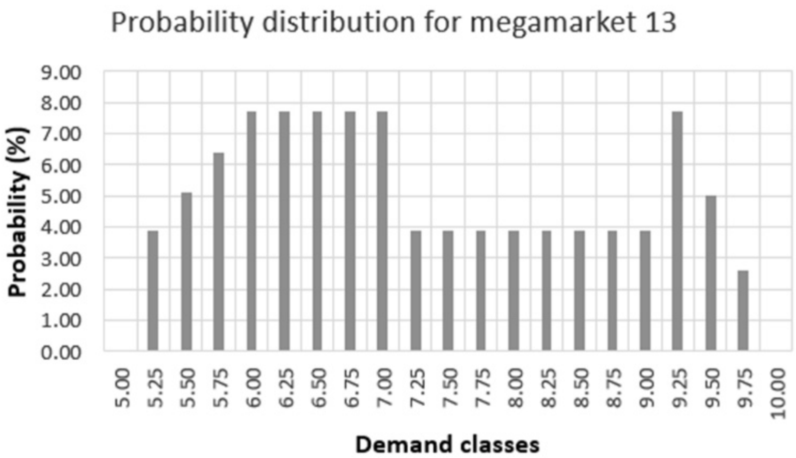


Figure 2: Example of a probability distribution for Fridays and Saturdays of the summer season

In the robust optimisation approach we first generate a scenario L, and use $X(L)$ as the initial solution, where $X(L)$ is the optimal solution of the linear program solving the deterministic transportation problem corresponding to scenario L. For the initial solution we first check if it coincides with the capacities and limitations of the warehouses and stores. For the initial solution, the first possible (minimal distributed quantities) and the last possible scenarios (maximal distributed quantities) we calculate the supply risks and check the robust conditions. In our case we request that the solution have at most 30% of the expected overstocking effects, as well as at most 30% of expected lost sales.

The robust optimisation approach was run 10 times for 1000 iterations. Figure 4 represents the iterative solution procedure (note that here the objective function is represented by the sum of expected overstocks and expected lost sales). As we can see, the expected supply risks amount up to €209 (€105 for overstocks and €104 for lost sales). The best solution from the tabu search procedure (without considering a robust 4-step approach) corresponds to the scenario with supply risk costs that amount up to €223.3 (€44.7 for overstocks and €178.6 for lost sales). Regarding the minimisation of supply risks, the criteria hierarchy definitely outperforms the simple tabu search procedure that uses only a simple or weighted sum of all criteria.

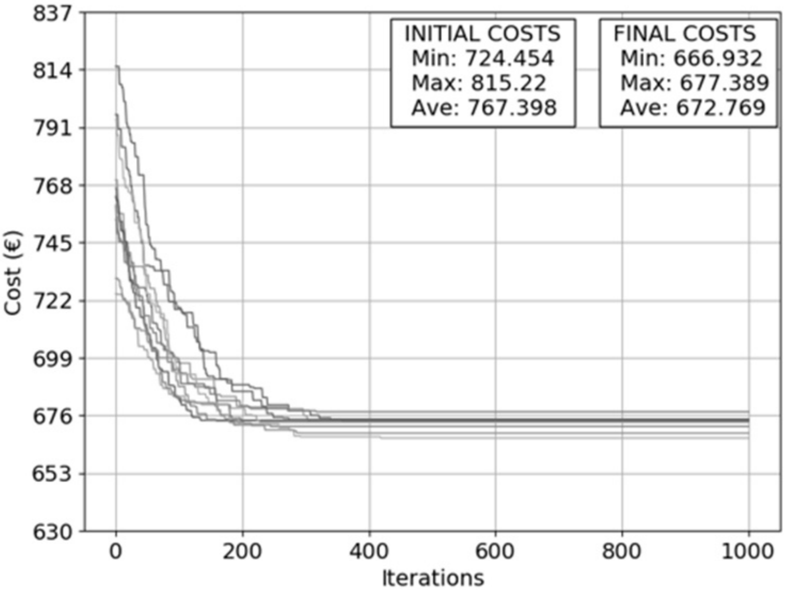


Figure 3: Convergence curve of the tabu search solution procedures (for a realistic case). The cost is the sum of expected risk costs and the distribution cost

For the best ten solutions (from the iterative improvement) we solve the linear program and evaluate the distribution costs. It turns out that the best solution with at most 30% of overstocking effects and at most 30% of lost sale realisations corresponds to the solution with total costs of €708 (distribution costs and supply risks). Nevertheless, although this solution has higher total costs by approximately €41 in comparison to the solution from the tabu search with the sum of the criteria (see also Figure 3), it has lower supply risk costs. Supply risks are also much more balanced (1:1 as compared to the previous result 1:4). The solution obtained by the robust optimisation approach is represented by matrix X, with 2 rows (2 warehouses) and 18 columns (18 megamarkets):

$$X = \begin{bmatrix} 12.5 & 10.5 & 17.75 & 9.75 & 22.75 & 14.0 & 0 & 8.25 & 0 & 8.5 & 0 & 0 & 0 & 0 & 10.0 & 11.0 & 20.25 & 9.75 \\ 0 & 0 & 0 & 0 & 0 & 0 & 12.0 & 0 & 6.75 & 0 & 25.25 & 11.0 & 7.5 & 7.75 & 0 & 0 & 0 & 0 \end{bmatrix}$$

From the distribution plan we see that the first warehouse should supply most of the megamarkets and the second warehouse, megamarkets 7, 9, 11, 12, 13 and 14. In the case of megamarket 13 we note that the appropriate stock level for Fridays and Saturdays of the summer season is seven and a half packages, while with the previous approach we have obtained six and a half packages. Note that we can distribute only a whole number of packages, so in this case we would distribute the difference between the needed level (the result from tactical planning) and the actual stock level, rounded up to the closest integer. We see that with the robust optimisation approach the distributed quantities are higher, as are also distribution costs and overstocking effects. As opposed to this, the costs of lost sales are lower and, most importantly, the costs of supply risk are balanced. The solution obtained can be also called a balanced or compromised solution.

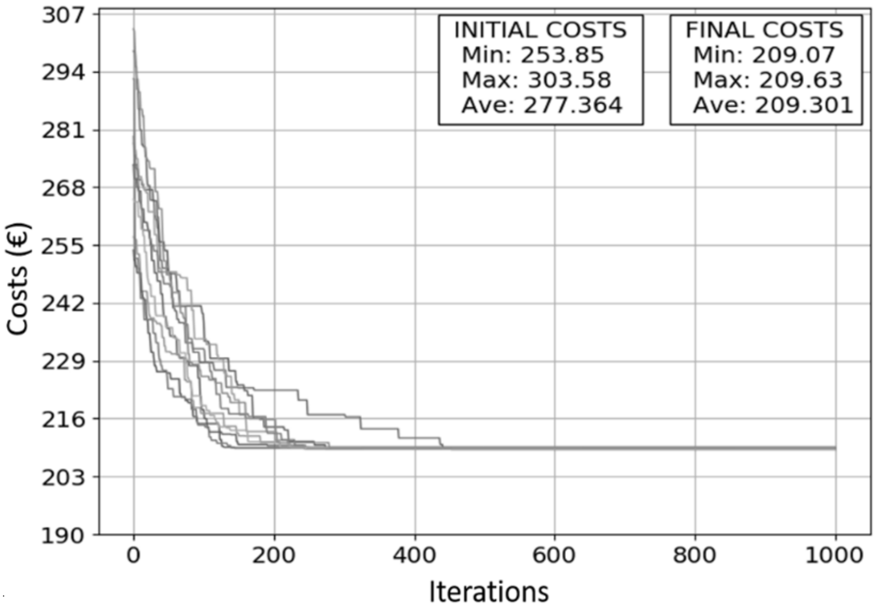


Figure 4: Convergence curve of a robust optimisation approach (for a realistic case). The cost is the expected risk costs

Algorithms have been implemented in a Python environment (the code was not optimised). Search runs were done on an Intel Xeon E3-1230 v3 (8M Cache, 3.3 GHz) processor. For the experimental case of 100 stores, 10 demand classes, 8 initially selected solutions and with 10 000 tabu search iterations performed for each solution, the basic tabu search computation (without considering a robust 4-step approach) took 40 minutes, and the robust computation took roughly 15 seconds. With the incorporation of robust criteria and criteria hierarchy, we have significantly increased the computation speed. Here we need to note that computational time increases when the number of stores changes, while the number of demand classes does not affect significantly the complexity of the problem.

5 Conclusion

This paper presents the robust optimisation approach for the inventory-allocation problem, which is appropriate for tactical planning of a retail supply chain product flow. We have considered a product whose sales figures are independent from those of other products. First, we randomly generated an initial solution (representing a distribution plan with defined inventory levels and allocation of resources), having at most some pre-defined percent of supply risk realisation (robust conditions). Then we used a tabu search algorithm to search for solutions with minimal supply risks (overstocking effects and lost sales). In a previous paper we have shown that local search, the most basic metaheuristics, is a very competitive choice. In fact, the tabu search was shown to be very efficient, therefore we have incorporated this solution procedure into the robust optimisation approach.

The initial solution was further evaluated by taking into the account also the distribution costs assessments. It was shown that exclusion of time prohibitive linear programming from the iterative improvement solution procedure greatly speeds up the computations. Therefore, the implementation that improves separately the distribution cost and the cost of risks is definitely reliable and also allows interactive decision making (e.g. defining robust criteria or choosing an appropriate cost function for optimisation).

From the general introductory discussion, we can conclude that the example discussed here is another argument supporting the claim that simple (meta)heuristics are usually a competitive choice when solving hard optimisation problems.

There are many interesting directions for future research. First of all, we will try to upgrade the model to deal with the distribution plans of the substitutable products. Of course, the optimisation problem will be harder, therefore we might also consider possible improvements of the heuristic solution procedure. For future research we also left out the natural extension that would include a dynamical self-adapting mechanism. Here the comparative model for the operative planning is one of the interesting research avenues, where the distribution quantities are going to be defined with a difference between an actual inventory and the pre-defined maximum level of a certain inventory (resulting from tactical planning).

Acknowledgments

The authors wish to thank Tomaž Kokolj for coding the algorithms and for assistance in running the computational experiments. The authors appreciate the constructive remarks by two anonymous reviewers that helped us to considerably improve the presentation of the paper.

References

Aarts E.H.L., Lenstra J.K. (1997), *Local Search Algorithms*, John Wiley & Sons, Chichester.

Anholcer M. (2016), *Bi-criteria Stochastic Generalized Transportation Problem: Expected Cost and Risk Minimization*, Multiple Criteria Decision Making [online publication], 1 November 2016, doi:10.22367/mcdm.2016.11.01.

Brest J., Žerovnik J. (1999), *An Approximation Algorithm for the Asymmetric Traveling Salesman Problem*, Ricerca Operativa, 28, 59-67.

Donnelly P., Welsh D. (1983), *Finite Particle Systems and Infection Models*, Mathematical Proceedings of the Cambridge Philosophical Society, 94(1), 167-182.

Ferreira A., Žerovnik J. (1993), *Bounding the Probability of Success of Stochastic Methods for Global Optimization*, Computers and Mathematics with Applications, 25, 1-8.

Franca R.B., Jones E.C., Richards C.N., Carlson J.P. (2010), *Multi-objective Stochastic Supply Chain Modelling to Evaluate Tradeoffs Between Profit and Quality*, International Journal of Production Economics, 127, 292-299.

Grossman I.E., Guillén-Gosálbez G. (2010), *Scope for the Application of Mathematical Programming Techniques in the Synthesis and Planning of a Sustainable Processes*, Computers and Chemical Engineering, 34(9), 1365-1376.

Ikica B., Povh J., Žerovnik J. (2019), *Clustering Via a Modified Petford–Welsh Algorithm* [in preparation].

Kirkpatrick S., Gelatt C.D., Vecchi M.P. (1983), *Optimization by Simulated Annealing*, Science, 220 (4598), 671-680.

Liu S., Papageorgiou L.G. (2013), *Multiobjective Optimisation of Production, Distribution and Capacity Planning of Global Supply Chains in the Process Industry*, Omega, 41(2), 369-382.

Pesek I., Schaerf A., Žerovnik J. (2007), *Hybrid Local Search Techniques for the Resource--constrained Project Scheduling Problem*, Lecture Notes in Computer Science, 4771, 57-68.

Petford A.D., Welsh D.J.A. (1989), *A Randomised 3-colouring Algorithm*, Discrete Mathematics, 74(1-2), 253-261.

Sarimveis H., Patrinos P., Tarantilis C.D., Kiranoudis C.T. (2008), *Dynamic Modelling and Control of Supply Chain Systems: A Review*, Computers and Operations Research, 35, 3530-3561.

Sorensen K., Sevaux M., Glover F. (2016), *A History of Metaheuristics*, [in:] R. Marti, P.M. Pardalos, M.G. Resende (eds.), *Handbook of Heuristics*, Springer, International Publishing.

Talbi E.G. (2009), *Metaheuristics: From Design to Implementation*, John Wiley & Sons, New Jersey.

Ubeda S., Žerovnik J. (1997), *A Randomized Algorithm for a Channel Assignment Problem*, Speedup, 11, 14-19.

Vizinger T., Kokolj T., Žerovnik, J. (2017), *A Robust Optimization Approach for Better Planning of a Retail Supply Chain Product Flow*, [in:] L. Zadnik Stirn, M. Kljajič Borštar, J. Žerovnik, S. Drobne (eds.), Proceedings of the 14th International Symposium on Operational Research, 220-226 (Bled, Slovenia, 27-29 September 2017).

Vizinger T., Žerovnik J. (2018), *Coordination of a Retail Supply Chain Distribution Flow*, Tehnični Vjesnik, 25, 5, 1298-1305.

Vizinger T., Žerovnik J. (2019), *A Stochastic Model for Better Planning of Product Flow in Retail Supply Chains*, Journal of the Operational Research Society, 70(11), 1900-1914.

(www 1) Combinatorial Optimization. http://en.wikipedia.org/wiki/Combinatorial_optimization (accessed: 25.04.2019).

(www 2) P vs NP Problem. http://www.claymath.org/millenium-problems/p-vs-np-problem (accessed: 25.04.2019).

(www 3) Frequency Assignment Problem. http://fap.zib.de/index.php (accessed: 25.04.2019)

Zupan H., Herakovič N., Žerovnik J. (2016), *A Heuristic for the Job Shop Scheduling Problem,* [in:] G. Papa, M. Mernik (eds.), Proceedings of the Seventh International Conference on Bioinspired Optimization Methods and their Applications − BIOMA, 187-198.

Žerovnik J. (1994), *A Randomized Algorithm for K-colourability*. Discrete Mathematics, 131(1-3), 379-393.

Žerovnik J. (1995), *A Heuristics for the Probabilistic Traveling Salesman Problem*, [in:] V. Rupnik, M. Bogataj (Eds.), Symposium on Operation Research'95 (Ljubljana: Slovenian Society Informatika, 165-172.

Žerovnik J. (2000), *On Temperature Schedules for Generalized Boltzmann Machine*, Neural Network World, 10, 495-503.

Žerovnik J. (2003), *Simulated Annealing Type Metaheuristics − to Cool or Not to Cool*, [in:] L. Zadnik Stirn, M. Bastič, S. Drobne (eds.), 7th International Symposium on Operational Research in Slovenia (Ljubljana: Slovenian Society Informatika).

Žerovnik J. (2015), *Heuristics for NP-hard Optimization Problems − Simpler is Better!?*, Logistics & Sustainable Transport, 6(1), 1-10.

MULTIPLE CRITERIA DECISION MAKING
Vol. 14 2019

Kazimierz Zaras[*]
Jean-Charles Marin[**]
Bryan Trudel[***]

IDENTIFYING STRATEGIC DEVELOPMENT OBJECTIVES FOR EUROPEAN UNION STATES USING THE DOMINANCE-BASED ROUGH SET APPROACH: THE CASE OF POLAND

DOI: 10.22367/mcdm.2019.14.09

Abstract

The use of the dominance-based rough set approach (DRSA) to help identify and prioritize strategic political, economic, sociological and technological (PEST) objectives for European Union (EU) countries is presented. The countries are first grouped into three categories: [A] those that are doing well according to the selected indicators; [B] those that need support to acquire category A status; [C] those ranked the lowest and needing special support with regard to the criteria considered. The categories correspond to tertiles within the average ranking of all EU countries. DRSA then provides decision rules based on PEST needs in order to improve the development and classification of the country. We conclude that by using this methodology, the EU could identify the strategic objectives to be given priority in order to stimulate its economic development or to improve the economic and sociological status of any country in the union. The case of Poland, a category C country from an economic perspective, is of particular interest.

Keywords: international development, European Union, international aid, economic growth, strategic objectives, rough-set theory, dominance-based rough set approach, selection of portfolio projects, multicriteria analysis, sustainable development.

[*] Université du Québec en Abitibi-Témiscamingue, Rouyn-Noranda, Canada, e-mail: kazimierz.zaras@uqat.ca
[**] Université du Québec en Abitibi-Témiscamingue, Rouyn-Noranda, Canada, e-mail: marj103@uqat.ca
[***] Université du Québec en Abitibi-Témiscamingue, Rouyn-Noranda, Canada, e-mail: bryan.trudel@uqat.ca

1 Introduction

This study proposes a systematic approach to helping Poland and the European Union identify strategic objectives to improve their status as compared to similar economies, using a combination of statistics and dominance-based rough set approach (DRSA). The approach began with a selection of statistical data drawn from various references. The selected variables included in our database were grouped into four different perspectives, namely: political, economic, sociological and technological. The countries were then ranked from each perspective, to obtain a weighted average. The final step was the use of DRSA to identify decision rules and conditions applicable specifically to Poland. These conditions represent strategic objectives that could be pursued in order to improve the development of this country relative to others in the EU.

1.1 Review of the literature

Proposed initially by Pawlak (1982, 1991) and then by Pawlak and Slowinski (1994), the rough set theory is a mathematical tool devised to support decision--making processes. Since its introduction, it has been used in many fields such as medicine, banking, engineering, learning, location selection, pharmacology, finance, market analysis and economics (Pawlak, 2002; Greco et al., 1999, 2001; Zaras, 2004; Zaras et al., 2012; Ho et al., 2016; Renaud et al., 2007; Marin et al., 2014; Prema and Umamaheswari, 2016; Songbian, 2016; Emam et al., 2017). It was later extended by Greco, Matarazo and Slowinski (2001, 1999) and renamed “dominance-based rough set approach”. Zaras then enlarged it to include mixed data, such as deterministic, probabilistic and fuzzy sets (2004). The purpose of the present study is to use DRSA to identify strategic policies that EU decision makers and leaders could implement in order to stimulate the development of the EU or of any of its member nations. For this purpose, 22 variables were selected, which were categorized as political, economic, sociological or technological. DRSA is expected to aid the decision maker to prioritize strategic objectives, based on actual data and results obtained for Poland.

1.2 Interactive approach

The proposed interactive approach (Figure 1) begins with the selection of indicators representing the four perspectives: political, economic, sociological and technological. The next step is to collect data from various databases. The multicriteria classification is then carried out to divide the countries into three categories for geographical analysis and production of the decision table. The

DRSA method is then used to obtain the decision rules by induction, followed by the strategic objectives to be recommended by the central decision-maker (CDM), who implements the actions intended to improve a country's position in the ranking. Once actions are completed at the local decision-maker (LDM) level, an audit should be carried out to verify whether or not the ranking has indeed improved. For this purpose, the CDM returns to data collection and multi-criteria classification.

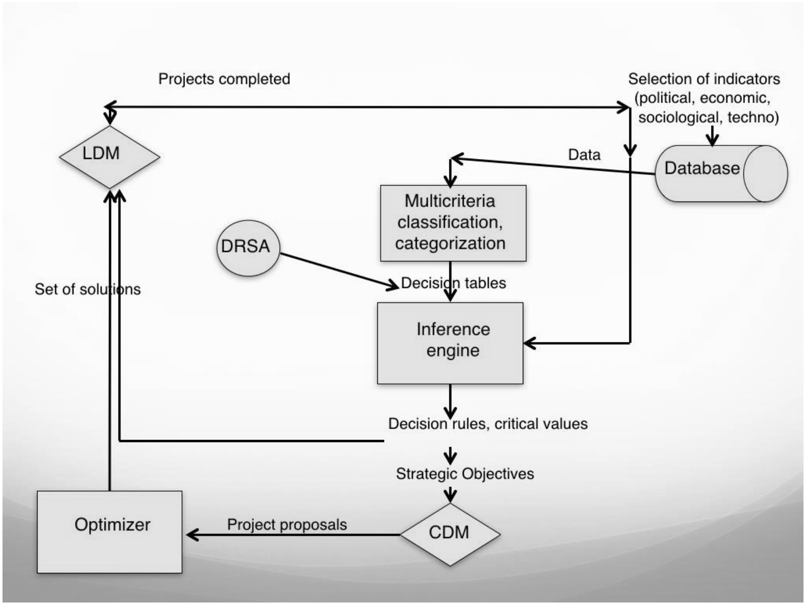


Figure 1: Interactive approach

2 Multicriteria classification

To obtain data for the 22 variables considered in this study, we searched the websites of the World Bank (2018), the United Nations (2018) and also the International Institute for Strategic Studies (2018) during the period from January 2018 to March 2018.

2.1 Political, economic, sociological and technological indicators

Data were categorized in one of the four perspectives, namely: political, economic, sociological and technological (PEST) as summarized in Table 1.

Table 1: Summary of the PEST indicators considered in this study

Perspective or measurement	Definition	Indicator	↑ = higher is better ↓ = lower is better	Poland
Political				
1.1 Global Peace Index	Number of deaths resulting directly from internal conflict involving at least one governmental armed force (2017)	Scale 1-5	↓	1.676
1.2 Military expenditure	Cash outlays of central or federal government to meet the costs of national armed forces (2017)	Scale 1-5		1.922
1.3 Corruption perception index	Based on ranking of countries according to the extent to which corruption is believed to exist (2017)	Scale 0-100	↑	62
1.4 Global competitiveness index	Competitiveness along various axes (2017)	Scale 1-7	↑	4.59
1.5 Ease of doing business index	Ease of completing business transactions (2017)	World rank	↓	27
1.6 Women in government	Proportion of seats held by women in national parliaments (2017)	%	↑	28
Economic				
2.1 Adjusted net national income per capita	Adjusted net national income per capita (Current USD, 2017)	$	↑	10,617.14
2.2 GNP per capita	Gross national product per capita (USD Constant, 2016)	$	↑	15,074.73
2.3 GNI per capita	Gross national income per capita Atlas method (Current USD, 2016)	$	↑	12,690
2.4 Unemployment	Unemployment, total (% of labor force, 2017)	%	↑	5.1
2.5 Exports of G&S	Exports of goods and services (% of GNP, 2017)	%	↑	52.26
Sociological				
3.1 Life expectancy, female	Life expectancy at birth, female (years, 2017)	Years	↑	82.2
3.2 Life expectancy, male	Life expectancy at birth, male (years, 2017).	Years	↑	74.4
3.3 School age	Average age when schooling is terminated (2017)	Years	↑	16
3.4 Urban population	Percentage of the population living in urban areas (2017)	%		60.53
3.5 Adolescent fertility	Number of births per 1,000 women aged 15-19 (2017)	Number	↓	13.03
3.6 Intentional homicides	Death inflicted deliberately on a person by another person (2017)	Scale 1-5	↓	1.35
Technological				
4.1Productivity of academia	Number of scientific articles published per 1 000 000 persons (2017)	Number	↑	157.38
4.2 Internet use	Percentage of active population using the Internet (2017)	%	↑	73.3
4.3 Fixed Internet	Fixed broadband Internet subscriptions per 100 persons (2017)	Number	↑	19.22
4.4 Secure Internet	Secure Internet servers per million persons (2017)	Number	↑	763.73
4.5 Mobile phones	Mobile cellular subscription per 100 persons (2017)	Number	↑	146.21

We thus selected six political, five economic, six sociological and five technological indicators. The political indicators are mostly complex, taking into account the opinions of international groups of experts, panels and think tanks. They are published every year; for example, the Global Peace Index is published by the Institute for Economics and Peace, military expenditure by the International Institute for Strategic Studies, the Corruption Perceptions Index by Transparency International, the Global Competitiveness Index by the World Economic Forum, the Index of ease of doing business by the World Bank, and so on. In Table 1, the values of the indicators are given for 2017, except for GNP and GNI, which are given for 2016. The status of Poland is indicated in the rightmost column.

2.2 Formulation of the multicriteria problem

Our first task was to obtain the overall ranking of the 28 countries based on the 22 indicators or criteria. This was then repeated for each perspective according to the respective criteria. This approach can be described using the AXE model, where:

A is a finite set of countries a_i, $i = 1, 2\ldots 28$;

X is a finite set of criteria X_k, $k = 1, 2\ldots 22$ or of criteria X_{kj} for each perspective j, where $k_j = 1, 2\ldots n_j$ and $\Sigma n_j = 22$

E is the set of evaluations measured by indicators $e_{i,k}$ with respect to criterion X_k or indicators $e_{i,kj}$ with respect to criterion X_{kj} for each perspective j.

The weighted average rank method was used to obtain the ranking of the countries. They were ranked from the most to the least preferable with respect to each indicator in relation to each criterion. Thereafter, since the weights of the indicators were considered equal at the outset, we calculated the weighted average rank for each country. This enabled us to obtain the ranking of the countries with respect to a given perspective as well as for the overall classification.

For each perspective j, the weighted average of country i,

$$r_{i_j} = \sum_{k_j} w_{k_j} r_{k_j i} \tag{1}$$

The overall weighted average of country i,

$$r_i = \sum_{k} w_k r_{ki} \tag{2}$$

where: w_k is the weight of criterion k and w_{k_j} for perspective j; r_{ki} a rank of country i with respect to criterion k and $r_{k_j i}$ for perspective j.

Having obtained the rankings for the 28 countries, overall and for each perspective, the next step was to group them into categories A, B and C, as shown in Table 2.

From this table, we can deduce the following heuristic rules for our sorting task: category A countries earn an A score for at least two perspectives; category B countries receive at most one C score; category C countries receive a C score for at least two perspectives. Table 2 shows that Poland currently earns an overall B score. Decision makers may propose to take actions designed to improve its ranking relative to the rest of the European Union. By extracting decision rules, the DRSA explanatory method allows us to identify the criteria that are most relevant to achieving this as well as the critical values that need to be reached.

Table 2: Overall classification of the 28 UE countries, based on the four perspectives

Overall	European Union State	Political	Economical	Sociological	Technological
A	Netherlands	A	A	A	A
A	Denmark	A	A	A	A
A	Sweden	A	A	A	A
A	Luxembourg	B	A	A	A
A	Austria	A	A	A	A
A	Finland	A	B	A	A
A	Germany	A	A	B	A
A	Belgium	A	A	A	A
A	United Kingdom	B	A	B	A
A	Ireland	A	A	B	B
B	Spain	A	C	A	B
B	Slovenia	B	B	B	B
B	Malta	C	A	B	B
B	France	B	B	B	B
B	Czech Republic	B	B	B	B
B	Portugal	A	C	B	B
B	Italy	C	B	A	B
B	Estonia	B	B	C	A
B	Cyprus	C	B	B	B
B	Poland	B	C	B	B
C	Greece	C	C	A	C
C	Lithuania	B	C	C	B
C	Hungary	B	B	C	C
C	Slovak Republic	C	B	C	C
C	Latvia	C	C	C	B
C	Croatia	C	C	C	C
C	Bulgaria	C	C	C	C
C	Romania	C	C	C	C

2.3 Geographical analysis

Geographical analysis shows that the countries graded as category A are located mostly in northern Europe, the exception being Austria, which is in central Europe. The countries graded B are located in western, central and southern Europe, except for Estonia. The countries graded C are located in eastern and southern Europe.

Figure 2: Geographical analysis of the overall classification of the 28 UE countries

3 Applying the dominance-based rough set approach to determining strategic developmental objectives for Poland

This approach consists of searching for a reduced set of attributes that ensures the same quality of object classification as does the original set of attributes. In rough set theory, the decision problem is represented by a decision table whose rows represent the objects while the columns represent the attributes (Table 3).

Table 3: Decision table

	X_1	…	X_m	D
a_1	$e[(a_1),1]$	…	$e[(a_1),m]$	$e(a_1)$ = {A, B, or C}
a_2	$e[(a_2),1]$	…	$e[(a_2),m]$	$e(a_2)$ = {A, B, or C}
…	…	…	…	…
a_n	$e[(a_n),1]$	…	$e[(a_n),m]$	$e(a_n)$ = {A, B, or C}

In our approach, the objects are the 28 countries; two types of attributes are used: conditional and decisional. The conditional attributes represent the values of the indicators, and we have only one decisional attribute, which is represented by the grade category, A, B or C in the overall classification or with respect to a given perspective.

3.1 The decision rules

To obtain the decision rules, we used 4eMka2 software, which was developed by the Intelligent Decision Support Systems Laboratory (IDSS) at the Computing Science Institute of the Poznan University of Technology (Greco et al., 1999). Rules for the four perspectives combined are presented below in Table 4. Since we were interested in the most significant combination, we kept only rules with a minimal relative strength of 25% and those that were limited to three conditional criteria.

Table 4: Decision rules for all perspectives combined

#	Decision rules	Condition 1
1	Decision ≥ A	Corruption Perception Index ≥ 73
2	Decision ≥ B	Mobile cellular subscriptions ≥ 146.21
3	Decision ≥ B	GNI per capita ≥ $19,880 US

Rule 1 indicates that in order to earn a category A score, the Corruption Perception Index (CPI) must be at least 73. Rules 2 and 3 indicate that to be scored as category B, the number of mobile phones subscriptions per 100 habitants needs to be greater than 146.21 or GNI per capita at least $19,880.

Poland is thus potentially upgradable from B to A based on Rule 1, by improving its CPI to at least 73. We also know from the sorting problem that to move to category A, at least two perspectives must be scored A, and no C score is allowed. Table 5 describes the rules for each of the four PEST perspectives.

In the economic perspective, Poland received a C score. To upgrade to B, gross national income and exports of goods and services should be improved at the same time (Rule 13).

In the political perspective, Poland could improve the Perception of Corruption Index, which should be at least 90, or improve the competitiveness index, since the Ease of Doing Business condition is met. Cutting military spending is incompatible with Polish government's strategy.

Table 5: Decision rules for each PEST perspective

#	Decision Rule	Condition 1	Condition 2	Condition 3
	Political Perspective			
4	Decision ≥ A	Corruption index ≥ 90		
5	Decision ≥ A	Competitiveness index ≥ 5.65	Ease of doing business ≤ 28	
6	Decision ≥ A	Military expenditure ≤ 1.47	Women in government ≥ 30.6%	
7	Decision ≥ B	Military expenditure ≤ 1.47	Ease of doing business ≤ 52	
	Economic Perspective			
10	Decision ≥ A	Unemployment ≤ 3.8%	GNI per capita ≥ $43,850	
11	Decision ≥ B	Export of goods and services ≥ 121.58%		
12	Decision ≥ B	GNI per capita ≥ $56,990		
13	Decision ≥ B	Export of goods and services ≥ 82.87%	GNI per capita ≥ $41,820	
	Sociological Perspective			
14	Decision ≥ A	Years of schooling ≥ 19	Adolescent fertility ≤ 6.38 per 1000	
15	Decision ≥ B	Homicide ≤ 1.25		
16	Decision ≥ B	Years of schooling ≥ 19		
17	Decision ≥ B	Homicide ≤ 1.25	Adolescent fertility ≤ 6.38 per 1000	Urban population ≥ 59.28%
	Technological Perspective			
18	Decision ≥ A	Scientific articles ≥ 378.35		
19	Decision ≥ A	Mobile phones ≥ 148.68 per 100 persons		
20	Decision ≥ B	Mobile Phones ≥ 129.95 per 100 persons		
21	Decision ≥ B	Mobile phones ≥ 111 per 100 persons	Fixed Internet ≥ 38.01 subscriptions per 100	

In the sociological perspective, Poland could upgrade to category A by increasing years of schooling and reducing adolescent fertility. From the technological perspective, Poland could increase spending on research, which would increase the number of scientific papers, or it could increase the number of mobile phones. The strategic objective should be formulated to introduce the quantitative notion of increasing the value of the indicator to satisfy the condition based on the decision rule in relation to the current level.

4 Strategic decision-making

In this section, we show the practical application and usefulness of the decision rules for achieving sustainable political, economic, sociological and technological development in Poland. The decision rules set targets for the improvements specified in the strategic objectives. These targets are based on the statistical data used to extract the decisional rules.

4.1 Strategic objectives and measurements of performance

Table 6 describes various strategic objectives that would be appropriate for Poland. The decision rules set the targets that must be reached for each objective. It is possible that some decision rule conditions are already satisfied, in which case the objective would be to maintain them at their current values. All other values become objectives that would elevate the status of Poland from B to A. It is important to note that Poland is in category C economically and that at least two objectives listed in this perspective would have to be achieved.

Table 6: Strategic objectives and targets for Poland

All perspectives	**Strategic objective 1**	**Strategic objective 2**	**Strategic objective 3**
Decision rule #1	Improve the corruption perception index by 11 points		
Political perspective	**Strategic objective 1**	**Strategic objective 2**	**Strategic objective 3**
Decision rule #2	Improve the corruption perception index by 28 points		
Decision rule #3	Improve the competitiveness index by at least 1.06 points	Maintain the ease of doing business below 28 (currently 27)	
Decision rule #4	Reduce military expenditure by 0.46 points	Improve the proportion of seats held by women in national parliaments by 2.6%.	
Economic perspective	**Strategic objective 1**	**Strategic objective 2**	**Strategic objective 3**
Decision rule #5	Reduce unemployment by 1.3%	Improve the gross national income by $31,160 per capita	
Decision rule #6	Improve exports of goods and services by 69.32% of GNP		

Table 6 cont.

Economic perspective	**Strategic objective 1**	**Strategic objective 2**	**Strategic objective 3**
Decision rule #7	Improve the gross national income by $44,300 per capita		
Decision rule #8	Improve exports of goods and services by 30.61% of GNP	Improve the gross national income by $29,130 per capita	
Sociological perspective	Strategic objective 1	Strategic objective 2	Strategic objective 3
Decision rule #9	Increase schooling by 3 years	Reduce adolescent fertility by 6.65 per 1000	
Technological perspective	Strategic objective 1	Strategic objective 2	Strategic objective 3
Decision rule #10	Increase by 220.97 the number of scientific articles published per 1 000 000 persons		
Decision rule #11	Increase mobile cellular subscriptions by 2.47 per 100 persons		

1) In the overall classification, Poland could move from category B to A by increasing its Corruption Perception Indicator by at least 11 points.
2) For economic classification purposes, rules 6 and 7 are extremely demanding, while rule 8 would be easier to satisfy. Poland could upgrade from C to B by increasing GNI per capita by $29,130 and increasing exportations of goods and services by 30.6% of GNP.
3) From the political perspective, upgrading from B to A status by rule 3 would be easier for Poland provided that the ease of doing business index were maintained at its current level while the competitiveness index increased by at least 1.06 points.
4) From the sociological perspective, it is clear that Poland needs to increase schooling by at least three years and reduce its adolescent fertility index by at least 6.65 per 1000 to move from B to A status.
5) From the technological perspective, Poland could upgrade from B to A status most easily by focusing on rule 11, since there would be relatively few obstacles to increasing the number of mobile phone subscriptions by 2.47 or more per 100 persons.

5 Conclusions

In this study, it was shown that DRSA can be used to obtain a classification of European Union countries for the purpose of designing strategic goals intended to improve their political, economic, sociological and technological status. The decision rules showed the boundary values defining each category and the criterion values which were used to assign Poland to its category.

Overall, the Polish government appears to be effective. It is fighting corruption and the CBA agencies or others doing similar work have met with success in reducing the VAT gap.

The most direct way for Poland to improve its status would be to increase its GDP and per capita GNI. This would be achievable only in the long term. According to the data available, this would earn Poland an A classification, in line with the leading countries of the European Union. Per capita GNI would have to reach at least $43,850 and unemployment would have to be 3.8% or lower. A more realistic economic strategic objective that could be pursued in a shorter term would be to increase its exportation of goods and services.

Strategic objectives in the political realm are very close to being attained thanks to laws and regulations implemented to improve the ease of doing business index and the competitiveness index.

Our analysis indicates that the Polish government could further improve its political status by reducing its military expenditure index. However, in reality, the Polish government cannot do this because of its NATO obligations, which require raising military spending to 2.5% of GDP to ensure the security of the eastern front.

Certain sociological improvements would upgrade the overall classification of Poland from B to A status, particularly in years of schooling and life expectancy. Efforts could be deployed also to reduce the adolescent fertility rate, even though the birth rate needs to be increased overall just to maintain the population.

From the technological perspective, the objective of increasing the number of cell phones is very realistic even though considerable investment would be required to increase network capacity.

References

Emam O., Farhan M., Abohany A. (2017), *Faults Repairing Analysis Using Rough Sets after Implementation of Labor Force Redistribution Algorithm: A Case Study in Telecom Egypt*, Information Sciences Letter, 6(3), 39-48.

Field A. (2005), *Discovering Statistics Using SPSS*, 2nd Edition, SAGE Publications, New Delhi.

Greco S., Matarazzo B., Slowinski R. (1999), *The Use of Rough Sets and Fuzzy Sets in MCDM*, [in:] T. Gal, T. Hanne, T. Stewart (eds.), *Advances in Multiple Criteria Decision Making*, Kluwer Academic Publishers, Dordrecht, Boston, 14.1-14.59.

Greco S., Matarazzo B., Slowinski R. (2001), *Rough Sets Theory for MultiCriteria Decision Analysis*, European Journal of Operational Research, 129, 1-47.

Ho H.-Ch., Fann W.J.-D., Chiang H.-J., Nguyen P.-T., Pham D.-H., Nguyen P.-H., Nagai M. (2016), *Application of Rough Set, GSM and MSM to Analyze Learning Outcome − An Example of Introduction to Education*, Journal of Intelligent Learning Systems and Applications, 8, 23-38.

International Institute for Strategic Studies, IISS, viewed 4 January 2018, https://www.iiss.org

Marin J.-C., Zaras K., Boudreau-Trudel B. (2014), *Use of the Dominance-Based Rough Set Approach as a Decision Aid Tool for the Selection of Development Projects in Northern Quebec*, Modern Economy, 5, 723-741.

Pawlak Z. (1982), *Rough Set*, International Journal of Parallel Programming, 11, 341-356.

Pawlak Z. (1991), *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishing, Dordrecht, http://dx.doi.org/10.1007/978-94-011-3534-4

Pawlak Z. (2002), *Rough Set Theory and Its Applications*, Journal of Telecommunications and Information Theory, 3, 7-10.

Pawlak Z., Slowinski R. (1994), *Rough Set Approach to Multi-Attribute Decision Analysis*, European Journal of Operational Research, 72, 443-459, http://dx.doi.org/10.1016/0377-2217(94)90415-4

Prema S., Umamaheswari P. (2016), *Multitude Classifier Using Rough Set Jelinek-Mercer Naïve Bayes for Disease Diagnosis*, Circuits and Systems, 7, 701-708.

Renaud J., Thibault J., Lanouette R., Kiss L.N., Zaras K., Fonteix C. (2007), *Comparison of Two Multi-Criteria Methods: Net Flow and Rough Set Methods for Aid to Decision Making in a High Yield Pulping Process*, European Journal of Operational Research, 177(3), 1418-1432.

Songbian Z. (2016), *Business Intelligence from Customer Review Management Using Rough Set Model*, International Journal of Advanced Research, 4, 816-824.

United Nations 2018, UNData, viewed 4 January 2018, http://data.un.org/Explorer.aspx?d=UNODC

World Bank 2018, Indicators, viewed 4 January 2018. https://data.worldbank.org/indicator.

Zaras K. (2004), *Rough Approximation of a Preference Relation by a Multi-attribute Stochastic Dominance for Deterministic, Stochastic and Fuzzy Evaluation Problems*, European Journal of Operational Research, 159, 196-206.

Zaras K., Marin J.-C., Boudreau-Trudel B. (2012), *Dominance Rough Set Approach as a Decision--Making Method for the Selection of Sustainable Development Projects*, American Journal of Operational Research, 2, 506.