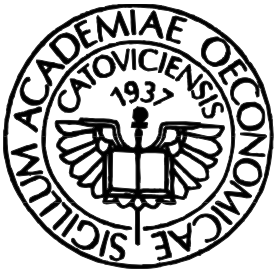
MULTIPLE CRITERIA DECISION MAKING '06

THE KAROL ADAMIECKI UNIVERSITY OF ECONOMICS IN KATOWICE

SCIENTIFIC PUBLICATIONS



MULTIPLE CRITERIA DECISION MAKING '06

Edited by Tadeusz Trzaskalik

Katowice 2007

ISBN 978-83-7246-965-6

**Publisher of The Karol Adamiecki
University of Economics in Katowice**

ul. 1 Maja 50, 40-287 Katowice, tel. +48 032 25 77 635, fax +48 032 25 77 643
www.ae.katowice.pl e-mail:wydawucz@ae.katowice.pl

CONTENTS

PREFACE

The book includes theoretical and applicational papers from the field of the multicriteria decision making. The authors are the faculty members of the Karol Adamiecki University of Economics in Katowice, Department of the Operations Research, and professors from Poland and abroad, collaborating with the Department.

The contents of particular paper is the following.

In the paper ***Dominance-based Rough Set Approach to Multiple Criteria Decision Support*** **R. Słowiński**, **S. Greco** and **B. Matarazzo** present DRSA for multiple criteria classification, choice and ranking as well as for decisions under risk. DRSA is compared with other decision support paradigms at an axiomatic level.

In the paper ***Production Planning and Control: an Approach Based on Rough Sets*** **K. Zaraś**, **H. Kane** and **M. Nowak** consider that problem in a job shop where the work flow is controlled by Kanban cards. Decision rules are induced, which are applied to chose the best solution from a large number o alternatives.

In the paper ***Sensitivity Analysis in Linear Vector Optimization*** **S. Sitarz** considers two situations: sensitivity analysis of efficient solutions and sensitivity analysis of dominating solutions.

In the paper ***Application of DEA Method to the Evaluation of the Efficiency of Polish Open Pension Funds in the Years 2004-06*** **D. Miszczyńska** and **M. Miszczyński** analyze the measurement of technical efficiency of these funds.

In the paper ***Multi-criteria Modelling of Integrated Asset & Liability Management in a Commercial Bank*** **J. Michnik** presents some models in a framework of deterministic and stochastic multiobjective programming as well as in the framework of interactive goal programming.

In the paper ***On Multicriteria Problems with Modification of Attributes*** **A. Skulimowski** proposes a mathematical model for multicriteria decision problems with alternatives which may change their properties in a direct response to external actions.

In the paper ***Enterprises Using Analytical Network Process*** **J. Ramik** investigates an economic diagnostic system in the situation of lack of data and proposes a diagnostic model working both with statistical and expert data.

In the paper ***AHP Application to Raw Materials Stock Management*** **T. Trzaskalik** and **S. Zawadzka** consider the problem of choosing logistics method for control the optimal level of stock for and solve it by means of AHP.

The volume editor wishes to acknowledge to the authorities of the Karol Adamiecki University of Economics for the support in editing the subsequent volume from the series ***Multiple Criteria Decision Making***, in the 70th anniversary of the University.

Tadeusz Trzaskalik

Roman Słowiński

Salvatore Greco

Benedetto Matarazzo

DOMINANCE-BASED ROUGH SET APPROACH TO MULTIPLE CRITERIA DECISION SUPPORT

Abstract

The utility of the rough set approach to multiple criteria decision support is related to the nature of both, the input preferential information available in decision analysis, and the output of the analysis. As to the input, the rough set approach requires a set of decision examples. This is convenient for the acquisition of preferential information from decision makers. Very often in multiple criteria decision support, this information has to be given in terms of preference model parameters, such as importance weights, substitution ratios and various thresholds. Producing such information requires a significant cognitive effort on the part of the decision maker. It is generally acknowledged that people often prefer to make exemplary decisions and cannot always explain them in terms of specific parameters. For this reason, the idea of inferring preference models from exemplary decisions provided by the decision maker is very attractive. Furthermore, the exemplary decisions may be inconsistent because of limited clear discrimination between values of particular criteria and because of hesitation on the part of the decision maker. These inconsistencies can convey important information that should be taken into account in the construction of the decision maker's preference model. The rough set approach is intended to deal with inconsistency and this is a major argument to support its application to multiple criteria decision analysis. The output of the analysis, i.e. the model of preferences in terms of "*if..., then...*" decision rules, is very convenient for decision support because it is intelligible and speaks the same language as the decision maker. The rough set approach adapted to multiple criteria decision support is called Dominance-based Rough Set Approach (DRSA). DRSA is concordant with the concept of granular computing, however, the granules are dominance cones in evaluation space and not bounded sets as it is the case in the basic rough set approach. It is also concordant with the paradigm of computing with words, as it exploits ordinal, and not necessarily cardinal, character of data. We present DRSA for multiple criteria classification, choice and ranking, as well as DRSA for decisions under risk. Finally, we compare DRSA with other decision support paradigms at an axiomatic level.

Keywords

Rough sets, multiple criteria decision support, decision under risk, knowledge discovery, preference model, decision rules.

INTRODUCTION

We present a knowledge discovery methodology for multiple attribute and multiple criteria decision support, which is based upon the concept of rough set proposed by Z. Pawlak [29, 30, 32]. Taking part in the development of rough set theory from the beginning, we adapted and extended its basic paradigm in many ways [31, 45, 46, 47]. For a long time, we also made attempts to employ rough set theory for decision support [33, 38, 39]. The standard rough set approach was not able, however, to deal with preference-ordered domains of attribute (then, called criteria) and preference-ordered decision classes, which are characteristic features of decision problems.

In the late 90's, adapting the standard rough set approach to knowledge discovery from preference-ordered data became a particularly challenging problem within the field of multiple criteria decision support. Why might it be so important? The answer is related to the nature of the input preferential information available in multiple criteria decision analysis and of the output of that analysis. As to the input, the rough set approach requires a set of decision examples. Such representation is convenient for the acquisition of preferential information from decision makers. Very often in multiple criteria decision analysis, this information has to be given in terms of preference model parameters, such as importance weights, substitution ratios and various thresholds. Producing such information requires a significant cognitive effort on the part of the decision maker. It is generally acknowledged that people often prefer to make exemplary decisions and cannot always explain them in terms of specific parameters.

For this reason, the idea of inferring preference models from exemplary decisions provided by the decision maker is very attractive. Furthermore, the exemplary decisions may be inconsistent because of limited clear discrimination between values of particular criteria and because of hesitation on the part of the decision maker. These inconsistencies cannot be considered as a simple error or as noise. They can convey important information that should be taken into account in the construction of the decision maker's preference model. The rough set approach is intended to deal with inconsistency and this is a major argument to support its application to multiple criteria decision analysis. The output of the analysis, i.e. the model of preferences in terms of decision rules, is very convenient for decision support because it is intelligible and speaks the same language as the decision maker.

An extension of the standard rough set approach which enables the analysis of preference-ordered data was proposed in [6, 7, 8, 11, 15]. This extension, called the Dominance-based Rough Set Approach (DRSA) is mainly based on

the substitution of the indiscernibility relation by a dominance relation in the rough approximation of decision classes. An important consequence of this fact is the possibility of inferring (from exemplary decisions) a preference model in terms of decision rules which are logical statements of the type "*if..., then...*". The separation of certain and uncertain knowledge about the decision maker's preferences results from the distinction of different kinds of decision rules, induced from lower approximations of decision classes or from the difference between upper and lower approximations (composed of inconsistent examples). Such a preference model is more general than the traditional functional models considered within multiattribute utility theory, or the relational models considered, for example, in outranking methods. This conclusion has been acknowledged by a thorough study of axiomatic foundations [16, 17, 42]. DRSA has also been used as a tool for inducing parameters of other preference models than the decision rules, like the relational outranking model used in multiple criteria choice problems [37].

As to the application side of the rough set approach, it has been used for discovering regularities in complex phenomena, like stormwater pollution [36], bankruptcy risk of firms applying for a bank credit [38], finding indications for a surgery treatment [31] and classification of Siberian forests [2]. A special attention has been paid to application of the rough set approach in clinical practice, to support some diagnostic and managerial decisions in hospital emergency rooms. This application required extension of the rough set approach to handle incomplete data. The results were implemented as a "decision making core" of a clinical decision support system developed on a mobile platform [27]. The system, called MET (*Mobile Emergency Triage*), supports triage of pediatric patients with various acute conditions. It underwent a clinical trial in the Children's Hospital of Eastern Ontario in Ottawa [50].

Since the first formulation of DRSA, we have proposed many extensions of the approach that make it a useful tool for many specific decision problems. In this survey, we characterize the basic DRSA approach and its main extensions (for complementary surveys see [18, 19, 20, 45]).

The chapter is organized as follows. In the next section, we introduce the concept of knowledge discovery from preference-ordered data. Then, we present the basic Dominance-based Rough Set Approach (DRSA) and in the following sections we review its main extensions. In the last section some conclusions are given and current research directions are outlined.

1. KNOWLEDGE DISCOVERY FROM PREFERENCE ORDERED DATA

The data set in which classification patterns are searched for is called the *learning sample*. The learning of patterns from this sample should take into account available *prior knowledge* that may include the following items (see [40]):

(i) Domains of attributes, i.e. sets of values that an attribute may take while being meaningful to the user.

(ii) A division of attributes into condition and decision attributes, which restricts the range of patterns to functional relations between condition and decision attributes.

(iii) A preference order in the domains of some attributes and a semantic correlation between pairs of these attributes, requiring the patterns to observe the dominance principle.

In fact, item (i) is usually taken into account in knowledge discovery. With this prior knowledge only, one can discover patterns called *association rules* which show strong relationships between values of some attributes, without fixing which attributes will be on the condition and which ones on the decision side in all rules.

If item (i) is combined with item (ii) in the prior knowledge, then one can consider a partition of the learning sample into decision classes defined by decision attributes. The patterns to be discovered have then the form of *decision trees* or *decision rules* representing functional relations between condition and decision attributes. These patterns are typically discovered by machine learning and data mining methods [28]. As there is a direct correspondence between a decision tree and rules, we will concentrate our attention on decision rules only.

As item (iii) is crucial for decision support, let us explain it in more detail. Consider an example of a data set concerning pupils' achievements in a high school. Suppose that among the attributes describing the pupils there are results in *Mathematics* (*Math*) and *Physics* (*Ph*). There is also a *General Achievement* (*GA*) result. The domains of these attributes are composed of three values: *bad*, *medium* and *good*. This information constitutes item (i) of prior knowledge. Item (ii) is also available because, clearly, *Math* and *Ph* are condition attributes while *GA* is a decision attribute. The preference order of the attribute values is obvious: *good* is better than *medium* and *bad*, and *medium* is better than *bad*. It is known, moreover, that both *Math* and *Ph* are semantically correlated with *GA*. This is, precisely, item (iii) of the prior knowledge.

Attributes with preference-ordered domains are called *criteria* because they involve an evaluation. We will use the name of *regular attributes* for those

attributes whose domains are not preference-ordered. *Semantic correlation between two criteria* (condition and decision) means that an improvement on one criterion should not worsen the evaluation on the second criterion, while other attributes and criteria are unchanged. In our example, an improvement of a pupil's score in *Math* or *Ph*, with other attribute values unchanged, should not worsen the pupil's general achievement (*GA*), but rather improve it. In general, semantic correlation between condition criteria and decision criteria requires that an object x dominating object y on all condition criteria (i.e. x having evaluations at least as good as y on all condition criteria) should also dominate y on all decision criteria (i.e. x should have evaluations at least as good as y on all decision criteria). This principle is called the *dominance principle* (or Pareto principle) and it is the only objective principle that is widely agreed upon in the multiple criteria comparisons of objects.

Let us consider two questions:

- What classification patterns can be drawn from the pupils' data set?
- How does item (iii) influences the classification patterns?

The answer to the first question is: "*if..., then...*" decision rules. Each decision rule is characterized by a *condition profile* and a *decision profile*, corresponding to vectors of threshold values of regular attributes and criteria in the condition and decision parts of the rule, respectively. The answer to the second question is that condition and decision profiles of a decision rule should observe the dominance principle if the rule has at least one pair of semantically correlated criteria spanned over the condition and decision part. We say that one profile *dominates* another if they both involve the same values of regular attributes and the values of criteria of the first profile are not worse than the values of criteria of the second profile.

Let us explain the dominance principle with respect to decision rules on the pupils' example. Suppose that two rules induced from the pupils' data set relate *Math* and *Ph* on the condition side, with *GA* on the decision side:

rule #1: if *Math* = *medium* and *Ph* = *medium*, then *GA* = *good*,

rule #2: if *Math* = *good* and *Ph* = *medium*, then *GA* = *medium*.

The two rules do not observe the dominance principle because the condition profile of rule #2 dominates the condition profile of rule #1, while the decision profile of rule #2 is dominated by the decision profile of rule #1. Thus, in the sense of the dominance principle, the two rules are inconsistent, i.e. they are wrong.

One could say that the above rules are true because they are supported by examples of pupils from the learning sample, but this would mean that the examples are also inconsistent. The *inconsistency* may come from many sources. Examples include:

- Missing attributes (regular ones or criteria) in the description of objects. Maybe the data set does not include such attributes as the *opinion of the pupil's tutor* expressed only verbally during an assessment of the pupil's *GA* by a school assessment committee.
- Unstable preferences of decision makers. Maybe the members of the school assessment committee changed their view on the influence of *Math* on *GA* during the assessment.

Handling these inconsistencies is of crucial importance for knowledge discovery. They cannot be simply considered as noise or error to be eliminated from data, or amalgamated with consistent data by some averaging operators. They should be identified and presented as uncertain patterns.

If item (iii) were ignored in prior knowledge, then the handling of the above mentioned inconsistencies would be impossible. Indeed, there would be nothing wrong with rules #1 and #2. They would be supported by different examples discerned by considered attributes.

It has been acknowledged by many authors that *rough set theory* provides an excellent framework for dealing with inconsistency in knowledge discovery [26, 30, 32, 34, 35, 39]. The paradigm of rough set theory is that of *granular computing*, because the main concept of the theory (rough approximation of a set) is built up of blocks of objects which are indiscernible by a given set of attributes, called *granules of knowledge*. In the space of regular attributes, the granules are bounded sets. Decision rules induced from rough approximation of a classification are also built up of such granules. While taking into account prior knowledge of type (i) and (ii), the rough approximation and the inherent rule induction ignore, however, prior knowledge of type (iii). In consequence, the resulting decision rules may be inconsistent with the dominance principle.

The authors have proposed an extension of the granular computing paradigm that enables us to take into account prior knowledge of type (iii), in addition to either (i) only [23], or (i) and (ii) together [8, 15, 40]. The combination of the new granules with the idea of rough approximation is called the *Dominance-based Rough Set Approach* (DRSA).

In the following, we present the concept of granules which permit us to handle prior knowledge of type (iii) when inducing decision rules.

Let U be a finite set of objects (universe) and let Q be a finite set of attributes divided into a set C of *condition attributes* and a set D of *decision attributes* where $C \cap D = \varnothing$. Also, let $X_C = \prod_{q=1}^{|C|} X_q$ and $X_D = \prod_{q=1}^{|D|} X_q$ be attribute spaces corresponding to sets of condition and decision attributes, respectively. The ele-

ments of X_C and X_D can be interpreted as possible evaluation of objects on attributes from set $C = \{1,\ldots,|C|\}$ and from set $D = \{1,\ldots,|D|\}$, respectively. Therefore, X_q is the set of possible evaluations of considered objects with respect to attribute q. The value of object x on attribute $q \in Q$ is denoted by x_q. Objects x and y are *indiscernible* by $P \subseteq C$ if $x_q = y_q$ for all $q \in P$ and, analogously, objects x and y are indiscernible by $R \subseteq D$ if $x_q = y_q$ for all $q \in R$. The sets of indiscernible objects are equivalence classes of the corresponding *indiscernibility relation* I_P or I_R. Moreover, $I_P(x)$ and $I_R(x)$ denote equivalence classes including object x. I_D generates a partition of U into a finite number of decision classes $\boldsymbol{Cl} = \{Cl_t, t = 1,...,n\}$. Each $x \in U$ belongs to one and only one class $Cl_t \in \boldsymbol{Cl}$.

The above definitions take into account prior knowledge of type (i) and (ii) only. In this case, the granules of knowledge are bounded sets in X_P and X_R ($P \subseteq C$ and $R \subseteq D$), defined by partitions of U induced by the indiscernibility relations I_P and I_R, respectively. Then, classification patterns to be discovered are functions representing granules $I_R(x)$ by granules $I_P(x)$ in the condition attribute space X_P, for any $P \subseteq C$ and for any $x \in U$.

If prior knowledge includes item (iii) in addition to (i) and (ii), then the indiscernibility relation is unable to produce granules in X_C and X_D that would take into account the preference order. To do so, the indiscernibility relation has to be substituted by a dominance relation in X_P and X_R ($P \subseteq C$ and $R \subseteq D$). Suppose, for simplicity, that all condition attributes in C and all decision attributes in D are criteria, and that C and D are semantically correlated.

Let $\succeq_q$ be a *weak preference relation* on U (often called *outranking*) representing a preference on the set of objects with respect to criterion $q \in \{C \cup D\}$. Now, $x_q \succeq_q y_q$ means "x_q is at least as good as y_q with respect to criterion q". On the one hand, we say that x *dominates* y with respect to $P \subseteq C$ (shortly, x *P-dominates* y) in the condition attribute space X_P (denoted by xD_Py) if $x_q \succeq_q y_q$ for all $q \in P$. Assuming, without loss of generality, that the domains of the criteria are numerical (i.e. $X_q \subseteq \boldsymbol{R}$ for any $q \in C$) and that they are ordered so that the preference increases with the value, we can say that xD_Py is equivalent to $x_q \geq y_q$ for all $q \in P$, $P \subseteq C$. Observe that for each $x \in X_P$, xD_Px, i.e. P-dominance is reflexive. On the other hand, the analogous definition holds in the decision attribute space X_R (denoted by xD_Ry), where $R \subseteq D$.

The dominance relations xD_Py and xD_Ry ($P \subseteq C$ *and* $R \subseteq D$) are directional statements where x is a subject and y is a referent.

If $x \in X_P$ is the referent, then one can define a set of objects $y \in X_P$ dominating x, called the *P-dominating set* (denoted by $D_P^+(x)$) and defined as $D_P^+(x) = \{y \in U: yD_Px\}$.

If $x \in X_P$ is the subject, then one can define a set of objects $y \in X_P$ dominated by x, called the *P-dominated set* (denoted by $D_P^-(x)$) and defined as $D_P^-(x) = \{y \in U: xD_P y\}$.

P-dominating sets $D_P^+(x)$ and *P*-dominated sets $D_P^-(x)$ correspond to *positive* and *negative dominance cones* in X_P, with the origin x.

With respect to the decision attribute space X_R (where $R \subseteq D$), the *R*-dominance relation enables us to define the following sets:

$Cl_R^{\geq x} = \{y \in U: yD_R x\}$, $Cl_R^{\leq x} = \{y \in U: xD_R y\}$.

$Cl_{t_q} = \{x \in X_D: x_q = t_q\}$ is a decision class with respect to $q \in D$. $Cl_R^{\geq x}$ is called the *upward union* of classes, and $Cl_R^{\leq x}$ is the *downward union* of classes. If $x \in Cl_R^{\geq x}$, then x belongs to class Cl_{t_q}, $x_q = t_q$, or better, on each decision attribute $q \in R$. On the other hand, if $x \in Cl_R^{\leq x}$, then x belongs to class Cl_{t_q}, $x_q = t_q$, or worse, on each decision attribute $q \in R$. The downward and upward unions of classes correspond to the *positive* and *negative dominance cones* in X_R, respectively.

In this case, the granules of knowledge are open sets in X_P and X_R defined by dominance cones $D_P^+(x)$, $D_P^-(x)$ ($P \subseteq C$) and $Cl_R^{\geq x}$, $Cl_R^{\leq x}$ ($R \subseteq D$), respectively. Then, classification patterns to be discovered are functions representing granules $Cl_R^{\geq x}$, $Cl_R^{\leq x}$ by granules $D_P^+(x)$, $D_P^-(x)$, respectively, in the condition attribute space X_P, for any $P \subseteq C$ and $R \subseteq D$ and for any $x \in X_P$.

In both cases above, the functions are sets of decision rules.

2. THE DOMINANCE-BASED ROUGH SET APPROACH (DRSA)

2.1. Granular computing with dominance cones

When discovering classification patterns, a set D of decision attributes is, usually, a singleton, $D = \{d\}$. Let us take this assumption for further presentation, although it is not necessary for the Dominance-Based Rough Set Approach. The decision attribute d makes a partition of U into a finite number of classes, $\boldsymbol{Cl} = \{Cl_t, t=1,...,n\}$. Each $x \in U$ belongs to one and only one class, $Cl_t \in \boldsymbol{Cl}$. The upward and downward unions of classes boil down, respectively, to:

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s$$

$$Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s$$

where $t = 1,...,n$. Notice that for $t = 2,...,n$ we have $Cl_t^{\geq} = U - Cl_{t-1}^{\leq}$, i.e. all the objects not belonging to class Cl_t or better, belong to class Cl_{t-1} or worse.

Let us explain how the rough set concept has been generalized to the Dominance-Based Rough Set Approach in order to enable granular computing with dominance cones (for more details, see [8, 11, 14, 18, 19, 20, 45]).

Given a set of criteria, $P \subseteq C$, the inclusion of an object $x \in U$ to the upward union of classes $Cl_t^{\geq}$, $t = 2,\ldots,n$, is *inconsistent with the dominance principle* if one of the following conditions holds:

- x belongs to class Cl_t or better but it is P-dominated by an object y belonging to a class worse than Cl_t, i.e. $x \in Cl_t^{\geq}$ but $D_P^+(x) \cap Cl_{t-1}^{\leq} \neq \emptyset$,
- x belongs to a worse class than Cl_t but it P-dominates an object y belonging to class Cl_t or better, i.e. $x \notin Cl_t^{\geq}$ but $D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset$.

If, given a set of criteria $P \subseteq C$, the inclusion of $x \in U$ to $Cl_t^{\geq}$, where $t = 2,\ldots,n$, is inconsistent with the dominance principle, we say that x belongs to $Cl_t^{\geq}$ *with some ambiguity*. Thus, x belongs to $Cl_t^{\geq}$ *without any ambiguity* with respect to $P \subseteq C$, if $x \in Cl_t^{\geq}$ and there is no inconsistency with the dominance principle. This means that all objects P-dominating x belong to $Cl_t^{\geq}$, i.e. $D_P^+(x) \subseteq Cl_t^{\geq}$. Geometrically, this corresponds to the inclusion of the complete set of objects contained in the positive dominance cone originating in x, in the positive dominance cone $Cl_t^{\geq}$ originating in Cl_t.

Furthermore, *x possibly belongs to* $Cl_t^{\geq}$ with respect to $P \subseteq C$ if one of the following conditions holds:

- According to decision attribute d, x belongs to $Cl_t^{\geq}$
- According to decision attribute d, x does not belong to $Cl_t^{\geq}$, but it is inconsistent in the sense of the dominance principle with an object y belonging to $Cl_t^{\geq}$.

In terms of ambiguity, x possibly belongs to $Cl_t^{\geq}$ with respect to $P \subseteq C$, if x belongs to $Cl_t^{\geq}$ with or without any ambiguity. Due to the reflexivity of the dominance relation D_P, the above conditions can be summarized as follows: *x possibly belongs* to class Cl_t or better, with respect to $P \subseteq C$, if among the objects P-dominated by x there is an object y belonging to class Cl_t or better, i.e. $D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset$. Geometrically, this corresponds to the non-empty intersection of the set of objects contained in the negative dominance cone originating in x, with the positive dominance cone $Cl_t^{\geq}$ originating in Cl_t.

For $P \subseteq C$, the set of all objects belonging to $Cl_t^{\geq}$ without any ambiguity constitutes the *P-lower approximation* of $Cl_t^{\geq}$, denoted by $\underline{P}(Cl_t^{\geq})$, and the set of all objects that possibly belong to $Cl_t^{\geq}$ constitutes the *P-upper approximation* of $Cl_t^{\geq}$, denoted by $\overline{P}(Cl_t^{\geq})$. More formally, we can say:

$$\underline{P}(Cl_t^{\geq}) = \{x \in U: D_P^+(x) \subseteq Cl_t^{\geq}\}$$
$$\overline{P}(Cl_t^{\geq}) = \{x \in U: D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset\}$$

where t=1,...,n. Analogously, one can define the *P-lower approximation* and the *P-upper approximation* of $Cl_t^{\leq}$ as follows:

$$\underline{P}(Cl_t^{\leq}) = \{x \in U: D_P^-(x) \subseteq Cl_t^{\leq}\}$$
$$\overline{P}(Cl_t^{\leq}) = \{x \in U: D_P^+(x) \cap Cl_t^{\leq} \neq \emptyset\}$$

where t=1,...,n. The *P*-lower and *P*-upper approximations so defined satisfy the following *inclusion properties* for each $t \in \{1,\ldots,n\}$ and for all $P \subseteq C$:

$$\underline{P}(Cl_t^{\geq}) \subseteq Cl_t^{\geq} \subseteq \overline{P}(Cl_t^{\geq})$$
$$\underline{P}(Cl_t^{\leq}) \subseteq Cl_t^{\leq} \subseteq \overline{P}(Cl_t^{\leq}).$$

All the objects belonging to $Cl_t^{\geq}$ and $Cl_t^{\leq}$ with some ambiguity constitute the *P-boundary* of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, denoted by $Bn_P(Cl_t^{\geq})$ and $Bn_P(Cl_t^{\leq})$, respectively. They can be represented, in terms of upper and lower approximations, as follows:

$$Bn_P(Cl_t^{\geq}) = \overline{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq})$$
$$Bn_P(Cl_t^{\leq}) = \overline{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq})$$

where t = 1,...,n. The *P*-lower and *P*-upper approximations of the unions of classes $Cl_t^{\geq}$ and $Cl_t^{\leq}$ have an important *complementarity property*. It says that if object x belongs without any ambiguity to class Cl_t or better, then it is impossible that it could belong to class Cl_{t-1} or worse, i.e. $\underline{P}(Cl_t^{\geq}) = U - \overline{P}(Cl_{t-1}^{\leq})$, t = 2,...,n.

Due to the complementarity property, $Bn_P(Cl_t^{\geq}) = Bn_P(Cl_{t-1}^{\leq})$, for t = 2,...,n, which means that if x belongs with ambiguity to class Cl_t or better, then it also belongs with ambiguity to class Cl_{t-1} or worse.

From the knowledge discovery point of view, *P*-lower approximations of unions of classes represent *certain knowledge* given by criteria from $P \subseteq C$, while *P*-upper approximations represent *possible knowledge* and the *P*-boundaries contain *doubtful knowledge* given by the criteria from $P \subseteq C$.

The above definitions of rough approximations are based on a strict application of the dominance principle. However, when defining non-ambiguous objects, it is reasonable to accept a limited proportion of negative examples, partic-

ularly for large data tables. This extended version of the Dominance-Based Rough Set Approach is called the Variable-Consistency Dominance-Based Rough Set Approach model [21].

For any $P \subseteq C$, we say that $x \in U$ belongs to $Cl_t^{\geq}$ *with no ambiguity at consistency level* $l \in (0, 1]$, if $x \in Cl_t^{\geq}$ and at least $l \times 100\%$ of all objects $y \in U$ dominating x with respect to P also belong to $Cl_t^{\geq}$, i.e.

$$\frac{card\left(D_P^+(x) \cap Cl_t^{\geq}\right)}{card\left(D_P^+(x)\right)} \geq l$$

The level l is called the *consistency level* because it controls the degree of consistency between objects qualified as belonging to $Cl_t^{\geq}$ without any ambiguity. In other words, if $l<1$, then at most $(1-l)\times 100\%$ of all objects $y \in U$ dominating x with respect to P do not belong to $Cl_t^{\geq}$ and thus contradict the inclusion of x in $Cl_t^{\geq}$.

Analogously, for any $P \subseteq C$ we say that $x \in U$ belongs to $Cl_t^{\leq}$ *with no ambiguity at consistency level* $l \in (0, 1]$, if $x \in Cl_t^{\leq}$ and at least $l \times 100\%$ of all the objects $y \in U$ dominated by x with respect to P also belong to $Cl_t^{\leq}$, i.e.

$$\frac{card\left(D_P^-(x) \cap Cl_t^{\leq}\right)}{card\left(D_P^-(x)\right)} \geq l$$

Thus, for any $P \subseteq C$, each object $x \in U$ is either ambiguous or non-ambiguous at consistency level l with respect to the upward union $Cl_t^{\geq}$ ($t = 2,...,n$) or with respect to the downward union $Cl_t^{\leq}$ ($t = 1,...,n-1$).

The concept of non-ambiguous objects at some consistency level l leads naturally to the definition of P-lower approximations of the unions of classes $Cl_t^{\geq}$ and $Cl_t^{\leq}$ which can be formally presented as follows:

$$\underline{P}^l\left(Cl_t^{\geq}\right) = \{x \in Cl_t^{\geq} : \frac{card\left(D_P^+(x) \cap Cl_t^{\geq}\right)}{card\left(D_P^+(x)\right)} \geq l\}$$

$$\underline{P}^l\left(Cl_t^{\leq}\right) = \{x \in Cl_t^{\leq} : \frac{card\left(D_P^-(x) \cap Cl_t^{\leq}\right)}{card\left(D_P^-(x)\right)} \geq l\}$$

Given $P \subseteq C$ and consistency level l, we can define the *P-upper approximations* of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, denoted by $\overline{P}^l\left(Cl_t^{\geq}\right)$ and $\overline{P}^l\left(Cl_t^{\leq}\right)$, respectively, by complementation of $\underline{P}^l\left(Cl_{t-1}^{\leq}\right)$ and $\underline{P}^l\left(Cl_{t+1}^{\geq}\right)$ with respect to U as follows:

$$\overline{P}^l\left(Cl_t^{\geq}\right) = U - \underline{P}^l\left(Cl_{t-1}^{\leq}\right)$$

$$\overline{P}^l\left(Cl_t^{\leq}\right) = U - \underline{P}^l\left(Cl_{t+1}^{\geq}\right)$$

$\overline{P}^{l}\left(Cl_t^{\geq}\right)$ can be interpreted as the set of all the objects belonging to $Cl_t^{\geq}$, which are *possibly ambiguous* at consistency level *l*. Analogously, $\overline{P}^{l}\left(Cl_t^{\leq}\right)$ can be interpreted as the set of all the objects belonging to $Cl_t^{\leq}$, which are *possibly ambiguous* at consistency level *l*. The *P-boundaries* (*P*-doubtful regions) of $Cl_t^{\geq}$ and $Cl_t^{\leq}$ are defined as:

$$BnP(Cl_t^{\geq}) = \overline{P}^{l}\left(Cl_t^{\geq}\right) - \underline{P}^{l}\left(Cl_t^{\geq}\right)$$
$$BnP(Cl_t^{\leq}) = \overline{P}^{l}\left(Cl_t^{\leq}\right) - \underline{P}^{l}\left(Cl_t^{\leq}\right)$$

where $t = 1,...,n$. The *variable consistency* model of the Dominance-based Rough Set Approach provides some degree of flexibility in assigning objects to lower and upper approximations of the unions of decision classes. It can easily be demonstrated that for $0<l'<l\leq 1$ and $t = 2,...,n$,

$$\underline{P}^{l}\left(Cl_t^{\geq}\right) \subseteq \underline{P}^{l'}\left(Cl_t^{\geq}\right) \text{ and } \overline{P}^{l'}\left(Cl_t^{\geq}\right) \subseteq \overline{P}^{l}\left(Cl_t^{\geq}\right)$$

For every $P\subseteq C$, the objects being consistent in the sense of the dominance principle with all upward and downward unions of classes are the objects *P-correctly classified*. For every $P\subseteq C$, the *quality of approximation of classification* $\boldsymbol{Cl}$ by the set of criteria *P* is defined as the ratio between the number of *P*-correctly classified objects and the number of all the objects in the data sample set. Since the objects which are *P*-correctly classified are those that do not belong to any *P*-boundary of unions $Cl_t^{\geq}$ and $Cl_t^{\leq}$, $t = 1,...,n$, the quality of approximation of classification $\boldsymbol{Cl}$ by set of criteria *P*, can be written as

$$\gamma_P(\boldsymbol{Cl}) = \frac{\left|\left(U - \left(\bigcup_{t\in\{1,...,n\}} Bn_P\left(Cl_t^{\leq}\right)\right) \cup \left(\bigcup_{t\in\{1,...,n\}} Bn_P\left(Cl_t^{\geq}\right)\right)\right|}{|U|} = \frac{\left|\left(U - \left(\bigcup_{t\in\{1,...,n\}} Bn_P\left(Cl_t^{\geq}\right)\right)\right)\right|}{|U|}$$

$\gamma_P(\boldsymbol{Cl})$ can be seen as a measure of the quality of knowledge that can be extracted from the data table, where *P* is the set of criteria and $\boldsymbol{Cl}$ is the considered classification.

Each minimal subset $P\subseteq C$ such that $\gamma_P(\boldsymbol{Cl}) = \gamma_C(\boldsymbol{Cl})$ is called a *reduct* of $\boldsymbol{Cl}$ and is denoted by $RED_{\boldsymbol{Cl}}$. Note that a decision table can have more than one reduct. The intersection of all reducts is called the *core* and is denoted by $CORE_{\boldsymbol{Cl}}$. Criteria from $CORE_{\boldsymbol{Cl}}$ cannot be removed from the data sample set without deteriorating the knowledge to be discovered. This means that in set *C* there are three categories of criteria:

- *Indispensable* criteria included in the core,
- *Exchangeable* criteria included in some reducts but not in the core,

– *Redundant* criteria being neither indispensable nor exchangeable, thus not included in any reduct.

Note that reducts are minimal subsets of attributes and criteria conveying the relevant knowledge contained in the learning sample. This knowledge is relevant for the explanation of patterns in a given decision table but not necessarily for prediction.

It has been shown in [8, 12] that the quality of classification satisfies properties of set functions which are called *fuzzy measures*. For this reason, we can use the quality of classification for the calculation of indices which measure the relevance of particular attributes and/or criteria, in addition to the strength of interactions between them. The useful indices are: the value index and interaction indices of Shapley and Banzhaf; the interaction indices of Murofushi-Soneda and Roubens; and the Möbius representation. All these indices can help to assess the interdependence of the considered attributes and criteria, and can help to choose the best reduct.

2.2. Induction of decision rules

The dominance-based rough approximations of upward and downward unions of classes can serve to induce a generalized description of the objects contained in the decision table in terms of "*if..., then...*" decision rules. For a given upward or downward union of classes, $Cl_t^{\geq}$ or $Cl_s^{\leq}$, the decision rules induced under a hypothesis that objects belonging to $\underline{P}(Cl_t^{\geq})$ or $\underline{P}(Cl_s^{\leq})$ are positive and all the others are negative, suggests an assignment to "class Cl_t or better", or to "class Cl_s or worse", respectively. On the other hand, the decision rules induced under a hypothesis that objects belonging to the intersection $\overline{P}(Cl_s^{\leq}) \cap \overline{P}(Cl_t^{\geq})$ are positive and all the others are negative, are suggesting an assignment to some classes between Cl_s and Cl_t ($s<t$).

In the case of preference-ordered data it is meaningful to consider the following five types of decision rules:

1) *Certain* $D_{\geq}$*-decision rules*. These provide lower profile descriptions for objects belonging to $Cl_t^{\geq}$ without ambiguity:

$$\textit{if } x_{q1} \succeq_{q1} r_{q1} \textit{ and } x_{q2} \succeq_{q2} r_{q2} \textit{ and } \ldots x_{qp} \succeq_{qp} r_{qp}, \textit{ then } x \in Cl_t^{\geq},$$

where for each $w_q, z_q \in X_q$, "$w_q \succeq_q z_q$" means "w_q is <u>at least</u> as good as z_q"

2) *Possible* $D_{\geq}$*-decision rules*. Such rules provide lower profile descriptions for objects belonging to $Cl_t^{\geq}$ with or without any ambiguity:

$$\textit{if } x_{q1} \succeq_{q1} r_{q1} \textit{ and } x_{q2} \succeq_{q2} r_{q2} \textit{ and } \ldots x_{qp} \succeq_{qp} r_{qp}, \textit{ then } x \text{ possibly belongs to } Cl_t^{\geq}$$

3) *Certain* $D_{\leq}$*-decision rules*. These give upper profile descriptions for objects belonging to $Cl_t^{\leq}$ without ambiguity:

$$\textit{if } x_{q1} \preceq_{q1} r_{q1} \textit{ and } x_{q2} \preceq_{q2} r_{q2} \textit{ and } \ldots x_{qp} \preceq_{qp} r_{qp}, \textit{ then } x \in Cl_t^{\leq},$$

where for each $w_q, z_q \in X_q$, "$w_q \preceq_q z_q$" means "w_q is at most as good as z_q"

4) *Possible* $D_{\leq}$*-decision rules*. These provide upper profile descriptions for objects belonging to $Cl_t^{\leq}$ with or without any ambiguity:

$$\textit{if } x_{q1} \preceq_{q1} r_{q1} \textit{ and } x_{q2} \preceq_{q2} r_{q2} \textit{ and } \ldots x_{qp} \preceq_{qp} r_{qp}, \textit{ then } x \text{ possibly belongs to } Cl_t^{\leq}$$

5) *Approximate* $D_{\geq\leq}$*-decision rules*. These represent simultaneously lower and upper profile descriptions for objects belonging to $Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t$ without the possibility of discerning the actual class:

$$\textit{if } x_{q1} \succeq_{q1} r_{q1} \textit{ and} \ldots x_{qk} \succeq_{qk} r_{qk} \textit{ and } x_{qk+1} \preceq_{qk+1} r_{qk+1} \textit{ and } \ldots x_{qp} \preceq_{qp} r_{qp},$$
$$\textit{then } x \in Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t.$$

In the left hand side of a $D_{\geq\leq}$-decision rule we can have "$x_q \succeq_q r_q$" and "$x_q \preceq_q r'_q$", where $r_q \leq r'_q$, for the same $q \in C$. Moreover, if $r_q = r'_q$, the two conditions boil down to "$x_q \sim_q r_q$", where for each $w_q, z_q \in X_q$, "$w_q \sim_q z_q$" means "w_q is indifferent to z_q".

An object $x \in U$ *supports* decision rule r if its description is matching both the condition part and the decision part of the rule. We also say that decision rule r *covers* object x if it matches at least the condition part of the rule. Each decision rule is characterized by its *strength* defined as the number of objects supporting the rule.

A *minimal* rule is an implication where we understand that there is no other implication with a left hand side which has at least the same weakness (which means that it uses a subset of elementary conditions and/or weaker elementary conditions) and which has a right hand side that has at least the same strength (which means, a $D_{\geq}$- or a $D_{\leq}$-decision rule assigning objects to the same union or sub-union of classes, or a $D_{\geq\leq}$-decision rule assigning objects to the same or larger set of classes).

The rules of type 1) and 3) represent certain knowledge extracted from the data table, while the rules of type 2) and 4) represent possible knowledge. Rules of type 5) represent doubtful knowledge.

The rules of type 1) and 3) are *exact* if they do not cover negative examples; they are *probabilistic*, otherwise. In the latter case, each rule is characterized by a confidence ratio, representing the probability that an object matching left hand side of the rule matches also its right hand side. Probabilistic rules concord to the Variable-Consistency Dominance-based Rough Set Approach model mentioned above.

We will now comment upon the application of decision rules to some objects described by criteria from C. When applying $D_{\geq}$-decision rules to an object x, it is possible that x either matches the left hand side of at least one decision rule or it does not. In the case of at least one such match, it is reasonable to conclude that x belongs to class Cl_t, because it is the lowest class of the upward union $Cl_t^{\geq}$ which results from intersection of all the right hand sides of the rules covering x. More precisely, if x matches the left hand side of rules $\rho_1, \rho_2,\ldots,\rho_m$, having right hand sides $x \in Cl_{t1}^{\geq}$, $x \in Cl_{t2}^{\geq},\ldots, x \in Cl_{tm}^{\geq}$, then x is assigned to class Cl_t, where $t = \max\{t1,t2,\ldots,tm\}$. In the case of no matching, we can conclude that x belongs to Cl_1, i.e. to the worst class, since no rule with a right hand side suggesting a better classification of x is covering this object.

Analogously, when applying $D_{\leq}$-decision rules to the object x, we can conclude that x belongs either to class Cl_z, (because it is the highest class of the downward union $Cl_t^{\leq}$ resulting from the intersection of all the right hand sides of the rules covering x) or to class Cl_n, i.e. to the best class, when x is not covered by any rule. More precisely, if x matches the left hand side of rules ρ_1, $\rho_2,\ldots,\rho_m$, having right hand sides $x \in Cl_{t1}^{\leq}$, $x \in Cl_{t2}^{\leq},\ldots, x \in Cl_{tm}^{\leq}$, then x is assigned to class Cl_t, where $t = \min\{t1,t2,\ldots,tm\}$. In the case of no matching, it is concluded that x belongs to the best class Cl_n because no rule with a right hand side suggesting a worse classification of x is covering this object.

Finally, when applying $D_{\geq\leq}$-decision rules to x, it is possible to conclude that x belongs to the union of all the classes suggested in the right hand side of the rules covering x.

A set of decision rules is *complete* if it is able to cover all objects from the decision table in such a way that consistent objects are re-classified to their original classes and inconsistent objects are classified to clusters of classes which refer to this inconsistency. Each set of decision rules that is complete and non-redundant is called *minimal*. Note that an exclusion of any rule from this set makes it non-complete.

In the case of the Variable-Consistency Dominance-based Rough Set Approach, the decision rules are induced from the P-lower approximations whose composition is controlled by the user-specified consistency level l. Consequently, the value of confidence α for the rule should be constrained from the bottom. It is reasonable to require that the smallest accepted confidence level of the rule should not be lower than the currently used consistency level l. Indeed, in the worst case, some objects from the P-lower approximation may create a rule using all the criteria from P thus giving a confidence $\alpha \geq l$.

Observe that the syntax of decision rules induced from dominance-based rough approximations uses the concept of dominance cones: each condition profile is a dominance cone in X_C, and each decision profile is a dominance cone in X_D. In both cases the cone is positive for $D_{\geq}$-rules and negative for $D_{\leq}$-rules.

Also note that dominance cones which correspond to condition profiles can originate in any point of X_C, without the risk of being too specific. Thus, in contrast to traditional granular computing, the condition attribute space X_C need not be discretized.

Some procedures for rule induction from rough approximations have been proposed in [22, 26, 49].

In [3], a new methodology for the induction of monotonic decision trees from dominance-based rough approximations of preference-ordered decision classes has been proposed.

2.3. An illustrative example

To illustrate the application of the DRSA to multiple criteria classification, we will use a part of some data provided by a Greek industrial bank ETEVA which finances industrial and commercial firms in Greece [6, 48]. A sample composed of 39 firms has been chosen for the study in co-operation with the ETEVA's financial manager. The manager has classified the selected firms into three classes of bankruptcy risk. The sorting decision is represented by decision attribute d making a trichotomic partition of the 39 firms:

D = A means "acceptable",
d = U means "uncertain",
d = NA means "non-acceptable".

The partition is denoted by $\boldsymbol{Cl} = \{Cl_A, Cl_U, Cl_{NA}\}$ and, obviously, class Cl_A is better than Cl_U which is better than Cl_{NA}.

The firms were evaluated using the following twelve criteria (↑ means *preference increasing with value* and ↓ means *preference decreasing with value*):

- A_1 = earnings before interests and taxes/total assets, ↑
- A_2 = net income/net worth, ↑
- A_3 = total liabilities/total assets, ↓
- A_4 = total liabilities/cash flow, ↓
- A_5 = interest expenses/sales, ↓
- A_6 = general and administrative expense/sales, ↓
- A_7 = managers' work experience, ↑ (very low = 1, low = 2, medium = 3, high = 4, very high = 5)

- A_8 = firm's market niche/position, ↑ (bad = 1, rather bad = 2, medium = 3, good = 4, very good = 5)
- A_9 = technical structure-facilities, ↑ (bad = 1, rather bad = 2, medium = 3, good = 4, very good = 5)
- A_{10} = organization-personnel, ↑ (bad = 1, rather bad = 2, medium = 3, good = 4, very good = 5)
- A_{11} = special competitive advantage of firms, ↑ (low = 1, medium = 2, high = 3, very high = 4)
- A_{12} = market flexibility, ↑ (very low = 1, low = 2, medium = 3, high = 4, very high = 5)

The first six criteria are cardinal (financial ratios) and the last six are ordinal. The data table is presented in Table 1.

The main questions to be answered by the knowledge discovery process were the following:

- Is the information contained in Table 1 consistent ?
- What are the reducts of criteria ensuring the same quality of approxi-mation of the multiple criteria classification as the whole set of criteria ?
- What decision rules can be extracted from Table 1 ?
- What are the minimal sets of decision rules ?

We will answer these questions using the DRSA. The first result from this approach is a discovery that the financial data matrix is *consistent* for the complete set of criteria C. Therefore, the C-lower and C-upper approximations of $Cl^{\leq}_{NA}$, $Cl^{\leq}_{U}$ and $Cl^{\geq}_{U}$, $Cl^{\geq}_{A}$ are the same. In other words, the quality of approximation of all upward and downward unions of classes, as well as the quality of classification, is equal to 1.

The second discovery is a set of 18 *reducts* of criteria ensuring the same quality of classification as the whole set of 12 criteria:

$RED^1_{Cl} = \{A_1, A_4, A_5, A_7\}$, $RED^2_{Cl} = \{A_2, A_4, A_5, A_7\}$,

$RED^3_{Cl} = \{A_3, A_4, A_6, A_7\}$, $RED^4_{Cl} = \{A_4, A_5, A_6, A_7\}$,

$RED^5_{Cl} = \{A_4, A_5, A_7, A_8\}$, $RED^6_{Cl} = \{A_2, A_3, A_7, A_9\}$,

$RED^7_{Cl} = \{A_1, A_3, A_4, A_7, A_9\}$, $RED^8_{Cl} = \{A_1, A_5, A_7, A_9\}$,

$RED^9_{Cl} = \{A_2, A_5, A_7, A_9\}$, $RED^{10}_{Cl} = \{A_4, A_5, A_7, A_9\}$,

$RED^{11}_{Cl} = \{A_5, A_6, A_7, A_9\}$, $RED^{12}_{Cl} = \{A_4, A_5, A_7, A_{10}\}$,

$RED^{13}_{Cl} = \{A_1, A_3, A_4, A_7, A_{11}\}$, $RED^{14}_{Cl} = \{A_2, A_3, A_4, A_7, A_{11}\}$,

$RED^{15}_{Cl} = \{A_4, A_5, A_6, A_{12}\}$, $RED^{16}_{Cl} = \{A_1, A_3, A_5, A_6, A_9, A_{12}\}$,

$RED^{17}_{Cl} = \{A_3, A_4, A_6, A_{11}, A_{12}\}$, $RED^{18}_{Cl} = \{A_1, A_2, A_3, A_6, A_9, A_{11}, A_{12}\}$.

All the eighteen subsets of criteria are equally good and sufficient for the perfect approximation of the classification performed by ETEVA's financial manager on the 39 firms. The core of ***Cl*** is empty ($CORE_{Cl} = \varnothing$) which means that no criterion is indispensable for the approximation. Moreover, all the criteria are exchangeable and no criterion is redundant.

The third discovery is the set of *all* decision rules. We obtained 74 rules describing $Cl_{NA}^{\leq}$, 51 rules describing $Cl_{U}^{\leq}$, 75 rules describing $Cl_{U}^{\geq}$ and 79 rules describing $Cl_{A}^{\geq}$.

Table 3

Financial data matrix

Firm	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	d
F1	16.4	14.5	59.82	2.5	7.5	5.2	5	3	5	4	2	4	A
F2	35.8	67.0	64.92	1.7	2.1	4.5	5	4	5	5	4	5	A
F3	20.6	61.75	75.71	3.6	3.6	8.0	5	3	5	5	3	5	A
F4	11.5	17.1	57.1	3.8	4.2	3.7	5	2	5	4	3	4	A
F5	22.4	25.1	49.8	2.1	5.0	7.9	5	3	5	5	3	5	A
F6	23.9	34.5	48.9	1.7	2.5	8.0	5	3	4	4	3	4	A
F7	29.9	44.0	57.8	1.8	1.7	2.5	5	4	4	5	3	5	A
F8	8.7	5.4	27.4	3.3	4.5	4.5	5	2	4	4	1	4	A
F9	25.7	29.7	46.8	1.7	4.6	3.7	4	2	4	3	1	3	A
F10	21.2	24.6	64.8	3.7	3.6	8.0	4	2	4	4	1	4	A
F11	18.32	31.6	69.3	4.4	2.8	3.0	4	3	4	4	3	4	A
F12	20.7	19.3	19.7	0.7	2.2	4.0	4	2	4	4	1	3	A
F13	9.9	3.5	53.1	4.5	8.5	5.3	4	2	4	4	1	4	A
F14	10.4	9.3	80.9	9.4	1.4	4.1	4	2	4	4	3	3	A
F15	17.7	19.8	52.8	3.2	7.9	6.1	4	4	4	4	2	5	A
F16	14.8	15.9	27.94	1.3	5.4	1.8	4	2	4	3	2	3	A
F17	16.0	14.7	53.5	3.9	6.8	3.8	4	4	4	4	2	4	A
F18	11.7	10.01	42.1	3.9	12.2	4.3	5	2	4	2	1	3	A
F19	11.0	4.2	60.8	5.8	6.2	4.8	4	2	4	4	2	4	A
F20	15.5	8.5	56.2	6.5	5.5	1.8	4	2	4	4	2	4	A
F21	13.2	9.1	74.1	11.21	6.4	5.0	2	2	4	4	2	3	U
F22	9.1	4.1	44.8	4.2	3.3	10.4	3	4	4	4	3	4	U
F23	12.9	1.9	65.02	6.9	14.01	7.5	4	3	3	2	1	2	U
F24	5.9	-27.7	77.4	-32.2	16.6	12.7	3	2	4	4	2	3	U
F25	16.9	12.4	60.1	5.2	5.6	5.6	3	2	4	4	2	3	U
F26	16.7	13.1	73.5	7.1	11.9	4.1	2	2	4	4	2	3	U
F27	14.6	9.7	59.5	5.8	6.7	5.6	2	2	4	4	2	4	U
F28	5.1	4.9	28.9	4.3	2.5	46.0	2	2	3	3	1	2	U
F29	24.4	22.3	32.8	1.4	3.3	5.0	2	3	4	4	2	3	U
F30	29.7	8.6	41.8	1.6	5.2	6.4	2	3	4	4	2	3	U
F31	7.3	-64.5	67.5	-2.2	30.1	8.7	3	3	4	4	2	3	NA
F32	23.7	31.9	63.6	3.5	12.1	10.2	3	2	3	4	1	3	NA
F33	18.9	13.5	74.5	10.0	12.0	8.4	3	3	3	4	3	4	NA
F34	13.9	3.3	78.7	25.5	14.7	10.1	2	2	3	4	3	4	NA
F35	-13.3	-31.1	63.0	-10.0	21.2	23.1	2	1	4	3	1	2	NA
F36	6.2	-3.2	46.1	5.1	4.8	10.5	2	1	3	3	2	3	NA
F37	4.8	-3.3	71.9	34.6	8.6	11.6	2	2	4	4	2	3	NA
F38	0.1	-9.6	42.5	-20.0	12.9	12.4	1	1	4	3	1	3	NA
F39	13.6	9.1	76.0	11.4	17.1	10.3	1	1	2	1	1	2	NA

The fourth discovery is the finding of *minimal sets* of decision rules. Several minimal sets were found. One of them is shown below. The number in parenthesis indicates the number of objects which support the corresponding rule, i.e. the rule strength:

1. *if* $f(x,A_3)\geq 67.5$ *and* $f(x,A_4)\geq -2.2$ *and* $f(x,A_6)\geq 8.7$, *then* $x\in Cl^{\leq}_{NA}$, (4)
2. *if* $f(x,A_2)\leq 3.3$ *and* $f(x,A_7)\leq 2$, *then* $x\in Cl^{\leq}_{NA}$, (5)
3. *if* $f(x,A_3)\geq 63.6$ *and* $f(x,A_7)\leq 3$ *and* $f(x,A_9)\leq 3$, *then* $x\in Cl^{\leq}_{NA}$, (4)
4. *if* $f(x,A_2)\leq 12.4$ *and* $f(x,A_6)\geq 5.6$, *then* $x\in Cl^{\leq}_{U}$, (14)
5. *if* $f(x,A_7)\leq 3$, *then* $x\in Cl^{\leq}_{U}$, (18)
6. *if* $f(x,A_2)\geq 3.5$ *and* $f(x,A_5)\leq 8.5$, *then* $x\in Cl^{\geq}_{U}$, (26)
7. *if* $f(x,A_7)\geq 4$, *then* $x\in Cl^{\geq}_{U}$, (21)
8. *if* $f(x,A_1)\geq 8.7$ *and* $f(x,A_9)\geq 4$, *then* $x\in Cl^{\geq}_{U}$, (27)
9. *if* $f(x,A_2)\geq 3.5$ *and* $f(x,A_7)\geq 4$, *then* $x\in Cl^{\geq}_{A}$, (20)

As the minimal set of rules is complete and composed of $D_{\geq}$-decision rules and $D_{\leq}$-decision rules only, application of these rules to the 39 firms will result in their exact re-classification to classes of risk.

Minimal sets of decision rules represent the most concise and non-redundant knowledge representations. The above minimal set of 9 decision rules uses 8 criteria and 18 elementary conditions, i.e. 3.85% of descriptors from the data matrix.

The well-known machine discovery methods cannot deal with multiple criteria classification because they do not consider preference orders in the domains of attributes and among the classes. There are multiple criteria decision analysis methods for such classification. However, they are not discovering classification patterns from data. They simply apply a preference model, like the utility function in scoring methods, to a set of objects to be classified. In this sense, they are not knowledge discovery methods at all.

Comparing the DRSA to the standard rough set approach, we can notice the following differences between the two approaches. The standard rough set approach extracts knowledge about a partition of U into classes which are not preference-ordered. The granules used for knowledge representation are sets of objects which are indiscernible by a set of condition attributes.

In the case of the DRSA and multiple criteria classification, the condition attributes are criteria and the classes are preference-ordered. The extracted knowledge concerns a collection of upward and downward unions of classes and the granules used for knowledge representation are sets of objects defined using the dominance relation. This is the main difference between the standard rough set approach and the DRSA.

There are three notable advantages of the DRSA over the standard rough set approach. The first one is the ability to handle criteria, preference-ordered classes and inconsistencies in the set of decision examples that the standard rough set approach is simply not able to discover. Consequently, the rough approximations separate the certain information from the doubtful, which is taken into account in rule induction. The second advantage is the ability to analyze a data matrix without any preprocessing of data. The third advantage lies in the richer syntax of decision rules that are induced from rough approximations. The elementary conditions of decision rules resulting from DRSA use relations from $\{\leq,=,\geq\}$, while those resulting from the standard rough set approach only use =. The DRSA syntax is more understandable to practitioners. The minimal sets of DRSA decision rules are smaller than the minimal sets which result from the standard rough set approach.

3. THE DOMINANCE-BASED ROUGH SET APPROACH FOR MULTIPLE CRITERIA CHOICE AND RANKING

One of the very first extensions of the DRSA concerned preference-ordered data representing pairwise comparisons (i.e. binary relations) between objects on both, condition and decision attributes [7, 8, 11]. Note that while classification is based on the absolute evaluation of objects, choice and ranking refer to pairwise comparisons of objects. In this case, the patterns (i.e. decision rules) to be discovered from the data characterize a comprehensive binary relation on the set of objects. If this relation is a preference relation and if, from among the condition attributes, there are some criteria which are semantically correlated with the comprehensive preference relation, then the data set (serving as the learning sample) can be considered to be preferential information of a decision maker in a multiple criteria choice or ranking problem. In consequence, the comprehensive preference relation characterized by the decision rules discovered from this data set can be considered as a *preference model* for the decision maker. It may be used to explain the decision policy of the decision maker and to recommend a good choice or preference ranking with respect to new objects.

Let us consider a finite set A of objects evaluated by a finite set of criteria C. The best choice (or the preference ranking) in set A is semantically correlated with the criteria from set C. The preferential information concerning the multiple criteria choice or ranking problem is a data set in the form of a pairwise comparison table, which includes pairs of some *reference objects* from a subset $B \subseteq A \times A$. This is described by preference relations on particular criteria from C and a

comprehensive preference relation. One such example is a weak preference relation called the *outranking relation*. By using the DRSA for the analysis of the pairwise comparison table, we can obtain a rough approximation of the outranking relation by a dominance relation. The decision rules induced from the rough approximation are then applied to the complete set A of the objects associated with the choice or ranking. As a result, one obtains a four-valued outranking relation on this set. In order to obtain a recommendation, it is advisable to use an exploitation procedure based on the net flow score of the objects. We present this methodology in more detail below.

3.1. The pairwise comparison table as preferential information and as a learning sample

A set of reference objects represent a decision problem and a decision maker can express the preferences by pairwise comparisons. In the following, xSy denotes the presence, while $xS^c y$ denotes the absence of the outranking relation for a pair of objects $(x,y)\in A\times A$.

For each pair of reference objects $(x,y)\in B\subseteq A\times A$, the decision maker can select one of the three following possibilities:

1) object x is as good as y, i.e. xSy,
2) object x is worse then y, i.e. $xS^c y$,
3) the two objects are incomparable at the present stage.

An $m\times(n+1)$ pairwise comparison table, denoted by $\boldsymbol{S}_{PCT}$, is then created on the basis of this information. The first n columns correspond to the criteria from set C. The last, i.e. the $(n+1)$-th, column represents the comprehensive binary preference relation S or S^c. The m rows are pairs from B. For each pair in $\boldsymbol{S}_{PCT}$, a difference between criterion values is put in the corresponding column. If the decision maker judges that two objects are incomparable, then the corresponding pair does not appear in $\boldsymbol{S}_{PCT}$.

We will define $\boldsymbol{S}_{PCT}$ more formally. For any criterion $g_i\in C$, let T_i be a finite set of binary relations defined on A on the basis of the evaluations of objects from A with respect to the considered criterion g_i, such that for every $(x,y)\in A\times A$ exactly one binary relation $t\in T_i$ is verified. More precisely, given the domain V_i of $g_i\in C$, if $v_i', v_i'' \in V_i$ are the respective evaluations of $x,y\in A$ by means of g_i and $(x,y)\in t$, with $t\in T_i$, then for each $w,z\in A$ having the same evaluations v_i', v_i'' by means of g_i, $(w,z)\in t$. Furthermore, let T_d be a set of binary relations defined on set A (comprehensive pairwise comparisons) such that at most one binary relation $t\in T_d$ is verified for every $(x,y)\in A\times A$.

The *pairwise comparison table* is defined as data table $\boldsymbol{S}_{PCT} = \langle B, C \cup \{d\}, T_C \cup T_d, f \rangle$, where $B \subseteq A \times A$ is a non-empty *set of exemplary pairwise comparisons of reference objects*, $T_C = \bigcup_{g_i \in C} T_i$, d is a decision corresponding to the comprehensive pairwise comparison (comprehensive preference relation), and $f: B \times (C \cup \{d\}) \rightarrow T_C \cup T_d$ is a total function such that $f[(x,y),q] \in T_i$ for every $(x,y) \in A \times A$ and for each $g_i \in C$, and $f[(x,y),q] \in T_d$ for every $(x,y) \in B$. It follows that for any pair of reference objects $(x,y) \in B$ there is verified one and only one binary relation $t \in T_d$. Thus, T_d induces a partition of B. In fact, the data table $\boldsymbol{S}_{PCT}$ can be seen as decision table, since the set of considered criteria C and the decision d are distinguished.

We assume that the exemplary pairwise comparisons made by the decision maker can be represented in terms of *graded preference relations* (for example "very large preference", "large preference", "strict preference", "strong preference" and "very strong preference"), denoted by P_q^h : For each $q \in C$ and for every $(x,y) \in A \times A$,

$$T_i = \{ P_i^h, h \in H_i \},$$

where H_i is a particular subset of the relative integers and

- $x P_i^h y$, $h>0$, means that object x is preferred to object y by degree h with respect to criterion g_i,
- $x P_i^h y$, $h<0$, means that object x is not preferred to object y by degree h with respect to criterion g_i,
- $x P_i^0 y$ means that object x is similar (asymmetrically indifferent) to object y with respect to criterion g_i.

Within the preference context, the similarity relation P_i^0, even if not symmetric, resembles the indifference relation. Thus, in this case, we call this similarity relation "asymmetric indifference". Of course, for each $g_i \in C$ and for every $(x,y) \in A \times A$,

$$[x P_i^h y, h>0] \Rightarrow [y P_i^k x, k \leq 0], [x P_i^h y, h<0] \Rightarrow [y P_i^k x, k \geq 0].$$

The set of binary relations T_d may be defined in a similar way, but $x P_d^h y$ means that object x is comprehensively preferred to object y by degree h. We are considering a pairwise comparison table where the set T_d is composed of two binary relations defined on A:

- x outranks y (denoted by xSy or $(x,y) \in S$), where $(x,y) \in B$,
- x does not outrank y (denoted by $xS^c y$ or $(x,y) \in S^c$), where $(x,y) \in B$, and $S \cup S^c = B$.

Observe that the binary relation S is reflexive, but not necessarily transitive or complete.

3.2. Rough approximation of the outranking and non-outranking relations specified in the pairwise comparison table

In the following we will distinguish between two types of evaluation scales of criteria: *cardinal* and *ordinal*. Let C^N be the set of criteria expressing preferences on a cardinal scale, and let C^O, be the set of criteria expressing preferences on an ordinal scale, such that $C^N \cup C^O = C$ and $C^N \cap C^O = \emptyset$. Moreover, for each $P \subseteq C$, we denote by P^O the subset of P composed of criteria expressing preferences on an ordinal scale, i.e. $P^O = P \cap C^O$, and by P^N we denote the subset of P composed of criteria expressing preferences on a cardinal scale, i.e. $P^N = P \cap C^N$. Of course, for each $P \subseteq C$, we have $P = P^N \cup P^O$ and $P^N \cap P^O = \emptyset$.

The meaning of the two scales is such that in the case of the cardinal scale we can specify the intensity of preference for a given difference of evaluations, while in the case of the ordinal scale, this is not possible and we can only establish an order of evaluations.

3.2.1. Multigraded dominance

Let $P = P^N$ and $P^O = \emptyset$. Given $P \subseteq C$ ($P \neq \emptyset$), $(x,y),(w,z) \in A \times A$, the pair of objects (x,y) is said to dominate (w,z) with respect to criteria from P (denoted by $(x,y)D_P(w,z)$), if x is preferred to y at least as strongly as w is preferred to z with respect to each $g_i \in P$. More precisely, "at least as strongly as" means "by at least the same degree", i.e. $h_i \geq k_i$, where $h_i, k_i \in H_i$, $x P_i^{hi} y$ and $w P_i^{ki} z$, for each $g_i \in P$.

Let $D_{\{i\}}$ be the dominance relation confined to the single criterion $g_i \in P$. The binary relation $D_{\{i\}}$ is reflexive ($(x,y)D_{\{i\}}(x,y)$, for every $(x,y) \in A \times A$), transitive ($(x,y)D_{\{i\}}(w,z)$ and $(w,z)D_{\{i\}}(u,v)$ imply $(x,y)D_{\{i\}}(u,v)$, for every $(x,y),(w,z),(u,v) \in A \times A$), and complete ($(x,y)D_{\{i\}}(w,z)$ and/or $(w,z)D_{\{i\}}(x,y)$, for all $(x,y),(w,z) \in A \times A$). Therefore, $D_{\{i\}}$ is a complete preorder on $A \times A$. Since the intersection of complete preorders is a partial preorder and $D_P = \bigcap_{g_i \in P} D_{\{i\}}$, $P \subseteq C$, then the dominance relation D_P is a partial preorder on $A \times A$.

Let $R \subseteq P \subseteq C$ and $(x,y),(u,v) \in A \times A$; then the following implication holds:

$$(x,y)D_P(u,v) \Rightarrow (x,y)D_R(u,v)$$

Given $P \subseteq C$ and $(x,y) \in A \times A$, we define the following:

- A set of pairs of objects dominating (x,y), called the *P-dominating set*, denoted by $D_P^+(x,y)$ and defined to be $\{(w,z) \in A \times A: (w,z)D_P(x,y)\}$,
- A set of pairs of objects dominated by (x,y), called the *P-dominated set*, denoted by $D_P^-(x,y)$ and defined as $\{(w,z) \in A \times A: (x,y)D_P(w,z)\}$.

The *P*-dominating sets and the *P*-dominated sets defined on *B* for all pairs of reference objects from *B* are "granules of knowledge" that can be used to express *P*-lower and *P*-upper approximations of the comprehensive outranking relations *S* and S^c, respectively:

$$\underline{P}(S) = \{(x,y)\in B:\ D_P^+(x,y)\subseteq S\}$$

$$\overline{P}(S) = \bigcup_{(x,y)\in S} D_P^+(x,y)$$

$$\underline{P}(S^c) = \{(x,y)\in B:\ D_P^-(x,y)\subseteq S^c\}$$

$$\overline{P}(S^c) = \bigcup_{(x,y)\in S^c} D_P^-(x,y)$$

It has been proved in [8] that

$$\underline{P}(S)\subseteq S\subseteq \overline{P}(S),\ \underline{P}(S^c)\subseteq S^c\subseteq \overline{P}(S^c)$$

Furthermore, the following complementarity properties hold:

$$\underline{P}(S) = B-\overline{P}(S^c),\ \overline{P}(S) = B-\underline{P}(S^c)$$

$$\underline{P}(S^c) = B-\overline{P}(S),\ \overline{P}(S^c) = B-\underline{P}(S)$$

The *P*-boundaries (*P*-doubtful regions) of *S* and S^c are defined as

$$Bn_P(S) = \overline{P}(S) - \underline{P}(S),\ Bn_P(S^c) = \overline{P}(S^c) - \underline{P}(S^c)$$

From the above it follows that $Bn_P(S) = Bn_P(S^c)$

The concepts of the quality of approximation, reducts and core can be extended also to the approximation of the outranking relation by multigraded dominance relations.

In particular, the coefficient

$$\gamma_P = \frac{\left|\underline{P}(S)\cup \underline{P}(S^c)\right|}{|B|}$$

defines the *quality of approximation of S* and S^c by $P\subseteq C$. It expresses the ratio of all pairs of reference objects $(x,y)\in B$ correctly assigned to *S* and S^c by the set *P* of criteria to all the pairs of objects contained in *B*. Each minimal subset $P\subseteq C$, such that $\gamma_P = \gamma_C$, is called a *reduct* of *C* (denoted by $RED_{\boldsymbol{S}_{PCT}}$). Note that $\boldsymbol{S}_{PCT}$ can have more than one reduct. The intersection of all *B*-reducts is called the *core* (denoted by $CORE_{\boldsymbol{S}_{PCT}}$).

It is also possible to use the Variable Consistency Model on $\boldsymbol{S}_{PCT}$ [41] but being aware that some of the pairs in the positive or negative dominance sets belong to the opposite relation but at least $l\times100\%$ of pairs belong to the correct one. Then the definition of the lower approximations of S and S^c boils down to:

$$\underline{P}(S)=\left\{(x,y)\in B:\frac{\left|D_P^+(x,y)\cap S\right|}{\left|D_P^+(x,y)\right|}\geq l\right\}$$

$$\underline{P}(S^c)=\left\{(x,y)\in B:\frac{\left|D_P^-(x,y)\cap S^c\right|}{\left|D_P^-(x,y)\right|}\geq l\right\}$$

3.2.2. Dominance without degrees of preference

The degree of graded preference considered above is defined on a cardinal scale of the strength of preference. However, in many real world problems, the existence of such a quantitative scale is rather questionable. This is the case with ordinal scales of criteria. In this case, the dominance relation is defined directly on evaluations $g_i(x)$ for all objects $x\in A$. Let us explain this latter case in more detail.

Let $P = P^O$ and $P^N = \varnothing$, then, given $(x,y),(w,z)\in A\times A$, the pair (x,y) is said to dominate the pair (w,z) with respect to criteria from P (denoted by $(x,y)D_P(w,z)$), if, for each $g_i\in P$, $g_i(x)\geq g_i(w)$ and $g_i(z)\geq g_i(y)$.

Let $D_{\{i\}}$ be the dominance relation confined to the single criterion $g_i\in P^O$. The binary relation $D_{\{i\}}$ is reflexive, transitive, but non-complete (it is possible that *not* $(x,y)D_{\{i\}}(w,z)$ and *not* $(w,z)D_{\{i\}}(x,y)$ for some $(x,y),(w,z)\in A\times A$). Therefore, $D_{\{i\}}$ is a partial preorder. Since the intersection of partial preorders is also a partial preorder and $D_P=\bigcap_{g_i\in P} D_{\{i\}}$, $P = P^O$, then the dominance relation D_P is a partial preorder.

If some criteria from $P\subseteq C$ express preferences on a quantitative or a numerical non-quantitative scale and others on an ordinal scale, i.e. if $P^N \neq \varnothing$ and $P^O \neq \varnothing$, then, given $(x,y),(w,z)\in A\times A$, the pair (x,y) is said to dominate the pair (w,z) with respect to criteria from P, if (x,y) dominates (w,z) with respect to both P^N and P^O. Since the dominance relation with respect to P^N is a partial preorder on $A\times A$ (because it is a multigraded dominance) and the dominance with respect to P^O is also a partial preorder on $A\times A$ (as explained above), then the dominance D_P, being the intersection of these two dominance relations, is a partial preorder. In consequence, all the concepts introduced in the previous section can be restored using this specific definition of dominance.

3.3. Induction of decision rules from rough approximations of outranking and non-outranking relations

Using the rough approximations of S and S^c defined in 3.2.1 and 3.2.2, it is possible to induce a generalized description of the preferential information contained in a given $\boldsymbol{S}_{PCT}$ in terms of suitable decision rules. The syntax of these rules is based on the concept of *upward cumulated preferences* (denoted by $P_i^{\geq h}$) and *downward cumulated preferences* (denoted by $P_i^{\leq h}$), having the following interpretation:

- $x\, P_i^{\geq h}\, y$ means "x is preferred to y with respect to g_i by at least degree h",
- $x\, P_i^{\leq h}\, y$ means "x is preferred to y with respect to g_i by at most degree h".

Exact definition of the cumulated preferences, for each $(x,y)\in A\times A$, $g_i\in C$ and $h\in H_i$, can be represented as follows:

- $x\, P_i^{\geq h}\, y$ if $x\, P_i^{k}\, y$, where $k\in H_i$ and $k\geq h$,
- $x\, P_i^{\leq h}\, y$ if $x\, P_i^{k}\, y$, where $k\in H_i$ and $k\leq h$.

Let also $G_i = \{g_i(x),\ x\in A\}$, $g_i\in C^O$. The decision rules have then the following syntax:

1. *Certain* $\mathrm{D}_{\geq}$*-decision rules*:

if $x\, P_{i1}^{\geq h(i1)}\, y$ *and...* $x\, P_{ie}^{\geq h(ie)}\, y$ *and* $g_{ie+1}(x)\geq r_{ie+1}$ *and* $g_{ie+1}(y)\leq s_{ie+1}$ *and...* $g_{ip}(x)\geq r_{ip}$ *and* $g_{ip}(y)\leq s_{ip}$, *then* xSy,

where $P = \{g_{i1},...,g_{ip}\}\subseteq C$, $P^N = \{g_{i1},...,g_{ie}\}$, $P^O = \{g_{ie+1},...,g_{ip}\}$, $(h(i1),...,h(ie))\in H_{i1}\times...\times H_{ie}$ and $(r_{ie+1},...,r_{ip}),(s_{ie+1},...,s_{ip})\in G_{ie+1}\times...\times G_{ip}$. These rules are supported by pairs of objects from the P-lower approximation of S only.

2. *Certain* $\mathrm{D}_{\leq}$*-decision rules*:

if $x\, P_{i1}^{\leq h(i1)}\, y$ *and...* $x\, P_{ie}^{\leq h(ie)}\, y$ *and* $g_{ie+1}(x)\leq r_{ie+1}$ *and* $g_{ie+1}(y)\geq s_{ie+1}$ *and...* $g_{ip}(x)\leq r_{ip}$ *and* $g_{ip}(y)\geq s_{ip}$, *then* xS^cy,

where $P = \{g_{i1},...,g_{ip}\}\subseteq C$, $P^N = \{g_{i1},...,g_{ie}\}$, $P^O = \{g_{ie+1},...,g_{ip}\}$, $(h(i1),...,h(ie))\in H_{i1}\times...\times H_{ie}$ and $(r_{ie+1},...,r_{ip}),(s_{ie+1},...,s_{ip})\in G_{ie+1}\times...\times G_{ip}$. These rules are supported by pairs of objects from the P-lower approximation of S^c only.

3. *Approximate* $\mathrm{D}_{\geq\leq}$*-decision rules*:

if $x\, P_{i1}^{\geq h(i1)}\, y$ *and...* $x\, P_{ie}^{\geq h(ie)}\, y$ *and* $x\, P_{ie+1}^{\leq h(ie+1)}\, y$... $x\, P_{if}^{\leq h(if)}\, y$ *and* $g_{if+1}(x)\geq r_{if+1}$ *and* $g_{if+1}(y)\leq s_{if+1}$ *and...* $g_{ig}(x)\geq r_{ig}$ *and* $g_{ig}(y)\leq s_{ig}$ *and* $g_{ig+1}(x)\leq r_{ig+1}$ *and* $g_{ig+1}(y)\geq s_{ig+1}$ *and...* $g_{ip}(x)\leq r_{ip}$ *and* $g_{ip}(y)\geq s_{ip}$, *then* xSy *or* xS^cy,

where $O' = \{g_{i1},...,g_{ie}\}\subseteq C$, $O'' = \{g_{ie+1},...,g_{if}\}\}\subseteq C$, $P^N = O'\cup O''$, O' and O'' are not necessarily disjoint, $P^O = \{g_{if+1},...,g_{ip}\}$, $(h(i1),...,h(if))\in H_{i1}\times...\times H_{if}$, $(r_{if+1},...,r_{ip}),(s_{if+1},...,s_{ip})\in G_{if+1}\times...\times G_{ip}$. These rules are supported by pairs of objects from the P-boundary of S and S^c only.

3.4. Use of decision rules for decision support

The decision rules induced from a given $\boldsymbol{S}_{PCT}$ describe the comprehensive preference relations S and S^c either exactly ($D_{\geq}$- and $D_{\leq}$-decision rules) or approximately ($D_{\geq\leq}$-decision rules). A set of these rules covering all pairs of $\boldsymbol{S}_{PCT}$ represents a preference model from the decision maker who gave the pairwise comparison of reference objects. The application of these decision rules on a new subset $M \subseteq A$ of objects induces a specific preference structure on M.

In fact, any pair of objects $(u,v) \in M \times M$ can match the decision rules in one of four ways:

- at least one $D_{\geq}$-decision rule and neither $D_{\leq}$- nor $D_{\geq\leq}$-decision rules,
- at least one $D_{\leq}$ -decision rule and neither $D_{\geq}$- nor $D_{\geq\leq}$-decision rules,
- at least one $D_{\geq}$-decision rule and at least one $D_{\leq}$-decision rule, or at least one $D_{\geq\leq}$-decision rule, or at least one $D_{\geq\leq}$-decision rule and at least one $D_{\geq}$- and/or at least one $D_{\leq}$-decision rule,
- no decision rule.

These four ways correspond to the following four situations of outranking, respectively:

- uSv and *not* uS^cv, i.e. *true* outranking (denoted by uS^Tv)
- uS^cv and *not* uSv, i.e. *false* outranking (denoted by uS^Fv)
- uSv and uS^cv, i.e. *contradictory* outranking (denoted by uS^Kv)
- *not* uSv and *not* uS^cv, i.e. *unknown* outranking (denoted by uS^Uv)

The four above situations, which together constitute the so-called *four-valued outranking* [24], have been introduced to underline the presence and absence of *positive* and *negative* reasons for the outranking. Moreover, they make it possible to distinguish contradictory situations from unknown ones.

A final *recommendation* (choice or ranking) can be obtained upon a suitable exploitation of this structure, i.e. of the presence and the absence of outranking S and S^c on M. A possible exploitation procedure consists of calculating a specific score, called the Net Flow Score, for each object $x \in M$:

$S_{nf}(x) = S^{++}(x) - S^{+-}(x) + S^{-+}(x) - S^{--}(x)$, where

$S^{++}(x) = \text{card}(\{y \in M$: there is at least one decision rule which affirms $xSy\})$

$S^{+-}(x) = \text{card}(\{y \in M$: there is at least one decision rule which affirms $ySx\})$

$S^{-+}(x) = \text{card}(\{y \in M$: there is at least one decision rule which affirms $yS^cx\})$

$S^{--}(x) = \text{card}(\{y \in M$: there is at least one decision rule which affirms $xS^cy\})$

The recommendation in ranking problems consists of the total preorder determined by $S_{nf}(x)$ on M. In choice problems, it consists of the object(s) $x^* \in M$ such that $S_{nf}(x^*) = \max_{x \in M} \{S_{nf}(x)\}$.

The above procedure has been characterized with reference to a number of desirable properties in [24].

3.5. An illustrative example

Let us suppose that a company managing a chain of warehouses wants to buy some new warehouses. To choose the best proposals or to rank them all, the managers of the company decide to analyze first the characteristics of eight warehouses already owned by the company (reference objects). This analysis should give some indications for the choice and ranking of the new proposals. Eight warehouses belonging to the company have been evaluated by the following three criteria: capacity of the sales staff (A_1), perceived quality of goods (A_2) and high traffic location (A_3). The domains (scales) of these attributes are presently composed of three preference-ordered echelons: $V_1 = V_2 = V_3 = \{$sufficient, medium, good$\}$. The decision attribute (d) indicates the profitability of warehouses, expressed by the *Return On Equity* (*ROE*) ratio (in %). Table 2 presents a decision table which represents this situation.

Table 2

Decision table with reference objects

Warehouse	A_1	A_2	A_3	d (*ROE*%)
1	good	medium	good	10.35
2	good	sufficient	good	4.58
3	medium	medium	good	5.15
4	sufficient	medium	medium	-5
5	sufficient	medium	medium	2.42
6	sufficient	sufficient	good	2.98
7	good	medium	good	15
8	good	sufficient	good	-1.55

With respect to the set of criteria $C = C^N = \{A_1,A_2,A_3\}$, the following multigraded preference relations P_i^h, $i = 1,2,3$, are defined:

- $x\,P_i^0\,y$ (and $y\,P_i^0\,x$), meaning that x is *indifferent* to y with respect to A_i, if $f(x,A_i) = f(y,A_i)$,
- $x\,P_i^1\,y$ (and $y\,P_i^{-1}\,x$), meaning that x is *preferred* to y with respect to A_i, if $f(x,A_i) =$ good and $f(y,A_i) =$ medium, or if $f(x,A_i) =$ medium and $f(y,A_i) =$ sufficient,
- $x\,P_i^2\,y$ (and $y\,P_i^{-2}\,x$), meaning that x is *strongly preferred* to y with respect to A_i, if $f(x,A_i) =$ good and $f(y,A_i) =$ sufficient.

Using the decision attribute, the comprehensive outranking relation was built as follows: warehouse x is at least as good as warehouse y with respect to profitability (xSy) if

$$ROE(x) \geq ROE(y) - 2\%.$$

Otherwise, i.e. if $ROE(x) < ROE(y) - 2\%$, warehouse x is *not* at least as good as warehouse y with respect to profitability ($xS^c y$).

The pairwise comparisons of the reference objects result in $\boldsymbol{S}_{PCT}$. The rough set analysis of the $\boldsymbol{S}_{PCT}$ leads to the conclusion that the set of decision examples on the reference objects is inconsistent. The quality of approximation of S and S^c by all criteria from set C is equal to 0.44. Moreover, $RED_{\boldsymbol{S}_{PCT}} = CORE_{\boldsymbol{S}_{PCT}} = \{A_1, A_2, A_3\}$. This means that no criterion is superfluous.

The C-lower approximations and the C-upper approximations of S and S^c, obtained by means of multigraded dominance relations, are:

$\underline{C}(S)$ = {(1,2),(1,4),(1,5),(1,6),(1,8),(3,2),(3,4),(3,5),(3,6),(3,8),(7,2),(7,4),(7,5),(7,6),(7,8)}

$\underline{C}(S^c)$ = {(2,1),(2,7),(4,1),(4,3),(4,7),(5,1),(5,3),(5,7),(6,1),(6,3),(6,7),(8,1),(8,7)}

All the remaining 36 pairs of reference objects belong to the C-boundaries of S and S^c, i.e. $Bn_C(S) = Bn_C(S^c)$.

The following minimal $D_{\geq}$-decision rules and $D_{\leq}$-decision rules can be induced from lower approximations of S and S^c, respectively (the figures within parentheses represent the pairs of objects supporting the corresponding rules):

if $x P_1^{\geq 1} y$ *and* $x P_2^{\geq 1} y$, *then* xSy; ((1,6),(3,6),(7,6))

if $x P_2^{\geq 1} y$ *and* $x P_3^{\geq 0} y$, *then* xSy; ((1,2),(1,6),(1,8),(3,2),(3,6),(3,8),(7,2),(7,6),(7,8))

if $x P_2^{\geq 0} y$ *and* $x P_3^{\geq 1} y$, *then* xSy; ((1,4),(1,5),(3,4),(3,5),(7,4),(7,5))

if $x P_1^{\leq -1} y$ *and* $x P_2^{\leq -1} y$, *then* $xS^c y$; ((6,1),(6,3),(6,7))

if $x P_2^{\leq 0} y$ *and* $x P_3^{\leq -1} y$, *then* $xS^c y$; ((4,1),(4,3),(4,7),(5,1),(5,3),(5,7))

if $x P_1^{\leq 0} y$ *and* $x P_2^{\leq -1} y$ *and* $x P_3^{\leq 0} y$, *then* $xS^c y$; ((2,1),(2,7),(6,1),(6,3),(6,7),(8,1),(8,7))

Moreover, it is possible to induce five minimal $D_{\geq\leq}$-decision rules from the boundary of approximation of S and S^c:

if $x P_2^{\leq 0} y$ *and* $x P_2^{\geq 0} y$ *and* $x P_3^{\leq 0} y$ *and* $x P_3^{\geq 0} y$, *then* xSy *or* $xS^c y$;
((1,1),(1,3),(1,7),(2,2),(2,6),(2,8),(3,1),(3,3),(3,7),(4,4),(4,5),(5,4),(5,5),(6,2),(6,6),(6,8),(7,1),(7,3),(7,7),(8,2),(8,6),(8,8))

if $x P_2^{\leq -1} y$ *and* $x P_3^{\geq 1} y$, *then* xSy *or* $xS^c y$; ((2,4),(2,5),(6,4),(6,5),(8,4),(8,5))

if $x P_2^{\geq 1} y$ *and* $x P_3^{\leq -1} y$, *then* xSy *or* $xS^c y$; ((4,2),(4,6),(4,8),(5,2),(5,6),(5,8))

if $x P_1^{\geq 1} y$ *and* $x P_2^{\leq 0} y$ *and* $x P_3^{\leq 0} y$, *then* xSy *or* $xS^c y$; ((1,3),(2,3),(2,6),(7,3),(8,3),(8,6)),

if $x P_1^{\geq 1} y$ *and* $x P_2^{\leq -1} y$, *then* xSy *or* $xS^c y$; ((2,3),(2,4),(2,5),(8,3),(8,4),(8,5))

Using all the above decision rules and the Net Flow Score exploitation procedure on ten other warehouses proposed for purchase, the managers can obtain the result presented in Table 3. The DRSA gives a clear recommendation:

- For the **choice problem** it suggests the selection of warehouse 2' and 6', having maximum score (11)
- For the **ranking problem** it suggests the ranking presented in the last column of Table 3, as follows:

$$(2',6') \rightarrow (8') \rightarrow (9') \rightarrow (1') \rightarrow (4') \rightarrow (5') \rightarrow (3') \rightarrow (7',10')$$

Table 3

Ranking of warehouses for sale by decision rules and the Net Flow Score procedure

Warehouse for sale	A_1	A_2	A_3	Net Flow Score	Ranking
1'	good	sufficient	medium	1	5
2'	sufficient	good	good	11	1
3'	sufficient	medium	sufficient	-8	8
4'	sufficient	good	sufficient	0	6
5'	sufficient	sufficient	medium	-4	7
6'	sufficient	good	good	11	1
7'	medium	sufficient	sufficient	-11	9
8'	medium	medium	medium	7	3
9'	medium	good	sufficient	4	4
10'	medium	sufficient	sufficient	-11	9

Summary

We briefly presented the contribution of the DRSA to multiple criteria choice and ranking problems. Let us point out the main features of the described methodology:

- The decision maker is asked for the preference information necessary to deal with a multiple criteria decision problem in terms of exemplary decisions.
- The rough set analysis of preferential information supplies some useful elements of knowledge about the decision situation. These are: the relevance of particular attributes and/or criteria, information about their interaction, minimal subsets of attributes or criteria (reducts) conveying important knowledge contained in the exemplary decisions and the set of the non-reducible attributes or criteria (core).
- The preference model induced from the preferential information is expressed in a natural and comprehensible language of "*if*..., *then*..." decision rules. The decision rules concern pairs of objects and from them we can determine either the presence or the absence of a comprehensive preference relation. The conditions for the presence are expressed in "at least" terms, and for the absence in "at most" terms, on particular criteria.

- The decision rules do not convert ordinal information into numeric but keep the ordinal character of input data due to the syntax proposed.
- Heterogeneous information (qualitative and quantitative, ordered and non-ordered) and scales of preference (ordinal, cardinal) can be processed within the DRSA, while classical methods consider only quantitative ordered evaluations (with rare exceptions).
- No prior discretization of the quantitative domains of criteria is necessary.

Rough approximations of a comprehensive preference relation can be defined using other types of dominance than the Pareto dominance used in this section. In [43], the Lorenz dominance has been used for rough approximations, permitting induction of more robust decision rules, i.e. certain decision rules supported by consistent pairs of objects characterized by equitable distributions of intensities of preference on considered criteria

4. DRSA FOR DECISION UNDER RISK

In [13], we opened a new avenue for applications of the rough set concept. This avenue leads to the classical problem of *decision under risk*. To adapt the DRSA to this problem, we substituted the dominance relation by *stochastic dominance relation* defined on a set of objects meaning acts. We considered the case of traditional additive probability distribution over a set of states of the world, however, the model is rich enough to handle non-additive probability distributions and even qualitative ordinal distributions. The adapted DRSA gives a representation of DM's preferences under risk in terms of "*if..., then*..." decision rules induced from rough approximation of preference ordered classification of acts described in terms of outcomes in uncertain states of the world. The preference ordered classification constitutes, in this case, preferential information acquired from the DM.

4.1. DRSA based on stochastic dominance

To apply DRSA to decision under risk, we consider the following basic elements:

- a set $S = \{s_1, s_2, \ldots, s_s\}$ of states of the world, or simply *states*, which are supposed to be mutually exclusive and collectively exhaustive,
- an a priori probability distribution P over the states of the world: more precisely, the probabilities of states $s_1, s_2, \ldots, s_s$ are $p_1, p_2, \ldots, p_s$, respectively $(p_1 + p_2 + \ldots + p_s = 1, p_i \geq 0, i = 1,\ldots s)$,

- a set $A = \{A_1, A_2, \ldots, A_m\}$ of *acts*,
- a set $X = \{x_1, x_2, \ldots, x_r\}$ of consequences or outcomes that, for the sake of simplicity, are supposed to be expressed in monetary terms, thus $X \subseteq \boldsymbol{R}$,
- a function g: $A \times S \rightarrow X$ assigning to each act-state pair $(A_i, s_j) \in A \times S$ a consequence $x_h \in X$,
- a set of classes $\boldsymbol{Cl} = \{Cl_1, Cl_2, \ldots, Cl_n\}$, such that $Cl_1 \cup Cl_2 \cup \ldots \cup Cl_n = A$, $Cl_p \cap Cl_q = \emptyset$ for each $p,q \in \{1,2\ldots,n\}$ with $p \neq q$; the classes of $\boldsymbol{Cl}$ are preference ordered according to an increasing order of their indices, in the sense that for each $A_i, A_j \in A$, if $A_i \in Cl_p$ and $A_j \in Cl_q$ with $p > q$, then A_i is preferred to A_j,
- a function e: $A \rightarrow \boldsymbol{Cl}$ assigning each act $A_i \in A$ to a class $Cl_t \in \boldsymbol{Cl}$.

In this context, two different types of dominance relations can be considered:

1) (classical) *dominance*: given $A_i, A_j \in A$, A_i dominates A_j iff for each possible state of the world act A_i gives an outcome at least as good as act A_j; more formally, $g(A_i,s_k) \geq g(A_j,s_k)$, for each $s_k \in S$,
2) *stochastic dominance*: given $A_i, A_j \in A$, A_i stochastically dominates A_j iff for each outcome $x \in X$, act A_i gives an outcome at least as good as x with a probability at least as large as act A_j.

Case 1) corresponds to a model in which the utility is state dependent while case 2) corresponds to a model of decision under risk proposed by. We consider the second case.

On the basis of an a priori probability distribution P, we can assign to each subset of states of the world $W \subseteq S$ ($W \neq \emptyset$) the probability $P(W)$ that one of the states in W is verified, i.e. $P(W) = \sum_{i:s_i \in W} p_i$ and then to build up the set $\boldsymbol{\Pi}$ of all the possible values $P(W)$, i.e.

$$\boldsymbol{\Pi} = \{\pi \in [0,1]: \pi = P(W), W \subseteq S\}$$

We define the following function z: $A \times S \rightarrow \boldsymbol{\Pi}$, assigning to each act-state pair $(A_i,s_j) \in A \times S$ a probability $\pi \in \boldsymbol{\Pi}$, as follows:

$$z(A_i,s_j) = \sum_{r:g(A_i,s_r) \geq g(A_i,s_j)} p_r$$

Therefore, $z(A_i,s_j)$ represents the probability of obtaining by act A_i an outcome whose value is at least $g(A_i,s_j)$.

On the basis of function $z(A_i, s_j)$, we can define the function ρ: $A \times \boldsymbol{\Pi} \rightarrow X$ as follows:

$$\rho(A_i,\pi) = \min_{j:z(A_i,s_j) \geq \pi} g(A_i,s_j)$$

Thus, $\rho(A_i,\pi) = x$ means that by act A_i one can gain *at least* x with a probability greater than or equal to π.

Using function $z(A_i, s_j)$, we can also define function ρ': $A \times \boldsymbol{\Pi} \rightarrow X$ as follows:

$$\rho'(A_i,\pi) = \max_{j:z(A_i,s_j)\leq\pi} g(A_i,s_j)$$

$\rho'(A_i,\pi) = x$ means that by act A_i one can gain *at most x* with a probability smaller than or equal to π.

If the elements of $\boldsymbol{\Pi}$, $0 = \pi_{(1)}, \pi_{(2)}, \ldots, \pi_{(w)} = 1$ ($w = \text{card}(\boldsymbol{\Pi})$), are reordered in such a way that $\pi_{(1)} \leq \pi_{(2)} \leq \ldots \leq \pi_{(w)}$, then we have $\rho(A_i,\pi_{(j)}) = \rho'(A_i, 1-\pi_{(j-1)})$.

Therefore, $\rho(A_i,\pi_{(j)}) \leq x$ is equivalent to $\rho'(A_i, 1-\pi_{(j-1)}) \geq x$, $A_i \in A$, $\pi_{(j)} \in \boldsymbol{\Pi}$, $x \in X$.

Given $A_i, A_j \in A$, A_i stochastically dominates A_j if and only if $\rho(A_i,\pi) \geq \rho(A_j,\pi)$ for each $\pi \in \boldsymbol{\Pi}$. This is equivalent to say: given $A_i, A_j \in A$, A_i stochastically dominates A_j if and only if $\rho'(A_i,\pi) \leq \rho'(A_j,\pi)$ for each $\pi \in \boldsymbol{\Pi}$.

We can apply DRSA in this context considering the following correspondence:

- the universe U is the set of acts A,
- the set of condition attributes (criteria) C is the set $\boldsymbol{\Pi}$,
- the domain V_π of each criterion $\pi \in \boldsymbol{\Pi}$ is the set X,
- the single decision attribute d specifies classification of acts from A into classes from $\boldsymbol{Cl}$,
- the information function f is a function $\boldsymbol{f}$ such that for all $A_i \in$ A and $\pi \in \boldsymbol{\Pi}$, $\boldsymbol{f}(A_i,\pi) = \rho(A_i,\pi)$ and $\boldsymbol{f}(A_i,d) = e(A_i)$,
- the dominance relation on U is the stochastic dominance relation on A.

The aim of DRSA to decision under risk is to explain the preferences of the DM, represented by his/her assignments of the acts from A to the classes of $\boldsymbol{Cl}$, in terms of decision rules involving stochastic dominance on partial profiles corresponding to outcomes x for some probabilities π.

4.2. An illustrative example

The following example illustrates the approach. Let us consider

- set $S = \{s_1, s_2, s_3\}$ of states of the world,
- a priori probability distribution P over the states of the world defined as follows: $p_1 = 0.25$, $p_2 = 0.35$, $p_3 = 0.40$,
- set $A = \{A_1, A_2, A_3, A_4, A_5, A_6\}$ of acts,
- set $X = \{0, 10, 15, 20, 30\}$ of consequences,
- set of classes $\boldsymbol{Cl} = \{Cl_1, Cl_2, Cl_3\}$, where Cl_1 is the set of *bad* acts, Cl_2 is the set of *medium* acts, Cl_3 is the set of *good* acts,
- function $g: A \times S \rightarrow X$ assigning to each act-state pair $(A_i,s_j) \in A \times S$ a consequence $x_h \in X$, and a function e: $A \rightarrow \boldsymbol{Cl}$ assigning each act $A_i \in$ A to a class $Cl_t \in \boldsymbol{Cl}$, as presented in Table 4.

Table 4

Acts, consequences and assignment to classes from ***Cl***

	p_j	A_1	A_2	A_3	A_4	A_5	A_6
s_1	0.25	30	0	15	0	20	10
s_2	0.35	10	20	0	15	10	20
s_3	0.40	10	20	20	20	20	20
d		*good*	*medium*	*medium*	*bad*	*medium*	*good*

DRSA is applied on Table 5 including the values of function $\rho(A_i,\pi)$. Let us explain what mean the entries in Table 5. If we consider the column of act, say A_3, we see that by act A_3,

- the value 20 in the row corresponding to 0.25 means that the outcome is at least 20 with a probability of at least 0.25,
- the value 15 in the row corresponding to 0.65 means that the outcome is at least 15 with a probability of at least 0.65,
- the value 0 in the row corresponding to 0.75 means that the outcome is at least 0 with a probability of at least 0.75.

Table 5

Acts, values of function $\rho(A_i,\pi)$ and assignment to classes from ***Cl***

	A_1	A_2	A_3	A_4	A_5	A_6
0.25	30	20	20	20	20	20
0.35	10	20	20	20	20	20
0.40	10	20	20	20	20	20
0.60	10	20	15	15	20	20
0.65	10	20	15	15	20	20
0.75	10	20	0	15	10	20
1	10	0	0	0	10	10
d	*good*	*medium*	*medium*	*bad*	*medium*	*good*

If we consider the row corresponding to 0.65, then

- the value 10 relative to A_1, means that by act A_1 the outcome is at least 10 with a probability of at least 0.65,
- the value 20 relative to A_2, means that by act A_2 the outcome is at least 20 with a probability of at least 0.65, and so on.

Applying DRSA, we approximate the following *upward* and *downward unions* of classes:

- $Cl_2^{\geq} = Cl_2 \cup Cl_3$, i.e. the set of the acts at least *medium*,
- $Cl_3^{\geq} = Cl_3$, i.e. the set of the acts (at least) *good*,

- $Cl_1^{\leq} = Cl_1$, i.e. the set of the acts (at most) *bad*,
- $Cl_2^{\leq} = Cl_1 \cup Cl_2$, i.e. the set of the acts at most *medium*.

The first result of the DRSA is a discovery that the data table (Table 5) is not consistent. Indeed, Table 5 shows that act A_4 stochastically dominates act A_3, however act A_3 is assigned to a better class (*medium*) than act A_4 (*bad*). Therefore, act A_3 cannot be assigned without doubt to the class of at least *medium* acts as well as act A_4 cannot be assigned without doubt to the class of (at most) *bad* acts. In consequence, lower approximation and upper approximation of $Cl_2^{\geq}$, $Cl_3^{\geq}$ and $Cl_1^{\leq}$, $Cl_2^{\leq}$ are equal, respectively, to

- $\underline{C}\left(Cl_2^{\geq}\right) = \{A_1,A_2,A_5,A_6\} = Cl_2^{\geq} - \{A_3\}$,
  $\overline{C}\left(Cl_2^{\geq}\right) = \{A_1,A_2,A_3,A_4,A_5,A_6\} = Cl_2^{\geq} \cup \{A_4\}$,
- $\underline{C}\left(Cl_3^{\geq}\right) = \{A_1,A_6\} = Cl_3^{\geq}$, $\overline{C}\left(Cl_3^{\geq}\right) = \{A_1,A_6\} = Cl_3^{\geq}$,
- $\underline{C}\left(Cl_1^{\leq}\right) = \varnothing = Cl_1^{\leq} - \{A_4\}$, $\overline{C}\left(Cl_1^{\leq}\right) = \{A_3,A_4\} = Cl_1^{\leq} \cup \{A_3\}$,
- $\underline{C}\left(Cl_2^{\leq}\right) = \{A_2,A_3,A_4,A_5\} = Cl_2^{\leq}$, $\overline{C}\left(Cl_2^{\leq}\right) = \{A_2,A_3,A_4,A_5\} = Cl_2^{\leq}$.

Since there are two inconsistent acts on a total of six acts (A_3,A_4), then the quality of approximation (quality of classification) is equal to 4/6.

The second discovery is one reduct of condition attributes (criteria) ensuring the same quality of classification as the whole set $\boldsymbol{\Pi}$ of probabilities: $RED_{\boldsymbol{Cl}} = \{0.25, 0.75, 1\}$. This means that we can explain the preferences of the DM using the probabilities in $RED_{\boldsymbol{Cl}}$ only. $RED_{\boldsymbol{Cl}}$ is also the core because no probability value can be removed from $RED_{\boldsymbol{Cl}}$ without deteriorating the quality of classification.

The third discovery gives sets of decision rules describing the DM's preferences. Below, we are presenting one of minimal sets of decision rules covering all the acts [*within brackets there is a verbal interpretation of the corresponding decision rule*] (within parentheses there are acts supporting the corresponding rule):

1) *if* $\rho(A_i,0.25) \geq 30$, *then* $A_i \in Cl_3^{\geq}$,
   [*if the probability of gaining at least 30 is at least 0.25, then act* A_i *is (at least) good*] (A_1),
2) *if* $\rho(A_i,0.75) \geq 20$ *and* $\rho(A_i,1) \geq 10$, *then* $A_i \in Cl_3^{\geq}$,
   [*if the probability of gaining at least 20 is at least 0.75 and the probability of gaining at least 10 is (at least) 1 (i.e. for sure the gain is at least 10), then act* A_i *is (at least) good*] (A_6),
3) *if* $\rho(A_i,1) \geq 10$, *then* $A_i \in Cl_2^{\geq}$,
   [*if the probability of gaining at least 10 is (at least) 1 (i.e. for sure the gain is at least 10), then act* A_i *is at least medium*] (A_1, A_5, A_6),

4) *if* $\rho(A_i,0.75)\geq 20$, *then* $A_i \in Cl_2^{\geq}$,
   [*if the probability of gaining at least 20 is at least 0.75, then act* A_i *is at least medium*] (A_2, A_6),
5) *if* $\rho(A_i,0.25)\leq 20$ (*i.e.* $\rho'(A_i,1)\geq 20$) *and* $\rho(A_i,0.75)\leq 15$ (*i.e.* $\rho'(A_i,0.35)\geq 15$), *then* $A_i \in Cl_2^{\leq}$,
   [*if the probability of gaining at most 20 is (at least) 1 (i.e. for sure the gain is at most 20) and the probability of gaining at most 15 is at least 0.35, then act* A_i *is at most medium*] (A_3, A_4, A_5),
6) *if* $\rho(A_i,1)\leq 0$ *(i.e.* $\rho'(A_i,0.25)\geq 0$*), then* $A_i \in Cl_1 \cup Cl_2$,
[*if the probability of gaining at most 0 is at least 0.25, then act* A_i *is at most medium*] (A_2, A_3, A_4),
7) *if* $\rho(A_i,1)\geq 0$ *and* $\rho(A_i,1)\leq 0$ (i.e. $\rho(A_i,1)=0$) *and* $\rho(A_i,0.75)\leq 15$ (*i.e.* $\rho'(A_i,0.35)\geq 10$), *then* $A_i \in Cl_1 \cup Cl_2$,
   [*if the probability of gaining at least 0 is 1 (i.e. for sure the gain is at least 0) and the probability of gaining at most 15 is at least 0.35, then act* A_i *is bad or medium, with no possibility of assigning* A_i *to only one of the two classes because of ambiguous knowledge*] (A_3, A_4).

Minimal sets of decision rules represent the most concise and non-redundant knowledge contained in Table 4 (and, consequently, in Table 5). The above minimal set of 7 decision rules uses 3 attributes (probabilities 0.25, 0.75 and 1) and 11 elementary conditions, i.e. 26% of descriptors from the original data table (Table 5). For larger sets of exemplary acts, the representation in terms of decision rules is even more synthetic (the percentage of descriptors from the original data table is smaller).

Let us observe that we considered an additive probability distribution, however, an extension to non-additive probability, and even to a qualitative ordinal probability, is straightforward. If the elements of set Π are numerous (like in real applications), a subset $\Pi' \subset \Pi$ of the most significant probability values (e.g. 0, 0.1, 0.2, …, 0.9, 1) can be considered.

5. COMPARISON OF DRSA WITH OTHER DECISION SUPORT PARADIGMS

DRSA aims to give an effective answer to the central problem of any decision-aiding methodology concerning multiple criteria and/or multiple attribute classification, that is the aggregation of the multiple criteria and attributes into a single preference model. In this section, we propose to compare different para-

digms used to solve this central problem by different theories (see Table 6). In [17, 42], this comparison was made at the level of axiomatic foundations, which has no precedence in the theoretical research concerning multiple criteria classification. The axiomatic approach is interesting for at least three reasons:

- it exhibits differences between preference models and methods,
- it permits to interpret methods conceived for one model in terms of another model,
- knowing the basic axioms, one can pass from one method to another with different preference models.

Table 6

Different paradigms of aggregation and preference representation

Theory (paradigm)	**Main preoccupation (axiomatic basis)**	**The aggregation result evidences**
Social Choice Theory	Voting system or aggregation of rankings	Final ranking
Decision Theory	Definition of preference structures	Relation in A
Measurement Theory	Cancellation property	Function, like in conjoint measurement
Measure Theory, Fuzzy Sets	Capacity or fuzzy measure	Weights or interaction between criteria, like in Choquet integral or Sugeno integral
Machine Learning, Logical Analysis of Data, Rough Sets	Boolean or pseudo-Boolean function, decision rules or decision trees	Knowledge, like in knowledge discovery or data mining

Moreover, in [17, 42], we have considered aggregation of ordinal criteria that has been studied much less than that of cardinal criteria. Among several aggregation models, a particular interest has been paid recently for an integral proposed by Sugeno, able to deal with ordinal data; it has been considered the most general ordinal aggregation function of the max-min average type. It appears, however, that this function has some unpleasant limitations: the most important is the so-called commensurability, i.e. the evaluations with respect to each considered criterion should be defined on the same scale. Comparison of the Sugeno integral with the decision rule model at the axiomatic level permits to show other limitation of the former.

Below, we present the main results concerning the comparison of axiomatic foundations of the decision rule model and two traditional models: utility function and outranking relation.

5.1. Axiomatic foundations of multiple criteria classification problems and associated preference models

In this point we consider a finite or denumerable *product space* $X = \prod_{i=1}^{n} X_i$, where X_i is an evaluation scale of *criterion* $i = 1,\ldots,n$. With appropriate topological conditions we can also work with infinite non-denumerable space, but in this paper, for the sake of simplicity, we will skip this possibility.

The following result is a representation theorem for the multiple criteria classification problem, stating equivalence between a very simple cancellation property, a general utility function, a very general outranking relation and a set of decision rules. Let us mention that equivalence of the considered cancellation property and the utility function was already noted by Goldstein (1991), within the conjoint measurement approach, for the special case of three decision classes.

Theorem 1 [17]. The following four propositions are equivalent:

1) (*cancellation property*) for each $i = 1,\ldots,n$, for each $x_i, y_i \in X_i$ and $a_{-i}, b_{-i} \in X_{-i}$, and for each $r,s \in \{1,\ldots,m\}$:

$$\{(x_i a_{-i}) \in Cl_r \text{ and } (y_i b_{-i}) \in Cl_s\} \Rightarrow \{(y_i a_{-i}) \in Cl_r^{\geq} \text{ or } (x_i b_{-i}) \in Cl_s^{\geq}\}$$

2) (*utility function*) there exist:
- functions $g_i: X_i \rightarrow \boldsymbol{R}$ for each $i = 1,\ldots,n$, called criteria,
- function $f: \boldsymbol{R}^n \rightarrow \boldsymbol{R}$, non-decreasing in each argument, called utility function,
- m-1 ordered thresholds z_t, $t = 2,\ldots,m$, satisfying

$$z_2 < z_3 < \ldots. < z_m$$

such that for each $x \in X$ and each $t = 2,\ldots,m$

$$f[g_1(x_1), g_2(x_2),\ldots, g_n(x_n)] \geq z_t \Leftrightarrow x \in Cl_t^{\geq}$$

3) (*outranking function and relation*) there exist
- functions $g_i: X_i \rightarrow \boldsymbol{R}$, $i = 1,\ldots,n$, called criteria,
- function $s: \boldsymbol{R}^{2n} \rightarrow \boldsymbol{R}$, non-decreasing in each odd argument and non-increasing in each even argument, called outranking function,
- m-1 reference profiles p^t, $t = 2,\ldots,m$, satisfying

$$g_i(p^2) \leq g_i(p^3) \leq \ldots \leq g_i(p^m), \text{ for } i = 1,\ldots,n$$

such that for each $x \in X$ and each $t = 2,\ldots,m$

$$s[g_1(x_1), g_1(p^t), g_2(x_2), g_2(p^t),\ldots, g_n(x_n), g_n(p^t)] \geq 0 \Leftrightarrow x \in Cl_t^{\geq}$$

$$(\text{N.B. } s[g_1(x_1), g_1(p^t), g_2(x_2), g_2(p^t),\ldots, g_n(x_n), g_n(p^t)] \geq 0 \Leftrightarrow x\, S\, p^t,$$

where S is a binary outranking relation),

4) ("*at least*" *decision rules*) there exist:

- functions g_i: $X_i \rightarrow \boldsymbol{R}$ for each $i = 1,\ldots,n$, called criteria,
- a set of "at least" decision rules whose syntax is

$$\textit{if } g_{i1}(x_{i1}) \geq r_{i1} \textit{ and } g_{i2}(x_{i2}) \geq r_{i2} \textit{ and } \ldots \textit{ and } g_{ih}(x_{ih}) \geq r_{ih}, \textit{ then } x \in Cl_t^{\geq},$$

$$\text{with } \{i1,i2,\ldots,ih\} \subseteq \{1,\ldots,n\},\ t = 2,\ldots,m,$$

such that for each $y \in Cl_t$, $t = 2,\ldots,m$, there is at least one rule implying $y \in Cl_t^{\geq}$ and there is no rule implying $y \in Cl_r^{\geq}$, with $r>t$.

Let us remark that the above representation theorem for multiple criteria classification problem starts with a very weak axiomatic condition called cancellation property. Indeed, this property does not require existence of criterion functions g_i, $i = 1,\ldots,n$, or a dominance relation D on X in order to characterize the three preference models. Instead, the meaning of the above cancellation property is the following. Let us consider the binary large preference relation $\succeq_i$ defined on X_i, $i = 1,\ldots,n$, as follows: for each for each $x_i,y_i \in X_i$, for each for each $a_{-i} \in X_{-i}$ and for each $Cl_r \in \boldsymbol{Cl}$:

$$x_i \succeq_i y_i \Leftrightarrow [(y_i a_{-i}) \in Cl_r \Rightarrow (y_i a_{-i}) \in Cl_r^{\geq}]$$

Cancellation property ensures that the binary large preference relation $\succeq_i$ on X_i is a complete preorder, that is strongly complete (for each $x_i,y_i \in X_i$, $x_i \succeq_i y_i$ or $y_i \succeq_i x_i$) and transitive. Consequently, there exists a function g_i: $X_i \rightarrow \boldsymbol{R}$ such that for each $x_i,y_i \in X_i$

$$x_i \succeq_i y_i \Leftrightarrow g_i(x_i) \geq g_i(y_i)$$

On the basis of relations $\succeq_i$, $i = 1,...,n$, one can also define a dominance relation D on X as follows: for each $x,y \in X$

$$xDy \Leftrightarrow x_i \succeq_i y_i \text{ for all } i = 1,...,n$$

This is of course equivalent to

$$xDy \Leftrightarrow g_i(x_i) \geq g_i(y_i) \text{ for all } i = 1,...,n$$

Cancellation property 1) of Theorem 1 permits to state the following *condition of coherence between dominance relation D and classification* ***Cl***, for each $x,y \in X$

$$xDy \Rightarrow x \in Cl_r \text{ and } y \in Cl_s, \text{ with } r \geq s$$

For any subset of criteria $P \subseteq \{1,\ldots,n\}$ and for each pair $x,y \in X$ one can also define a dominance relation D_P on X:

$$xD_P y \Leftrightarrow x_i \succeq_i y_i \text{ for all } i \in P$$

which is equivalent to

$$xD_Py \Leftrightarrow g_i(x_i) \geq g_i(y_i) \text{ for all } i \in P$$

Dominance relations D_P, $P \subseteq \{1,\ldots,n\}$, are used in the condition part of decision rules. Being an intersection of complete preorders, binary relations D_P are partial preorders, i.e. they are reflexive and transitive.

Observe, moreover, that Theorem 1 regards a representation of classification ***Cl*** in terms of "lower bounds". Theorem 1 can be reformulated in terms of "upper bounds" in such a way that:

– condition of proposition 2) is expressed as

$$f[g_1(x_1), g_2(x_2),\ldots, g_n(x_n)] \leq w_t \Leftrightarrow x \in Cl_t^{\leq},$$

where w_t, $t = 1,\ldots,m\text{-}1$, are m-1 suitably ordered thresholds,

– condition of proposition 3) is expressed as

$$s[g_1(x_1), g_1(q^t), g_2(x_2), g_2(q^t),\ldots, g_n(x_n), g_n(q^t)] < 0 \Leftrightarrow x \in Cl_t^{\leq}$$

where q^t, $t = 1,\ldots,m\text{-}1$, are m-1 reference profiles q^t, such that q^{t+1} dominates q^t (i.e. q^{t+1} is at least as good as q^t with respect to each criterion $i = 1,\ldots,n$, and there is at least one criterion $j \in \{1,\ldots,n\}$ for which q^{t+1} is strictly preferred to q^t), $t = 1,\ldots,m\text{-}2$,

– condition of proposition 4) considers a set of decision rules whose syntax is

$$\textit{if } g_{i1}(x_{i1}) \leq r_{i1} \textit{ and } g_{i2}(x_{i2}) \leq r_{i2} \textit{ and } \ldots \textit{ and } g_{ih}(x_{ih}) \leq r_{ih}, \textit{ then } x \in Cl_r^{\leq}$$

with $\{i1,i2,\ldots,ih\} \subseteq \{1,\ldots,n\}$. These decision rules are called "*at most*" decision rules. The classification of $x \in X$ with "at most" decision rules is done according to the following procedure:

– $x \in Cl_t$ if and only if there exists a rule matching x that assigns x to $Cl_t^{\leq}$, and there exists no rule matching x that assigns x to $Cl_s^{\leq}$, where $s<t$;

– $x \in Cl_m$ if and only if there exists no rule matching x.

6.2. Conjoint measurement for multiple criteria classification problems with inconsistencies

The conjoint measurement model presented in point 6.1 cannot represent the inconsistency with the dominance principle considered within the DRSA. In this point we present a more general model of conjoint measurement that permits representation of this inconsistency. This model is based on the concepts of dominance-based rough approximation of upward and downward unions of classes $Cl_t^{\geq}$ and $Cl_t^{\leq}$.

The following concepts will be useful: for each $x \in X$, the *lower class* and the *upper class* of x, denoted by $r_*(x)$ and $r^*(x)$, respectively, are defined as follows

$$r_*(x) = \max\{s \in \{1,..,m\} : x \in \underline{C}\left(Cl_s^{\geq}\right)\}$$

$$r^*(x) = \min\{s \in \{1,..,m\} : x \in \underline{C}\left(Cl_s^{\leq}\right)\}$$

where $\underline{C}\left(Cl_s^{\geq}\right)$ and $\underline{C}\left(Cl_s^{\leq}\right)$ are the lower approximations of $Cl_s^{\geq}$ and $Cl_s^{\leq}$, respectively, with respect to set of criteria $C = \{g_1,g_2,\ldots,g_n\}$.

Theorem 2 [17]. For each set of binary relations $\succeq_i$, $i = 1,\ldots,n$, being complete preorders, and for each classification ***Cl*** there exist

- functions g_i: $X_i \rightarrow \boldsymbol{R}$, such that $x_i \succeq_i y_i \Leftrightarrow g_i(x_i) \geq g_i(y_i)$, $i = 1,\ldots,n$,
- functions $f^{\geq}$: $\boldsymbol{R}^n \rightarrow \boldsymbol{R}$ and $f^{\leq}$: $\boldsymbol{R}^n \rightarrow \boldsymbol{R}$, non-decreasing in each argument, such that
- $f^{\geq}[g_1(x_1),g_2(x_2),\ldots,g_n(x_n)] \leq f^{\leq}[g_1(x_1),g_2(x_2),\ldots,g_n(x_n)]$
- m-1 ordered thresholds z_t, t=2,…,m,

$$z_2 < z_3 < \ldots < z_m$$

such that for each object $x \in X$, functions $f^{\geq}$ and $f^{\leq}$ assign x to a lower and an upper class, respectively:

$$f^{\geq}[g_1(x_1),g_2(x_2),\ldots,g_n(x_n)] \geq z_t \Leftrightarrow x \in \underline{C}\left(Cl_t^{\geq}\right)$$

$$f^{\leq}[g_1(x_1),g_2(x_2),\ldots,g_n(x_n)] < z_t \Leftrightarrow x \in \underline{C}\left(Cl_{t-1}^{\leq}\right)$$

where $t = 2,\ldots,m$, $C = \{g_1,g_2,\ldots,g_n\}$

Inconsistency with the dominance principle can also be represented in terms of a set of "*at least*" and "*at most*" decision rules considered together. More formally, a set of "*at least*" and "*at most*" decision rules does not contradict the classification ***Cl*** if for each $x \in Cl_t$ there exists no "*at least*" decision rule for which $x \in Cl_s^{\geq}$, with $s>t$, and there exists no "*at most*" decision rule for which $x \in Cl_s^{\leq}$, with $s<t$. A set of decision rules is complete if for each $x \in \underline{C}\left(Cl_t^{\geq}\right)$ there exists a decision rule for which $x \in Cl_s^{\geq}$, with $s \geq t$, and for each $x \in \underline{C}\left(Cl_t^{\leq}\right)$ there exists a decision rule for which $x \in Cl_s^{\leq}$, with $s \leq t$. A set of decision rules represents the classification ***Cl*** if it does not contradict ***Cl*** and it is complete.

Theorem 3 [17]. For each set of binary relations $\succeq_i$, $i = 1,\ldots,n$, being complete preorders, and for each classification ***Cl***, there exists a set of decision rules representing the classification ***Cl***.

The advantage of the DRSA with respect to competitive methodologies is the possibility of handling partially inconsistent data that are often encountered in preferential information, due to hesitation of decision makers, unstable character of their preferences, imprecise or incomplete information and the like. Therefore, we proposed a general model of conjoint measurement that, using the basic concepts of DRSA (lower and upper approximations), is able to represent these inconsistencies by a specific utility function. We showed that these inconsistencies can also be represented in a meaningful way by "*if..., then...*" decision rules induced from rough approximations.

As DRSA to multiple-criteria classification problems and the underlying decision rules exploit only the ordinal properties of the scales of criteria, they are appropriate for aggregation of ordinal criteria. This challenging problem of multiple-criteria decision making has been solved until now by using some 'max-min' aggregation functions, with the most general one – the fuzzy integral proposed by Sugeno. In [17, 42], we have shown that the decision rule model following from DRSA has advantages over the integral of Sugeno, in particular, it can represent some (even consistent) preferences that the Sugeno integral cannot.

The characterization of the decision rule preference model given in this section shows clearly its extraordinary capacity of criteria aggregation in multiple criteria classification problems. The decision rule preference model, apart from its capacity of representation, fulfils the postulate of transparency and interpretability of preference models in decision aiding. The characterization shows that the decision rule preference model is a strong alternative to functional and relational preference models to which it is formally equivalent. Recently, similar benefits of the decision rule model have been proved with respect to multiple criteria choice and ranking problems [16].

CONCLUSIONS AND PROMISING AREAS

We presented a knowledge discovery paradigm for multiple attribute and multiple criteria decision support, based on the concept of rough sets. Rough set theory provides mathematical tools for dealing with granularity of information and possible inconsistencies in the description of objects. Considering this description as an input data about a decision problem, the knowledge discovery paradigm consists in searching for patterns in the data that facilitate an understanding of the decision maker's preferences and that permit to recommend a decision concordant with these preferences. An original component of this paradigm is taking into account prior knowledge about preference semantics in patterns to be discovered.

Knowledge discovery from preference-ordered data differs from usual knowledge discovery since the former involves preference orders in domains of attributes and in the set of decision classes. This requires that a knowledge discovery method applied to preference-ordered data respects the dominance principle. As this is not the case for the well-known methods of data mining and knowledge discovery, they are not able to discover all relevant knowledge contained in the analyzed data sample and, even worse, they may yield unreasonable discoveries, because inconsistent with the dominance principle. These deficien-

cies are repaired in DRSA based on the concept of rough approximations consistent with the dominance principle. DRSA permits, moreover, to apply rough set approach to some new fields, like multiple criteria decision making and decision under uncertainty. Many extensions proposed for DRSA make of this approach
a useful tool for practical applications. Let us mention below the most important ones:

- DRSA with missing values of attributes and criteria [9],
- fuzzy set extensions of DRSA [4, 10, 19, 46],
- DRSA for hierarchical decision making [1],
- induction of association rules from preference-ordered data sets [23].

DRSA gives, moreover, a methodology for building a preference model of a decision maker in terms of decision rules. The decision rules have a special syntax involving partial evaluation profiles and dominance relation on these profiles. The clarity of the rule representation of preferences permits to see the limits of other traditional aggregation functions: utility function and outranking relation. We proposed an axiomatic characterization of these aggregation functions in terms of conjoint measurement and in terms of a set of decision rules. The axioms of the "cancellation property" type are the weakest possible. In comparison to other studies on characterization of aggregation functions, our axioms do not require any preliminary assumption about the scales of criteria. A side-result of these investigations is the corollary that the decision rule aggregation (preference model) is the most general among the known aggregation functions.

The application of DRSA to analysis of data representing a preferential information supplies, moreover, some useful elements of knowledge about the decision situation; these are: the relevance of attributes and/or criteria, information about their interaction (from quality of approximation and its analysis using fuzzy measures theory), minimal subsets of attributes or criteria (reducts) conveying the relevant knowledge contained in the exemplary decisions, the set of the non-reducible attributes or criteria (core). Moreover, DRSA permits to handle heterogeneous information: qualitative and quantitative, preference-ordered or not, crisp and fuzzy, ordinal and cardinal, partially missing and inconsistent. Finally, the proposed methodology is based on elementary concepts and mathematical tools (sets and set operations, binary relations), without recourse to any algebraic or analytical structures; the main idea is very natural and even objective, in a certain sense, like the dominance relation is.

Due to the above features, DRSA contributes in a very promising way to many different areas, like:

- **knowledge discovery and data mining**, where without DRSA the preference order in data is ignored,
- **multiple criteria decision analysis**, for which DRSA is offering a natural, general and intelligible way of modeling DM's preferences in terms of "*if..., then...*" decision rules,
- **decision under risk**, where DRSA handles non-additive probability distributions and even qualitative ordinal distributions over possible states of the world, and offers a decision rule representation of DM's preferences,
- **approximate reasoning** based on fuzzy-rough modus ponens and gradual rules induced from fuzzy rough approximations,
- **fuzzy-rough control** involving gradual rules.

The DRSA leads to a preference model of a decision maker in terms of decision rules. The decision rules have a special syntax which involves partial evaluation profiles and dominance relations on these profiles. The clarity of the rule representation of preferences enables us to see the limits of other traditional aggregation functions: the utility function and the outranking relation. In several studies [16, 17, 20, 42] we proposed an axiomatic characterization of these aggregation functions in terms of conjoint measurement theory and in terms of a set of decision rules. In comparison to other studies on the characterization of aggregation functions, our axioms do not require any preliminary assumptions about the scales of criteria. A side-result of these investigations is that the decision rule aggregation (preference model) is the most general among the known aggregation functions. The decision rule preference model fulfils, moreover, the postulate of transparency and interpretability of preference models in decision support.

An interesting research problem concerns measuring attractiveness of decision rules taking into account three application perspectives: (i) knowledge representation, (ii) prediction of new decisions and (iii) interventions based on discovered rules in some other universe (see [44]). In order to choose attractiveness measures concordant with the above perspectives we analyzed semantics of particular measures which led us to a conclusion that the best suited measures for the above applications are: (i) support and certainty, (ii) a Bayesian confirmation measure [25], and (iii) two measures related to efficiency of intervention [5], respectively. These five measures induce a partial order in the set of rules giving a starting point for an interactive browsing procedure. For building a strategy of intervention, we proposed rules discovered using the DRSA – the "at least" type rules indicate opportunities for improving assignment of objects, and the "at most" type rules indicate threats for deteriorating assignment of objects.

REFERENCES

1. Dembczyński K., Greco S., Słowiński R.: Methodology of rough-set-based classification and sorting with hierarchical structure of attributesand criteria. "Control & Cybernetics" 2002, 31, pp. 891-920.
2. Flinkman M., Michalowski W., Nilsson S., Słowiński R., Susmaga R., Wilk Sz.: Use of rough sets analysis to classify Siberian forest ecosystem according to net primary production of phytomass. "INFOR" 2000, 38, No. 3, pp. 145-161.
3. Giove S. Greco S. Matarazzo B. Słowiński R.: Variable consistency monotonic decision trees. In: Rough Sets and Current Trends in Computing. Eds. J.J. Alpigini, J.F. Peters, A. Skowron, N. Zhong. LNAI 2475, Springer-Verlag, Berlin 2002, pp. 247-254.
4. Greco, S., Inuiguchi, M., Słowiński, R.: A new proposal for fuzzy rough approximations and gradual decision rule representation. In: Transactions in Rough Sets II, LNCS 3135, Springer-Verlag, Berlin 2004, pp. 332-355.
5. Greco S., Matarazzo B., Pappalardo N., Słowiński R.: Measuring expected effects of interventions based on decision rules. "Journal of Experimental and Applied Artificial Intelligence" 2005, 17, No.1-2, pp. 103-118.
6. Greco S., Matarazzo B., Słowiński R.: A new rough set approach to evaluation of bankruptcy risk. In: Operational Tools in the Management of Financial Risk. Ed. C. Zopounidis. Kluwer, Dordrecht 1998, pp. 121-136.
7. Greco S., Matarazzo, B., Słowiński, R.: Rough approximation of a preference relation by dominance relations. "European J. Operational Research" 1999, 117, pp. 63-83.
8. Greco S., Matarazzo B., Słowiński R.: The use of rough sets and fuzzy sets in MCDM. Chapter 14. In: Advances in Multiple Criteria Decision Making. Eds. Gal T., Stewart T., Hanne T. Kluwer, Dordrecht 1999, pp. 14.1-14.59.
9. Greco S., Matarazzo B., Słowiński R.: Dealing with missing data in rough set analysis of multi-attribute and multi-criteria decision problems. In: Decision Making: Recent Developments and Worldwide Applications. Eds. Zanakis S.H., Doukidis G., Zopounidis C. Kluwer, Dordrecht 2000, pp. 295-316.
10. Greco S., Matarazzo B., Słowiński R.: Fuzzy extension of the rough set approach to multicriteria and multiattribute sorting. In: Preferences and Decisions under Incomplete Knowledge. Eds. Fodor J., De Baets B., Perny P. Physica-Verlag, Heidelberg, 2000, pp.131-151
11. Greco S., Matarazzo B., Słowiński R.: Rough sets theory for multicriteria decision analysis. "European J. Operational Research" 2001, 129, pp. 1-47.
12. Greco S., Matarazzo B., Słowiński R.: Assessment of a value of information using rough sets and fuzzy measures. In: Fuzzy Sets and their Applications. Eds. Chojcan J., Łęski J. Silesian University of Technology Press, Gliwice 2001, pp. 185-193.

13. Greco S., Matarazzo B., Słowiński R.: Rough set approach to decisions under risk. In: Rough Sets and Current Trends in Computing. Eds. Ziarko W., Yao Y. LNAI 2005, Springer-Verlag, Berlin 2001, pp. 160-169.
14. Greco S., Matarazzo B., Słowiński R.: Rough sets methodology for sorting problems in presence of multiple attributes and criteria. "European J. Operational Research" 2002, 138, pp 247-259.
15. Greco S., Matarazzo B., Słowiński R.: Multicriteria classification. In: Handbook of Data Mining and Knowledge Discovery. Eds. Kloesgen W., Zytkow J. Oxford University Press, New York 2002, chapter 16.1.9, pp. 318-328.
16. Greco S, Matarazzo B, Słowiński R: Preference representation by means of conjoint measurement & decision rule model. In: Aiding Decisions with Multiple Criteria–Essays in Honor of Bernard Roy. Eds. Bouyssou D., Jacquet-Lagrèze E., Perny P., Słowiński R., Vanderpooten D., Vincke P. Kluwer, Dordrecht, 2002, pp. 263-313
17. Greco S., Matarazzo B., Słowiński R.: Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. "European J. Operational Research" 2004, 158, pp. 271-292.
18. Greco S., Matarazzo B., Słowiński R.: Dominance-Based Rough Set Approach to Knowledge Discovery (I) – General Perspective. Chapter 20. In: Zhong N., Liu J.: Intelligent Technologies for Information Analysis. Springer-Verlag, Berlin 2004, pp. 513-552.
19. Greco S., Matarazzo B., Słowiński R.: Dominance-Based Rough Set Approach to Knowledge Discovery (II) – Extensions and Applications. Chapter 21. In: Zhong N., Liu J. Intelligent Technologies for Information Analysis. Springer-Verlag, Berlin 2004, pp. 553-612.
20. Greco, S., Matarazzo, B., Słowiński, R.: Decision rule approach. Chapter 13 [in]: J.Figueira, S.Greco and M.Ehrgott (eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer-Verlag, New York, 2005, pp. 507-562
21. Greco S., Matarazzo B., Słowiński R., Stefanowski J.: Variable consistency model of dominance-based rough set approach. In: Ziarko W., Yao Y.: Rough Sets and Current Trends in Computing. LNAI 2005, Springer-Verlag, Berlin 2001, pp. 170-181
22. Greco S., Matarazzo B., Słowiński R., Stefanowski J.: An algorithm for induction of decision rules consistent with dominance principle. In: Rough Sets & Current Trends in Computing. Eds. Ziarko W., Yao Y. LNAI 2005, Springer-Verlag, Berlin 2001, pp. 304-313.
23. Greco S., Matarazzo B., Słowiński R., Stefanowski J.: Mining association rules in preference-ordered data. In: Foundations of Intelligent Systems. Eds. Hacid M.-S., Ras Z.W., Zighed D.A., Kodratoff Y. LNAI 2366, Springer-Verlag, Berlin 2002, pp. 442-450.
24. Greco S., Matarazzo B., Słowiński R., Tsoukias A.: Exploitation of a rough approximation of the outranking relation in multicriteria choice and ranking. [In]: Trends in Multicriteria Decision Making. Eds. Stewart T.J., van den Honert R.C. LNEMS 465, Springer-Verlag, Berlin 1998, pp. 45-60.

25. Greco S., Pawlak Z., Słowiński R.: Can Bayesian confirmation measures be useful for rough set decision rules? "Engineering Applications of Artificial Intelligence" 2004, 17, No 4, pp. 345-361.
26. Grzymala-Busse J.W.: A new version of the rule induction system LERS. "Fundamenta Informaticae" 1997, 31, pp. 27-39.
27. Michalowski W., Rubin S., Słowiński R., Wilk Sz.: Mobile clinical support system for pediatric emergencies. "Journal of Decision Support Systems" 2003, 36, pp. 161-176.
28. Michalski R.S., Bratko I. Kubat M.: Machine Learning and Data Mining – Methods and Applications. Eds. Wiley, New York 1998.
29. Pawlak Z.: Rough sets. Int. J. Information & Computer Sciences, 1982, 11, pp. 341-356.
30. Pawlak Z.: Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht 1991.
31. Pawlak Z., Grzymala-Busse J.W., Słowiński R., Ziarko W.: Rough sets. „Communications of the ACM" 1995, 38, pp. 89-95.
32. Pawlak Z., Słowiński K., Słowiński R.: Rough classification of patients after highly selective vagotomy for duodenal ulcer. "International Journal of Man-Machine Studies" 1986, 24, pp. 413-433.
33. Pawlak Z., Słowiński R.: Rough set approach to multi-attribute decision analysis. "European J. Operational Research" 1994, 72, pp. 443-459.
34. Polkowski L.: Rough Sets: Mathematical Foundations. Physica-Verlag, Heidelberg 2002.
35. Polkowski L., Skowron A.: Calculi of granules based on rough set theory: approximate distributed synthesis and granular semantics for computing with words. In: New Directions in Rough sets, Data Mining and Soft-Granular Computing. Eds. Zhong N., Skowron A., Ohsuga S. LNAI 1711, Springer-Verlag, Berlin 1999, pp. 20-28.
36. Rossi L., Słowiński R., Susmaga R.: Rough set approach to evaluation of stormwater pollution. "International Journal of Environment and Pollution" 1999, 12, No 2/3, pp. 232-250.
37. Roy B.: Multicriteria Methodology for Decision Aiding. Kluwer, Dordrecht 1996.
38. Słowiński R.: Rough set learning of preferential attitude in multi-criteria decision making. In: Methodologies for Intelligent Systems. Eds. Komorowski J., Ras Z.W. LNAI 689, Springer-Verlag, Berlin 1993, pp. 642-651.
39. Słowiński R.: Rough set approach to decision analysis. "AI Expert Magazine" 1995, 10, No 3, pp. 18-25.
40. Słowiński R., Greco S., Matarazzo B.: Rough set analysis of preference-ordered data. In: Rough Sets and Current Trends in Computing. Eds. Alpigini J.J., Peters J.F., Skowron A., Zhong N. LNAI 2475, Springer-Verlag, Berlin 2002, pp. 44-59.
41. Słowiński R., Greco S., Matarazzo B.: Mining decision-rule preference model from rough approximation of preference relation. In: Proc. 26th IEEE Annual Int. Conf. on Computer Software & Applications. Oxford 2002, pp. 1129-1134.

42. Słowiński R., Greco S., Matarazzo B.: Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle. "Control and Cybernetics" 2002, 31, pp. 1005-1035.
43. Słowiński R., Greco S.: Inducing robust decision rules from rough approximations of a preference relation. In: Artificial Intelligence and Soft Computing. Eds. Rutkowski L., Siekmann J., Tadeusiewicz R., Zadeh L.A. LNAI 3070, Springer-Verlag, Berlin 2004, pp. 118-132.
44. Słowiński R., Greco S.: Measuring attractiveness of rules from the viewpoint of knowledge representation, prediction and efficiency of intervention. Keynote lecture. In: Proc. 3rd Atlantic Web Intelligence Conference, AWIC'05, Łódź, June 6-9, 2005.
45. Słowiński R., Greco S., Matarazzo B.: Rough set based decision support. Chapter 16. In: Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques. Eds. Burke E.K., Kendall G. Springer-Verlag, New York 2005, pp. 475-527.
46. Słowiński R., Stefanowski J.: Rough set reasoning about uncertain data. "Fundamenta Informaticae" 1996, 27, pp. 229-243.
47. Słowiński R., Vanderpooten D.: A generalised definition of rough approximations. "IEEE Transactions on Data and Knowledge Engineering" 2000, 12, pp. 331-336.
48. Słowiński R., Zopounidis C.: Application of the rough set approach to evaluation of bankruptcy risk. "Intelligent Systems in Accounting, Finance & Management" 1995, 4, pp. 27-41.
49. Stefanowski J., Vanderpooten D.: Induction of decision rules in classification and discovery-oriented perspectives. "International J. Intelligent Systems" 2001, 16, No 1, pp. 13-28.
50. Wilk Sz., Słowiński R., Michalowski W., Greco S.: Supporting triage of children with abdominal pain in the emergency room. "European J. Operational Research" 2005, 160, No 3, pp. 696-709.

Kazimierz Zaraś

Hamdjatou Kane

Maciej Nowak

PRODUCTION PLANNING AND CONTROL: AN APPROACH BASED ON ROUGH SETS

Abstract

This paper deals with the problem of production process control in a job shop where the work flow is controlled by Kanban cards. Production may proceed differently depending on the lot size, number of Kanban cards used, and the decision rule for choosing the job waiting to be processed. The problem that arises consists in deciding which rule should be used, how many Kanbans should be allocated for each operation, and what lot size should be applied. Thus, the choice of the best triplet constitutes a multicriteria problem. We propose to solve the multicriteria problem by using Rough Set Approach. Taking into account operator's choices we use the dominance-based rough set approach to induce the decision rules, which can be applied to choose the best triplet from a large number of alternatives. This paper deals with the problem of production process control in a job shop where the work flow is controlled by Kaban cards. Production may proceed differently according to a lot size, number of Kanban cards used, and the decision rule for choosing the waiting job to process. The problem that arises consists in deciding which rule should be used, how many Kanbans should be allocated for each operation, and what lot size should be applied. Thus, the choice of the best triplet constitutes a multicriteria problem. We propose to solve the multicriteria problem by using Rough Set Approach. Based on the choice of the operator and using the dominance-based rough set approach we will be able to induce the decision rules, which can be applied to choose the best triplet from a large number of alternatives.

Keywords

Job shop, Kanban, multicriteria analysis, rough sets, stochastic dominance.

INTRODUCTION

This study assumes that Just-in-Time (JIT) approach is used for scheduling the production system. The work flow is controlled by Kanban cards. This technique is mainly used in a classic mass production environment with few product variations and levelled demand. Gravel and Price [4; 3] have shown how

this approach can be adapted to job-shop environment. A more recent variation of this problem is known as POLCA (Paired, Over-lapping, Loops of Cards with Authorisation); it is applicable in a job-shop environment where each job can be unique [8; 9].

Production may proceed differently depending on the lot size, number of Kanban cards used, and the decision rule for choosing the job waiting to be processed. The problem that arises consists in deciding which rule should be used, how many Kanbans should be allocated for each operation, and what lot size should be applied. In general, smaller lot sizes reduce work-in-progress, but also increase the number of machine set-ups. Increasing the number of allocated Kanbans improves machine utilisation, but may also increase average work-in-progress level. Finally, the performance of a scheduling rule depends on the performance measure used. Thus, the choice of the best triplet involving the Kanban lot size, the decision rule, and the number of Kanbans constitutes a multicriteria problem. Gravel et al. [5] considered a similar problem and used Electre method [13] to model outranking relations. They assumed that completion time of each operation is known and simulated each product separately to evaluate performance of the shop under various conditions (various products, various production environments). In their study, they assumed that the decision maker (DM) is risk-averse. Nowak et al. [11] proposed a modified approach for this problem. They assumed that the DM is risk-prone and several products are processed simultaneously in the shop. The probability distribution of the operation's completion times was determined by series of simulations for each decision alternative to analyse the performance of a shop. This paper deals with solving the problem of production process control as a multicriteria problem such as in [2 ;11] but by using the Rough Set approach. By application of the Rough Set approach we don't need the explicit information about criterion weights as it is necessary to have for preference modelling with the ELECTRE method. In practice we know that criterion weights determination is not the easy task. In the Rough Set approach the DM shows us how he does his job by ordering the alternatives from the efficient set; implicit weights are given by ranking the alternatives.

This paper is structured as follows: the problem is formulated as a multicriteria problem in Section 1. Section 2 presents the rough set approach to choose the control production parameters. In Section 3, we give a job shop production example.

1. PRODUCTION PROCESS CONTROL AS A MULTICRITERIA PROBLEM

In this paper, three performance criteria are considered: makespan; average work-in-progress level; number of set-ups. The first criterion is very important, since short execution times increase effective capacity of the shop and improve the service level. The average work-in-progress level reflects the effectiveness of the firm in reducing investment in semi-finished work. Finally, the number of machine set-ups indicates the number of times the operators have to adapt to a different operation. All tree criteria will be minimized.

The set of alternatives includes all triplets (the lot size, the number of kanban cards and the decision rule). The set of attributes includes all criteria (makespan, average stock and the number of set-ups). Performances of each alternative with respect to the attributes are evaluated by distribution functions. The knowledge base used for the construction of these functions was obtained by using a simulation model of the process where several products are manufactured simultaneously such as in Nowak [12].

The decision situation considered here may be conceived as a problem (**A**, **X**, **E**) where **A** is a finite set of alternatives (triplets), $i = 1, 2, ..., m$; **X** is a finite set of attributes (criteria) X_k, $k = 1, 2, ..., n$; and **E** is a set of evaluations of triplets with respect to the criteria:

$$\mathbf{E} = \begin{bmatrix} X_{11} & \cdots & X_{1k} & \cdots & X_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ X_{i1} & \cdots & X_{ik} & \cdots & X_{in} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ X_{m1} & \cdots & X_{mk} & \cdots & X_{mn} \end{bmatrix}$$

We assume that the attributes are probabilistically independent and also satisfy the independence conditions which allows us to use additive utility function.

Our approach consists in building global preferences on the set of parameter triplets by first comparing their distributional evaluations in relation to each criterion to model the partial preferences and then by aggregating them into global preferences. With respect to each criterion the preferences are modelled by using the Stochastic Dominances [18; 17; 10]. The comparison of alternatives can be conducted by means of First Degree Stochastic Dominance (FSD), Second Degree Inverse Stochastic Dominance (SISD) [1] and Third Degree Inverse Stochastic Dominance (TISD1 and TISD2). The FSD is defined, if the difference between two cumulated distributions is non-positive for all x, and for at least one

x this difference is strictly negative. The Second Inverse Stochastic Dominance (SISD) is defined, if the difference between two integrals from right to left on two cumulated distributions is non-positive for all x, and for at least one x this difference is strictly negative. The Third Inverse Stochastic Dominance (TISD2) is defined, if the difference between two double integrals from right to left on two cumulated distributions is non-positive for all x, and for at least one x this difference is strictly negative. Generally, if one of the inverse stochastic dominances is verified, it has been proven for increasing convex class of utility functions that the expected utility of distributional evaluation which dominates is greater or equal to the expected utility of distributional evaluation which is dominated.

Let F_{ik} and F_{jk} be cumulative distribution functions:

(1) F_{ik} FSD F_{jk} if and only if $F_{ik} \neq F_{jk}$ and $F_{ik} \leq F_{jk}$ for all $x \in [c, d]$.

(2) F_{ik} SISD F_{jk} if and only if $F_{ik} \neq F_{jk}$ and $\int_{x}^{d} F_{ik}(x_k)\,dx_k \leq \int_{x}^{d} F_{jk}(x_k)\,dx_k$

for all $x \in [c, d]$.

(3) F_{ik} TISD2 F_{jk} if and only if $F_{ik} \neq F_{jk}$ and $\int_{x}^{d}\int_{x}^{d} F_{ik}(y_k)\,dx_k\,dy_k \leq \int_{x}^{d}\int_{x}^{d} F_{jk}(y_k)\,dx_k dy_k$

for all x $\in [c, d]$.

(4) F_{ik} TISD1 F_{jk} if and only if $F_{ik} \neq F_{jk}$ and $\int_{c}^{x}\int_{x}^{d} F_{ik}(y_k)\,dx_k dy_k \leq \int_{c}^{x}\int_{x}^{d} F_{jk}(y_k)\,dx_k dy_k$

for all x $\in [c, d]$.

where $[c, d]$ is the interval of definition of two random variables X_{ik}, X_{jk}.

2. THE ROUGH SET APPROACH

The rough set approach is based on the Rough Set theory developed by Pawlak (1991), Pawlak and Slowinski (1994) and Greco, Matarazzo and Slowinski (1999). This theory was proposed in this paper for ranking a large number of parameter triplets from the efficient set. The Rough Set Theory relies on a tabular representation of the preferential information expressed by the DM.

These preferences are expressed using the following procedure. First, a small number (4-7) of parameter triplets chosen from different parts of effi-

cient set are presented , and the DM is asked to order triplets from the most preferred to the least preferred.

Second, this ranking represents the DM's preferences, which are noted in a decision table with respect to the decisional attributes.

Let B be a finite subset of parameter triplets which are considered by the DM as the basis for exemplary pairwise comparisons. In addition, let C be the set of attributes (condition attributes) describing the parameter triplet, and D, the decision attribute. The decision table is defined as the 4-tuple: $T = (H, C \cup D, V_C \cup V_D, g)$ where $H \subseteq B \times B$ is a finite set of pairs of parameter triplets, $C \cup D$ is the union of two subsets of attributes, called condition and decision attributes, $V_C \cup V_D$ is the union of the domains of these attributes respectively, and g: $H \times (C \cup D) \rightarrow V_C \cup D$ is a total function where $V_C = \cup V_k$.

This function is such that:

(1) $g[(a_i, a_j), k] = 1$, if f_{ik} SD_k f_{jk} is verified $\forall X_k \in C$, and $\forall (a_i, a_j) \in H$;

(2) $g[(a_i, a_j), k] = 0$, if f_{ik} not SD_k f_{jk} is verified $\forall X_k \in C$, and $\forall (a_i, a_j) \in H$; and

$g[(a_i, a_j), k] \in V_k$, $\forall X_k \in C$, and $\forall (a_i, a_j) \in H$ and $g[(a_i, a_j), D] \in V_D$, $\forall (a_i, a_j) \in H$.

In our decision table $g[(a_i, a_j), D]$ can also have two values on $H \subseteq B \times B$:

(1) $g[(a_i, a_j), D] = P$, if a_i is preferred to a_j $\forall (a_i, a_j) \in H_P$,

(2) $g[(a_i, a_j), D] = N$, if a_i is not preferred to a_j $\forall (a_i, a_j) \in H_N$.

These two values will be expressed with respect to the decisional attribute. The subset H_P expresses the preferences and H_N expresses non-preferences.

In general, the decision table can be presented as in Table 1.

Table 1

Decision table

		X1	X2	...	Xm	D
HP	(a_i, a_j)	$g[(a_i, a_j), 1]$	$g[(a_i, a_j), 2]$	...	$g[(a_i, a_j), m]$	$g[(a_i, a_j), D] = P$
	.	.	.	...	.	.
	.	.	.	...	.	.
HN	.	.	.	...	.	.
	.	.	.	...	.	.
	(a_s, a_t)	$g[(a_s, a_t), 1]$	$g[(a_s, a_t), 2]$	...	$g[(a_s, a_t), m]$	$g[(a_i, a_j), D] = N$

In the decision table, with respect to each conditional attribute, the pairwise evaluation of each ranked parameter triplet provides decision rules. In our approach, we suggest the approximation of the global preference relation P by the Multiattribute Stochastic Dominance for reduced number of attribute MSD_R. This dominance can be defined as follows:

Definition 1 [15; 14]

a_i MSD_R a_j if and only if f_{ik} SD_k f_{jk} for all $X_k \in R \subseteq X$ (1) (1)

The MSD_R is the particular case of the MSD dominance defined for given a_i, a_j $\in$ A by [6] as follows:

Definition 2 [6]

a_i MSD a_j if and only if f_{ik} SD_k f_{jk} for all $X_k \in X$ (2) (2)

In the Rough set theory, the approximation of the global preference relation P by MSD_R can be done by lower and upper approximations. According to Greco et al. [5], the lower approximation can be defined as follows:

$$Q_*(P) = \bigcup_{R \subseteq X} \{(MSD_R \cap H) \subseteq P\} \tag{3}$$

The application of the lower approximation allows us to induce the following kind of decision rules:

Rule: If a_j MSD_{R_3} a_j then a_j P a_j

The upper approximation (4) may contain the Multiattribute Stochastic Dominances for reduced number of attributes which leads to the conclusion for preference or non preference. These dominances usually introduce uncertainty in the induction of the decision rules and are referred to as the boundary region (5) which added to lower approximation give us an upper approximation of the preferences. According to Greco et al. [5], the upper approximation can be defined as follows:

$$Q^*(P) = \bigcap_{R \subseteq X} \{(MSD_R \cap H) \supseteq P\} \tag{4}$$

$$BN_Q(P) = Q^*(P) - Q_*(P) \tag{5}$$

The decision rules from upper approximation of the preference P are formulated as follow:

Rule: If a_i MSD_{R_2} a_j then a_i P a_j or a_i N a_j

These uncertain rules obtained from upper approximation must be discarded to eliminate inconsistencies. The decision rules obtained from the lower approximation of the global preference P are kept and they are used to rank all parameter triplets belonging to the efficient set.

All parameter triplets in the efficient set are compared two by two to determine if they satisfy a decision rule. If the comparison of two parameter triplets with the procedure used to define the set of rules leads to a decision rule, the score associated with the first parameter triplet is incremented by one and the score for the second parameter triplet is decremented by one. Following the comparisons of all parameter triplets in the efficient set, all parameter triplets are ranked in decreasing order of score.

3. APPLICATION

We consider a company which produces sport equipment. The firm manufacturers 12 different products which are processed simultaneously in the shop. The production process of each product includes a number of operations performed on different machines (see Table 2). The number and type of operations are different for each product. Parts may return to the same machining centre in the process. 24 devices are installed in the work centre: 6 machines of type M_1, 6 machines of M_2 type, 4 machines of M_3 type, 4 machines of M_4 type, 2 machines of M_5 type and 2 machines of M_6 type.

The production planning and control are organized according to the "Just-in-Time" rules. Production orders are broken into small Kanban lots treated individually. The firm uses Kanban cards to control the work flow. Each operation has its Kanban. One or more Kanbans may be used for each operation. Before starting his work, an operator has to choose one of the waiting operations. Scheduling rules are often used to determine the order in which operations should be processed on workstations. Thus, the worker is able to decide which job queuing at the station needs to be processed first. In our study, eight decision rules are considered: (1) The first come – first served (FCFS) rule; (2) The shortest processing time (SPT) rule; (3) The same job as previously (SJP) rule; (4) The shortest next queue (SNQ) rule; (5) The minimal total time of the rest operations on the path (MTP) rule, (6) The maximal number of Kanbans awaiting processing at the workstation (MKW) rule; (7) The maximal total number of Kanbans awaiting processing at all workstations on the path (MKP) rule; (8) The priority ratio (PR) rule.

Table 2

he number of operations and units of products

Products	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{11}	W_{12}
Number of operations	21	20	22	20	23	21	23	19	25	21	14	40
Number of units	120	120	120	120	120	120	60	60	60	60	60	60

The FCFS rule gives priority to a job that has been queuing at the station for the longest time. The SPT rule chooses the job with shortest planned processing time. The SJP rule assumes that the job which is the same as previously processed on the station should be chosen. The SNQ rule gives priority to the job for which the queue at the next station is the shortest. The MPT rule chooses the operation for which the total time for the remaining operations that have to be completed is minimal. The MKW rule assumes that the job with the greatest number of Kanbans waiting at the workstation should be performed first. The MKP rule is similar, but it considers all operations that have to be performed to complete the processing. The rule selects the operation, for which the total number of Kanbans for all operations on the path is maximal. Finally PR rule is based on the following ratio (6):

$$z_i = \begin{cases} 0 & \text{for } i \in K \\ l_i - \max\limits_{k \in E_i}\{l_k\} & \text{for } l_i > \max\limits_{k \in E_i}\{l_k\} \text{ for } i \in A \\ L & \text{otherwise} \end{cases} \tag{6}$$

Where:
A: the set of operations waiting for processing at the workstation i,
K: the set of final operations,
l_i: the number of unavailable Kanbans for operation i,
E_i: the set of operations which produce components, that are used together with the component produced by operation i in the next stage of the process.

Four values of lot size are considered: 5, 10, 15, and 20, while the number of Kanbans is assumed to be between 2 and 5 and eight scheduling rules. Thus, 128 triplets of parameters are considered. Three criteria are used for evaluating performance of the alternatives: makespan (measured in seconds); average work-in-progress level (measured by the average number of jobs queuing at stations); number of set-ups (the whole number of set-ups done on all stations).

The solution to the problem is as follows: (a) simulation of the production of selected products for each triplet of parameters; (b) construction of distribution functions for each triplet with respect to each attribute; (c) identification of stochastic dominances between triplets of parameters in relation to each attribute; (d) ranking of parameter triplets according to decision rules.

First, a series of one hundred simulations had been done for each triplet to build distributional evaluations with respect to each criterion. In our case, the set of alternatives includes 128 triplets but a certain number of triplets were rejected because of the constraints involved by the DM such that makespan cannot be longer than 85 hours by week; the average number of jobs waiting for all operations no more than 4600 by week; and the number of set-ups for all machines no more than 4000 by week. It was also assumed that the probability of reaching unsatisfactory attribute value should not exceed 0,05. The result of this verification was that 71 triplets were rejected. Next we started to identify types of stochastic dominance between alternatives with respect to attributes. According to prospect theory [7], we assumed that the decision-maker is risk-prone and so we used FSD, SISD, TISD1 and TISD2 (as defined in section 2) to explain relations between alternatives. Tables 3, 4 and 5 show the relations between selected alternative pairs explained by stochastic dominance with respect to the attribute X_1 (makespan). These dominances can be used to determine the multi-attribute stochastic dominance (MSD). By verification of the multi-attribute dominance rule (see definition 1, in Section 3) on the remaining subset of 57 alternatives, we obtained 44 efficient triplets as shown in Table 6.

Table 3

Stochastic dominance for attribute X_1 (makespan)

X_1	5_3_2	5_3_3	5_4_4	10_3_3	10_3_4	10_4_2	10_4_3	10_6_2	10_6_3
5_3_2	X	FSD	SISD	FSD	FSD	SISD	FSD	SISD	TISD1
5_3_3	X	X	X	FSD	FSD	X	SISD	X	X
5_4_4	X	TISD1	X	FSD	FSD	FSD	FSD	X	X
10_3_3	X	X	X	X	TISD1	X	X	X	X
10_3_4	X	X	X	X	X	X	X	X	X
10_4_2	X	TISD1	X	FSD	FSD	X	FSD	X	X
10_4_3	X	X	X	FSD	FSD	X	X	X	X
10_6_2	X	FSD	SISD	FSD	FSD	SISD	FSD	X	X
10_6_3	X	FSD	SISD	FSD	FSD	FSD	FSD	TSD	X

Table 4

Stochastic dominance for attribute X_2 (number of set-ups)

X_2	5_3_2	5_3_3	5_4_4	10_3_3	10_3_4	10_4_2	10_4_3	10_6_2	10_6_3
5_3_2	X	X	FSD	X	X	X	X	X	X
5_3_3	FSD	X	FSD	X	X	X	X	FSD	FSD
5_4_4	X	X	X	X	X	X	X	X	X
10_3_3	FSD	FSD	FSD	X	X	FSD	FSD	FSD	FSD
10_3_4	FSD	FSD	FSD	FSD	X	FSD	FSD	FSD	FSD
10_4_2	FSD	TISD1	FSD	X	X	X	X	FSD	FSD
10_4_3	FSD	FSD	FSD	X	X	FSD	X	FSD	FSD
10_6_2	FSD	X	FSD	X	X	X	X	X	X
10_6_3	FSD	X	FSD	X	X	X	X	TISD1	X

Table 5

Stochastic dominance for attribute X_3 (average stock)

X3	5_3_2	5_3_3	5_4_4	10_3_3	10_3_4	10_4_2	10_4_3	10_6_2	10_6_3
5_3_2	X	FSD	FSD	FSD	FSD	FSD	FSD	FSD	FSD
5_3_3	X	X	FSD	FSD	FSD	FSD	FSD	X	FSD
5_4_4	X	X	X	FSD	FSD	FSD	FSD	X	X
10_3_3	X	X	X	X	FSD	X	FSD	X	X
10_3_4	X	X	X	X	X	X	X	X	X
10_4_2	X	X	X	FSD	FSD	X	FSD	X	X
10_4_3	X	X	X	X	SSD	X	X	X	X
10_6_2	X	FSD	FSD	FSD	FSD	FSD	FSD	X	FSD
10_6_3	X	X	FSD	FSD	FSD	FSD	FSD	X	X

Table 6

Alternatives analyzed in the last step of the procedure

Alternative	Lot-size	Scheduling rule	No. of Kanbans	Alternative	Lot-size	Scheduling rule	No. of Kanbans
1	5	SJP	2	23	15	MKW	4
2	5	SJP	3	24	15	MKW	5
3	5	SNQ	4	25	15	MKP	2
4	10	SJP	3	26	15	MKP	3
5	10	SJP	4	27	15	MKP	4
6	10	SNQ	2	28	15	MKP	5
7	10	SNQ	3	29	15	PR	4
8	10	MKW	2	30	15	PR	5
9	10	MKW	3	31	20	SPT	2
10	10	MKW	4	32	20	SPT	5
11	10	MKW	5	33	20	SNQ	2
12	10	MKP	2	34	20	MKW	2
13	10	MKP	3	35	20	MKW	3
14	10	MKP	4	36	20	MKW	4
15	10	MKP	5	37	20	MKW	5
16	10	PR	5	38	20	MKP	2
17	15	SPT	2	39	20	MKP	3
18	15	SJP	2	40	20	MKP	4
19	15	SJP	3	41	20	MKP	5
20	15	SNQ	2	42	20	PR	3
21	15	MKW	2	43	20	PR	4
22	15	MKW	3	44	20	PR	5

Next, we attempt to build a decision table (Table 7) for pairwise comparison between 5 triplets chosen to make an exercise with industrial operator. The preferences in the decision table were supposed finally to be the same as those analyzed by ELECTRE method in Nowak et al. [11] case. The evaluations with respect to the decisional attribute partition the set of pairs of triplets H into: those which express preferences and those which express non preferences. With respect to the conditional attributes, for each pair of triplets, we can identify the Multiattribute Dominances for reduced number of attributes (MSD_R).

This table shows for the first three pairs of triplets, two attribute dominances with respect to X_2 and X_3.

Table 7

Decision table

0	1	1	P
0	1	1	P
0	1	1	P
1	0	1	P
0	1	0	P
0	1	0	P
1	0	1	P
1	0	0	P
1	0	1	P
1	0	1	P
0	0	0	N
1	0	0	N
1	0	0	N
0	1	0	N
1	0	0	N
1	0	0	N
0	1	0	N
0	1	1	N
0	1	0	N
0	1	0	N

In our approach, we suggest the approximation of the global preference relation P by the Multiattribute Stochastic Dominance to reduce the number of attribute MSD_R. The application of the lower approximation (3) for twenty examples (see Table 5) of the pairwise comparison of the triplets using the software package 4eMKA2 allows us to induce the first decision rule which is based on the two attribute dominances with respect to the attributes X_1 (makespan) and X_3 (number of set-ups).

Rule 1: If a_i $MSD_{x1,x3}$ a_j then a_i P a_j

By application of the upper approximation (4) of preferences we can identify the boundary region, which contains 15 pairs of triplets out of 20 in the decision table (Table 5). Larger boundary region implies the weaker quality of approximation. This is why the quality of approximation of preference is equal to only 0.21.

$$\gamma_Q(P) = \frac{|Q_*(P)|}{|Q^*(P)|} = \frac{4}{19} = 0.21 \tag{7}$$

Rules 2 and 3 are induced from upper approximation of preferences.

Rule 2: If a_i MSD_{X_2} a_j then a_i P a_j or a_i N a_j

Rule 3: If a_i MSD_{X_1} a_j then a_i P a_j or a_i N a_j

Finally, we keep the first certain decision rule to model the overall binary preference relation. The last step of the suggested methodology is to apply this decision rule to order the entire set of forty-four triplets. The extraction of the list of pairs of triplets supporting the decision rule of 44 triplets is presented in Table 6. For each triplet a_i we have:

$SC_{++}(a_i) = card(\{ a_j \in A$: there is at least one D_{++} decision rule stating that $a_i P a_j \})$,

$SC_{+-}(a_i) = card(\{ a_j \in A$: there is at least one D_{++} decision rule stating that $a_j P a_i \})$,

If we identify the pairs of triplets with the decision rule which it corresponds to, we have one of two following situations for each triplet. The triplet a_i dominates the others or is dominated by them.

To each triplet a_i, we assign a score $NFS(a_i)$ called the *net flow score* [5] where:

$$NFS(a_i) = S_{++}(a_i) - S_{+-}(a_i).$$

In the ranking problem, the final recommendation is the total pre-order established by $SNF(a_i)$ on the set of triplets shown in Table 9. By comparing two rankings obtained by the ELECTRE method and the Rough Set method, we find them very similar because the preferences were supposed to be the same in the decision table. Usually, the preferences are given by the DM while ranking small number of alternatives from the efficient set (Table 10).

Table 8

Pairs of alternatives supporting decision Rule 1

(5_3_2 ; 5_3_3)	(5_3_2 ; 5_4_4)	(5_3_2 ; 10_3_3)	(5_3_2 ; 10_3_4)	(5_3_2 ; 10_4_2)	(5_3_2 ; 10_4_3)
(5_3_2 ; 15_6_4)	(5_3_2 ; 15_6_5)	(5_3_2 ; 15_7_2)	(5_3_2 ; 15_7_3)	(5_3_2 ; 15_7_4)	(5_3_2 ; 15_7_5)
(5_3_2 ; 20_7_5)	(5_3_3 ; 10_3_3)	(5_3_3 ; 10_3_4)	(5_3_3 ; 10_4_3)	(5_3_3 ; 15_2_2)	(5_3_3 ; 15_3_2)
(5_3_3 ; 20_6_3)	(5_3_3 ; 20_6_4)	(5_3_3 ; 20_6_5)	(5_4_4 ; 10_3_3)	(5_4_4 ; 10_3_4)	(5_4_4 ; 10_4_2)
(5_4_4 ; 20_6_4)	(5_4_4 ; 20_6_5)	(10_3_3 ; 10_3_4)	(10_3_3 ; 15_3_2)	(10_3_3 ; 15_3_3)	(10_3_3 ; 20_2_5)
(10_4_2 ; 20_4_4)	(10_4_2 ; 20_6_2)	(10_4_2 ; 20_6_3)	(10_4_2 ; 20_6_4)	(10_4_2 ; 20_6_5)	(10_4_3 ; 10_3_4)
(10_6_2 ; 15_2_2)	(10_6_2 ; 15_3_2)	(10_6_2 ; 15_3_3)	(10_6_2 ; 15_4_2)	(10_6_2 ; 15_6_2)	(10_6_2 ; 15_6_3)
(10_6_2 ; 20_7_4)	(10_6_2 ; 20_7_5)	(10_6_3 ; 5_4_4)	(10_6_3 ; 10_3_3)	(10_6_3 ; 10_3_4)	(10_6_3 ; 10_4_2)
(10_6_3 ; 20_2_5)	(10_6_3 ; 20_4_2)	(10_6_3 ; 20_6_2)	(10_6_3 ; 20_6_3)	(10_6_3 ; 20_6_4)	(10_6_3 ; 20_6_5)
(10_6_4 ; 15_6_2)	(10_6_4 ; 15_6_3)	(10_6_4 ; 15_6_4)	(10_6_4 ; 15_6_5)	(10_6_4 ; 20_2_2)	(10_6_4 ; 20_2_5)
(10_6_5 ; 15_2_2)	(10_6_5 ; 15_3_2)	(10_6_5 ; 15_3_3)	(10_6_5 ; 15_4_2)	(10_6_5 ; 15_6_2)	(10_6_5 ; 15_6_3)
(10_7_2 ; 10_3_3)	(10_7_2 ; 10_3_4)	(10_7_2 ; 10_4_3)	(10_7_2 ; 10_7_5)	(10_7_2 ; 15_2_2)	(10_7_2 ; 15_3_2)
(10_7_2 ; 20_2_2)	(10_7_2 ; 20_2_5)	(10_7_2 ; 20_4_2)	(10_7_2 ; 20_6_2)	(10_7_2 ; 20_6_3)	(10_7_2 ; 20_6_4)
(10_7_3 ; 15_2_2)	(10_7_3 ; 15_3_2)	(10_7_3 ; 15_3_3)	(10_7_3 ; 15_4_2)	(10_7_3 ; 15_6_2)	(10_7_3 ; 15_6_3)
(10_7_3 ; 20_6_3)	(10_7_3 ; 20_6_4)	(10_7_3 ; 20_6_5)	(10_7_3 ; 20_7_2)	(10_7_3 ; 20_7_3)	(10_7_3 ; 20_7_4)

Table 9

Ranking of triplets according to the Rough Set approach

Triplet	S_{++}	S_{+-}	NFS	Rang	Triplet	S_{++}	S_{+-}	NFS	Rang
5_3_2	33	0	33	1	20_8_3	0	0	0	23
10_7_2	28	0	28	2	15_8_5	0	0	0	24
10_7_4	28	0	28	3	15_8_4	0	0	0	25
10_7_3	28	0	28	4	10_8_5	0	0	0	26
10_7_5	27	3	24	5	15_6_2	16	18	-2	27
10_6_2	24	1	23	6	15_6_3	15	19	-4	28
15_7_3	25	5	20	7	15_6_4	13	20	-7	29
10_6_4	21	1	20	8	15_6_5	12	21	-9	30
15_7_2	24	5	19	9	15_2_2	11	22	-11	31
10_6_3	20	1	19	10	10_4_3	5	16	-11	32
10_6_5	20	2	18	11	20_6_2	9	24	-15	33
15_7_5	23	6	17	12	20_6_4	7	26	-19	34
15_7_4	23	7	16	13	20_2_2	4	25	-21	35
5_4_4	15	5	10	14	20_6_5	6	27	-21	36
20_7_2	18	9	9	15	20_6_3	5	26	-21	37
5_3_3	18	10	8	16	15_4_2	3	25	-22	38
20_7_5	17	10	7	17	15_3_2	3	28	-25	39
10_4_2	14	7	7	18	10_3_3	4	30	-26	40
20_7_4	17	11	6	19	20_4_2	1	30	-29	41
20_7_3	17	11	6	20	10_3_4	1	33	-32	42
20_8_4	0	0	0	21	15_3_3	0	34	-34	43
20_8_5	0	0	0	22	20_2_5	0	35	-35	44

Table 10

Results obtained from Rough Set and Electre methods

Rough Set approach				Electre method			
Rank	Triplet	Rank	Triplet	Rank	Triplet	Rank	Triplet
1	5_3_2	23	20_8_3	1	10_7_2	23	5_3_3
2	10_7_2	24	15_8_5	2	10_7_3	24	10_4_2
3	10_7_4	25	15_8_4	3	5_3_2	25	10_8_5
4	10_7_3	26	10_8_5	4	10_7_4	26	15_6_2
5	10_7_5	27	15_6_2	5	15_7_2	27	15_6_3
6	10_6_2	28	15_6_3	6	15_7_3	28	15_2_2
7	15_7_3	29	15_6_4	7	10_7_5	29	15_6_4
8	10_6_4	30	15_6_5	8	15_7_5	30	15_6_5
9	15_7_2	31	15_2_2	9	15_7_4	31	20_6_2
10	10_6_3	32	10_4_3	10	20_8_4	32	5_4_4
11	10_6_5	33	20_6_2	11	20_8_5	33	20_2_2
12	15_7_5	34	20_6_4	12	10_6_2	34	20_6_4
13	15_7_4	35	20_2_2	13	20_8_3	35	20_6_5
14	5_4_4	36	20_6_5	14	10_6_4	36	20_6_3
15	20_7_2	37	20_6_3	15	20_7_2	37	10_3_3
16	5_3_3	38	15_4_2	16	10_6_3	38	10_4_3
17	20_7_5	39	15_3_2	17	10_6_5	39	15_3_2
18	10_4_2	40	10_3_3	18	20_7_3	40	15_4_2
19	20_7_4	41	20_4_2	19	15_8_5	41	15_3_3
20	20_7_3	42	10_3_4	20	20_7_4	42	10_3_4
21	20_8_4	43	15_3_3	21	20_7_5	43	20_4_2
22	20_8_5	44	20_2_5	22	15_8_4	44	20_2_5

CONCLUSION

The given procedure constitutes a dynamic decision aid in the production process control. By changing parameters and attributes, we can adapt it to other production environments.

We have used a Rough Set approach to choose the best triplet (Kanban lot size, the decision rule and the number of Kanbans). The set of decision rules induced by application of Rough Set techniques represents the preference model of the DM and can be used to order very large efficient set of triplets.

REFERENCES

1. Goovaerts M.J., de Vylder F., Haezendonck J.: Insurance premiums: Theory and Applications. North-Holland, Amsterdam 1984.
2. Gravel, M., Martel, J.M., Nadeau, R., Price, W. and Tremblay, R. A multicriterion view of optimal resource allocation in job-shop production. "European Journal of Operational Research" 1992, 61, pp. 230-244.
3. Gravel, M. and Price W.L.: Visual Interactive Simulation Shows How to Use the Kanban Method in Small Business. "Interfaces" 1991, 21, pp. 22-33.
4. Gravel M. and Price W.L.: Using the Kanban in job shop environment. "International Journal of Production Research" 1988, 26, pp.1105-1118.
5. Greco S., Matarazzo B. and Słowinski R.: Rough approximation of a preference relation by dominance relations. "European Journal of Operational Research" 1999, 117, pp. 63-83.
6. Huang C.C., Kira D. and Vertinsky I.: Stochastic dominance rules for multiattribute utility functions. "Review of Economics Studies" 1978, 41, pp. 611-616.
7. Kahneman D. and Tversky A.: Prospect theory: an analysis of decisions under risk. "Econometrica" 1979, 47, pp. 262-291.
8. Li J.W.: Investigating the efficacy of exercising JIT practices to support pull production control in a job shop environment. "Journal of Manufacturing Technology Management" 2005, Vol. 16, Issue 7, pp. 765-783.
9. Li J.W. and Barnes D.J.: Investigating the factors influencing the shop performance in a job shop environment with kanban-based production control. "International Journal of Production Research" 2000, Vol. 38, No 18, pp. 4683-4699.
10. Martel J.M. and Zaras K".: Une méthode multicritère de rangement de projets face au risque. "MS/OR Division" 1995, Vol. 16, No 2, pp. 135-144.
11. Nowak M., Trzaskalik T., Trzpiot G. and Zaras K. : Inverse stochastic dominance and its application in production process control. In: Multiple Objective and Goal Programming. Recent Developments. Eds. Trzaskalik T., Michnik J. Physica-Verlag, Heidelberg 2002, pp. 362-376.

12. Nowak M.: Wielokryterialne wspomaganie sterowania procesem produkcyjnym z wykorzystaniem dominacji stochastycznych. Ph.D. thesis. AE, Katowice 2002.
13. Roy B.: Méthodologie Multicritère d'Aide à la Décision. Economica, Paris 1985.
14. Zaras K.: Rough approximation of a preference relation by multi-attribute stochastic dominance for deterministic, stochastic and fuzzy decisions problems. "European Journal of operational Research" 2004, 159, pp. 196-206.
15. Zaras K.: Rough approximation of a preference relation by multi-attribute stochastic dominance for deterministic and stochastic evaluations problems. "European Journal of operational Research" 2001, 130, pp. 305-314.
16. Zaras K.: Rough approximation of pairwise comparisons described by a multi-attribute stochastic dominance. "Multicriteria Journal of Decision Analysis" 1999, 8, pp. 291-297.
17. Zaras K. and Martel J.M.: Multi-attribute analysis based on stochastic dominance. In: Models and Experiments in Risk and Rationality. Eds. Munier B., Machina M.J. Kluwer Academic Publishers, Dordrecht 1994, pp. 225-248.
18. Zaras K.: Les dominances stochastiques pour deux classes de fonction d'utilité: concaves et convexes. "RAIRO/RO" 1989, Vol. 23, No 1, pp. 57-65.

Sebastian Sitarz

SENSITIVITY ANALYSIS IN LINEAR VECTOR OPTIMIZATION

Abstract

In this paper the element-wise analysis approach to sensitivity analysis in linear vector optimization is presented. Two cases are considered: sensitivity analysis of efficient solutions and sensitivity analysis of dominating solutions. The results obtained allow to create methods based on the analysis of a simplex tableau. The presented approach allowed to obtain the intervals of the parameter for which a given solution is efficient or dominating.

Keywords

Multicriteria linear programming, sensitivity analysis, postoptimal analysis.

INTRODUCTION

Sensitivity analysis is an important tool in decision-making theory. Many research papers have been published in this field in recent years. Kuk et al. [4] consider three types of perturbation maps: perturbation maps, proper perturbation maps, and weak perturbation maps, corresponding to three kinds of solution concepts: minimality, proper minimality, and weak minimality with respect to a fixed ordering cone for a vector optimization problem. Sensitivity analysis for multiobjective linear programming problems based on scalarization was presented by Vetschera [7], although the volume-based sensitivity analysis was used. In the paper of Thuan and Luc [6] it is proved that if the data of a linear multiobjective programming problem are smooth functions of a parameter, then in the parameter space there is an open dense subset where the efficient solution set of the problem can be locally represented as a union of some faces whose vertices and directions are smooth functions of the parameter. Yildirim [8] presents a unifying geometric framework to extend the optimal partition approach to sensitivity analysis in convex conic optimization. Gunawan and Azarm [3] present a method to measure the multiobjective robustness of a design alternative using the sensitivity region concept and an approach using that measure to obtain robust Pareto solutions of multicriteria programming problems.

The paper presents sensitivity analysis of a chosen efficient (or dominating) solution in vector linear optimization. The case of parameterizing the objective function coefficient is considered. The methods presented allow to use analysis of simplex tableau. Here, the author examines the sensitivity of a single efficient (or dominating) solution. The postoptimization problem is presented in the following ways:

1. If and when a given efficient solution remains an efficient solution after a certain change of the objective function.

2. If and when a given dominating solution remains a dominating solution after a certain change of objective function.

The paper is structured as follows. Section 2 presents the basic theory and notation of linear vector optimization. Section 3 describes the formulation of the considered problems: sensitivity analysis of efficiency and sensitivity analysis of domination. For illustration, a numerical example is presented in Section 4. The last section consists of concluding remarks and further research.

1. LINEAR VECTOR OPTIMIZATION

We will consider the following vector linear optimization problem:

$$\text{VMax } \{\mathbf{Cx}: \mathbf{x}\in X\} \tag{1}$$

where

$X=\{\mathbf{x}\in R^n: \mathbf{Ax}\leq\mathbf{b}, \mathbf{x}\geq 0\}$ or $X=\{\mathbf{x}\in R^n: \mathbf{Ax}=\mathbf{b}, \mathbf{x}\geq 0\}$ – feasible region in decision space.

$\mathbf{x}\in R^n$ – vector of decision variables

$\mathbf{C}\in R^{n,k}$ – matrix of objective function coefficients

$\mathbf{A}\in R^{n,m}$ – full row rank matrix of constraint coefficients

$\mathbf{b}\in R^m$ – right hand side vector

We call $\hat{\mathbf{x}}\in X$ the dominating solution of (1) if

$$\underset{\mathbf{x}'\in X}{\forall} \mathbf{C}\hat{\mathbf{x}}\geq\mathbf{Cx}'$$

We will denote the set of all dominating solutions of the problem (i) by $X_{D(i)}$.

We call $\mathbf{x}^*\in X$ the efficient solution of (1) if

$$\sim \underset{\mathbf{x}'\in X}{\exists} \mathbf{Cx}^*\leq\mathbf{Cx}'\wedge\ \mathbf{Cx}^*\neq\mathbf{Cx}'$$

We will denote the set of all efficient solutions of the problem (i) by $X_{S(i)}$.

1.1. Simplex tableau

We will use the following notation for the problem (1):

$\mathbf{A}_B = [\mathbf{a}^{j_1}, \mathbf{a}^{j_2}, ..., \mathbf{a}^{j_m}]$– basic columns of $\mathbf{A}$

$B = \{j_1, ..., j_m\}$ – index set of base

$\mathbf{A}_N$ – nonbasic columns of $\mathbf{A}$

$\mathbf{x} = [\mathbf{x}_B, \mathbf{x}_N]$ – basic solution associated with B, $(\mathbf{x}_B=\mathbf{A}_B^{-1}\mathbf{b}\geq\mathbf{0}, \mathbf{x}_N=\mathbf{0})$

$\mathbf{C}_B$ – basic columns of $\mathbf{C}$

$\mathbf{C}_N$ – nonbasic columns of $\mathbf{C}$

$\overline{\mathbf{C}} = \mathbf{C} - \mathbf{C}_B\mathbf{A}_B^{-1}\mathbf{A}$ – reduced cost matrix

$\overline{\mathbf{C}}_N = \mathbf{C}_N - \mathbf{C}_B\mathbf{A}_B^{-1}\mathbf{A}$ – reduced cost matrix associated with nonbasic variables.

Using these symbols, we will denote a simplex tableau as presented in Table 1.

Table 1

Organization of a simplex tableau of the problem (1)

	$\mathbf{x}$	
$\mathbf{x}_B$	$\mathbf{A}_B^{-1}\mathbf{A}$	$\mathbf{A}_B^{-1}\mathbf{b}$
	$\overline{\mathbf{C}}$	

1.2. Testing efficiency

Consider the following single objective linear programming problem associated with the basic solution x*:

$$\begin{aligned} &\text{Max } \mathbf{1}^T\mathbf{v} \\ &-\overline{\mathbf{C}}_N\mathbf{y} + \mathbf{I}\mathbf{v} = \mathbf{0} \\ &\left[\mathbf{A}_B^{-1}\mathbf{A}_N\right]_D\mathbf{y} + \mathbf{I}\mathbf{s} = \mathbf{0} \\ &\mathbf{0}\leq\mathbf{y},\ \mathbf{0}\leq\mathbf{v},\ \mathbf{0}\leq\mathbf{s} \end{aligned} \tag{2}$$

where

$\mathbf{1} = [1,1,\ldots,1]^T$ vector of ones,

$\mathbf{I}$ – identity matrix.

$\left[\mathbf{A}_B^{-1}\mathbf{A}_N\right]_D$ – the rows of $\mathbf{A}_B^{-1}\mathbf{A}_N$ associated with degenerated basic variables.

Theorem 1 [5]

The solution **x*** is efficient if and only if the problem (2) has a bounded objective function value of zero.

2. SENSITIVITY ANALYSIS

We want to determine a region of the parameter t such that the feasible solution x* is an efficient (dominating) solution of the following problem:

$$\text{VMax } \{ \mathbf{D}_t^{ij} \mathbf{x} : \mathbf{x} \in X \} \tag{3}$$

where $\mathbf{D}_t^{ij}$ is this matrix obtained from matrix $\mathbf{C}$ by changing element c_{ij} into parameter t. To make the notation clear we will omit the indexes "ij" and "t", ie.: $\mathbf{D} := \mathbf{D}_t^{ij}$. The elements d_{kl} of matrix $\mathbf{D}$ are described as follows:

$$d_{kl} = \begin{cases} c_{kl}, & \text{if } (k,l) \neq (i,j) \\ t, & \text{if } (k,l) = (i,j) \end{cases}$$

Moreover, we will denote the reduced cost matrix of the problem (3) by $\overline{\mathbf{D}}$.

2.1. Testing efficiency

The problem test (2) constructed for the problem (3) has the following form:

$$\begin{aligned} &\text{Max } \mathbf{1}^{\mathrm{T}}\mathbf{v} \\ &-\overline{\mathbf{D}}_{\mathrm{N}}\,\mathbf{y} + \mathbf{I}\mathbf{v} = \mathbf{0} \\ &\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}_{\mathrm{N}}\right]_{\mathrm{D}}\mathbf{y} + \mathbf{I}\mathbf{s} = \mathbf{0} \\ &\mathbf{0} \le \mathbf{y},\ \mathbf{0} \le \mathbf{v},\ \mathbf{0} \le \mathbf{s} \end{aligned} \tag{4}$$

Let us discuss the effect of parameterizing the coefficient c_{ij} in the problem (1) on the problem (4). We will analyze the problem in two cases: when $j \notin B$ and $j \in B$.

Case: $j \notin B$.

In this case the reduced cost matrix has the form:

$$\overline{\mathbf{D}}_{\mathrm{N}} = \mathbf{C}_{\mathrm{N}}(t) - \mathbf{C}_{\mathrm{B}}\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}$$

In the above equation the parameter t appears only in one element c_{ij}. Hence

$$\left(\bar{d}_N\right)_{kl} = \begin{cases} \left(\bar{c}_N\right)_{kl} & \text{if } (k,\ l) \neq (i,\ j) \\ t - \mathbf{C}_B^i \left[\mathbf{A}_B^{-1}\mathbf{A}\right]^j & \text{if } (k,\ l) = (i,\ j) \end{cases}$$

where

$\mathbf{C}_{\mathrm{B}}^i$ – i-th row of matrix $\mathbf{C}_{\mathrm{B}}$

$\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}\right]^j$ j-th column of matrix $\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}\right]$

Therefore in a case where $j \notin \mathrm{B}$ the only one element of the constraint matrix $[\overline{\mathbf{D}}_{\mathrm{N}}, \mathrm{I}]$ depends on t.

Case: $j \in \mathrm{B}$

In this case the reduced cost matrix has the form:

$$\overline{\mathbf{D}}_{\mathrm{N}} = \mathbf{C}_{\mathrm{N}} - \mathbf{C}_{\mathrm{B}}(t)\,\mathbf{A}_{\mathrm{B}}^{-1}\,\mathbf{A}$$

In the above equation the parameter t appears only in i-th row of $\mathbf{C}_{\mathrm{B}}(t)$. Hence

$$\left(\bar{d}_{\mathrm{N}}\right)_{kl} = \begin{cases} \left(\bar{c}_{\mathrm{N}}\right)_{kl} & \text{if } k \neq i \\ \left(c_{\mathrm{N}}\right)_{il} - \mathbf{C}_{\mathrm{B}}^i(t)\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}\right]^l & \text{if } k = i \end{cases}$$

where

$\mathbf{C}_{\mathrm{B}}^i(t) = [c_{i1}, \ldots, c_{i,j-1},\ t,\ c_{i,j+1}, \ldots, c_{im}]$ – i-th row of matrix $\mathbf{C}_{\mathrm{B}}$.

$\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}\right]^l$ – l-th column of matrix $\left[\mathbf{A}_{\mathrm{B}}^{-1}\mathbf{A}\right]$

Therefore in a case where $j \in \mathrm{B}$ the only one row of the constraint matrix $[\overline{\mathbf{D}}_{\mathrm{N}}, \mathrm{I}]$ depends on t. This row is associated with objective function $\mathbf{c}_i^{\mathrm{T}}\mathbf{x}$

We have shown that parameterizing the coefficient c_{ij} in the problem (1) causes the paramaterizing constraint matrix in the problem (4). The methods of analyzing such problems were widely discussed by many authors. Below we present some of them in view of the results obtained earlier.

Consider a single objective parametric linear programming:

$$\begin{aligned} &\text{Max } \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &\mathbf{x} \in X = \{\ (\mathbf{A} + \mathbf{A}^* t)\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\ \} \end{aligned} \tag{5}$$

where $\mathbf{A}^*$ denotes the matrix consisting of coefficients of parameter t.

The simplest case, where only one element of the constraint matrix changes (A^* has all zero elements except one), is examined among others in Dinkelbach's book [1] which contains methods of sensitivity analysis for this case. Let us notice that such parameterization presents sensitivity analysis of efficient solution due to c_{ij} when $j \notin B$ (see chapter 3.1).

Gal's book [2] gives algorithms for finding problem solutions (5) in two cases. One case deals with parameterization of one constraint matrix row (A^* has all zero rows except one). The other case deals with parameterization of one column of constraint matrix (A^* has all zero columns except one).

2.2. Testing domination

We want to verify that the dominating solution $\hat{\mathbf{x}}$ of (1) remains the efficient solution of (3).

Point $\hat{\mathbf{x}}$ is dominating a solution of (3) if and only if it is an optimal solution for all objective functions. Thus, it is an optimal solution if the optimality condition (by means of simplex tableau) is satisfied. This condition is satisfied if $\overline{\mathbf{D}} \leq 0$ (all elements of matrix $\overline{\mathbf{D}}$ should be nonpositive).

Below, we present an example of sensitivity analysis in the presented model.

3. EXAMPLE

Consider the problem:

$$\begin{gathered} \text{VMax } [4x_1 - x_2,\ \ x_1 + 3x_2] \\ x_1 + x_2 \leq 6, \\ x_1 + 2x_2 \leq 10, \\ 0 \leq x_1,\ 0 \leq x_2 \end{gathered}$$

The set of feasible solutions X is a polyhedron with the extreme points: $\mathbf{x}^1 = [0, 0]$, $\mathbf{x}^2 = [0, 5]$, $\mathbf{x}^3 = [2, 4]$, $\mathbf{x}^4 = [6, 0]$. The set of all efficient solutions contains two edges: $\overline{\mathbf{x}^1\mathbf{x}^2}$ and $\overline{\mathbf{x}^2\mathbf{x}^3}$. Figure 1 presents the graphical illustration of this problem in the decision space.

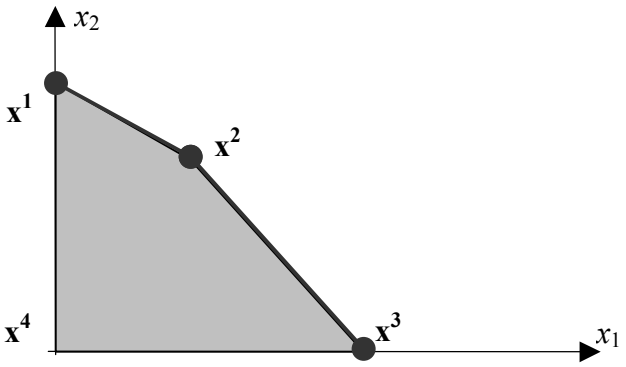


Fig. 1. Graphical illustration of the example in the decision space.

Let us analyze the sensitivity of efficiency for the extreme point $\mathbf{x}^2=[2,4]$ considering the coefficient $c_{22}=t$. The simplex tableau for the problem (1) associated with extreme point $\mathbf{x}^3$ is shown in Table 2.

Table 2

Simplex tableau related to x^2 (Example 1)

	x_1	x_2	x_3	x_4	
x_1	1	0	2	-1	2
x_2	0	1	−1	1	4
	0	0	−7	3	
	0	0	−2+t	1−t	

Using the reduced cost matrix presented in table 2:

$$\overline{\mathbf{D}}_N = \begin{bmatrix} -7 & 3 \\ -2+t & 1-t \end{bmatrix}$$

we obtain the problem test (4) for $\mathbf{x}^3$:

$$\text{Max } v_1 + v_2$$
$$7y_1 - 3y_2 + v_1 = 0$$
$$(2-t)y_1 + (-t+1)y_2 + v_2 = 0$$
$$y_1, y_2, v_1, v_2 \geq 0$$

The initial simplex tableau for this problem is presented in Table 3.

Table 3

Initial tableau for problem test considering point x^3 (Example 1)

	y_1	y_2	v_1	v_2	
v_1	7	−3	1	0	0
v_2	2−t	−1+t	0	1	0
	−9+t	4−t	0	0	

It is easy to determine for which t the solution presented in table 3 remains optimal:

$$(-9+t \le 0) \ \wedge (4-t \le 0) \ \Leftrightarrow t \in [4, 9]$$

Let $t \le 4$. Consider two cases: $t \le 1$ and $t \ge 1$

- if $t \le 1$ we have a nonpositive column accompanied by a positive reduced cost of y_2:

$$\begin{bmatrix} -3 \\ -1+t \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ for } t \le 1$$

This means that our linear program has an unbounded optimal value. Thus for $t \in (-\infty, 1]$ the extreme point $\mathbf{x}^2$ is not an efficient point.

- if $t > 1$ (and $t \le 4$). Pivoting y_2 into the basis, we obtain Table 4.

Table 4

Second simplex tableau for the problem test considering point x^2 (Example 1)

	y_1	y_2	v_1	v_2	
v_1	$\frac{-1+4t}{-1+t}$	0	1	$\frac{3}{-1+t}$	0
y_2	$\frac{2-t}{-1+t}$	1	0	$\frac{1}{-1+t}$	0
	$\frac{1-4t}{-1+t}$	0	0	$\frac{-4+t}{-1+t}$	

To determine for which t the solution presented in Table 4 remains optimal we solve the system of inequalities:

$$\frac{1-4t}{-1+t} \le 0 \text{ and } \frac{-4+t}{-1+t} \le 0.$$

Since $1<t\le 4$ this system is satisfied for each $t\in(1, 4>$.

Let $t\ge 9$. In this case, pivoting y_1 into the basis (table 3), we obtain Table 5.

Table 5

Third simplex tableau for the problem test considering point x^2 (Example 1)

	y_1	y_2	v_1	v_2	
y_1	1	$-\frac{3}{7}$	$\frac{1}{7}$	0	0
v_2	0	$-\frac{1}{7}+\frac{4}{7}t$	$-\frac{2}{7}+\frac{1}{7}t$	1	0
	0	$\frac{1}{7}-\frac{4}{7}t$	$\frac{9}{7}-\frac{1}{7}t$	0	

To determine for which t the solution presented in Table 4 remains optimal we solve the system of inequalities:

$$\frac{1}{7}-\frac{4}{7}t \le 0 \text{ and } \frac{9}{7}-\frac{1}{7}t \le 0$$

This system is satisfied for each $t\in[9, \infty)$

To summarize: The extreme point $\mathbf{x}^2=[2, 4]$ is efficient for $c_{22}=t\in(1, +\infty)$.

Investigating other extreme points, we obtain the intervals presented in Table 6.

Table 6

Intervals of the parameter t, for which the extreme points are efficient

	c_{11}	c_{12}	c_{21}	c_{22}
$\mathbf{x^1=[0, 5]^T}$	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(-\infty, 1{,}5)$	$(2, +\infty)$
$\mathbf{x^2=[2, 4]^T}$	$(0{,}5, +\infty)$	$(-\infty, 8)$	$(-\infty, 3)$	$(1, +\infty)$
$\mathbf{x^3=[6, 0]^T}$	$(1, +\infty)$	$(-\infty, 4)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$
$\mathbf{x^4=[0, 0]^T}$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

Now let us analyze the sensitivity of domination for the extreme point $\mathbf{x}^2=[2,4]$ considering the coefficient $c_{22}=t$. First, we look for such values of the parameter t for which the point x^2 is optimal for both criteria. Using the reduced cost matrix (all elements of matrix $\overline{\mathbf{D}}$ should be nonpositive) presented in Table 2, we obtain the following conditions:

$$-7 \le 0$$

$$3 \le 0$$

and

$$-2+t \le 0$$

$$1-t \le 0$$

This system is inconsistent. Thus, there is no such t that x^2 is a dominating point.

Investigating other extreme points, we obtain the intervals presented in Table 7.

Table 7

Intervals of the parameter t, for which the extreme points are dominating

	c_{11}	c_{12}	c_{21}	c_{22}
$\mathbf{x^1=[0, 5]^T}$	$(-\infty, 0{,}5]$	$[8, +\infty)$	$\varnothing$	$\varnothing$
$\mathbf{x^2=[2, 4]^T}$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
$\mathbf{x^3=[6, 0]^T}$	$\varnothing$	$\varnothing$	$[3, +\infty)$	$(-\infty, 1]$
$\mathbf{x^4=[0, 0]^T}$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

CONCLUSIONS AND FURTHER RESEARCH

The element-wise analysis approach to sensitivity analysis in linear vector optimization was presented. Two cases were considered: sensitivity analysis of efficient solutions and sensitivity analysis of dominating solutions. The results obtained allow to create methods based on analysis of a simplex tableau. The approach presented here allowed to obtain the intervals of parameter for which a given solution is efficient or dominating. We presented examples which let us analyse the described methods.

It is worth considering a case of vector perturbation (instead only of one element). Parameterizing the vector of the coefficients in cost matrix causes parameterizing of one column of constraint matrix in one-criterion linear program, which may be the subject of further research.

REFERENCES

1. Dinkelbach W.: Sensitivitätsanalysen und parametrische Programmierung. Springer, Berlin-New York 1969.
2. Gal T.: Postoptimal analyses, parametric programming, and related topics. Walter de Gruyter, Berlin 1995.
3. Gunawan S., Azarm S.: Multi-objective robust optimization using a sensitivity region concept. „Structural and Multidisciplinary Optimization" 2005, 29, pp. 50-60.
4. Kuk H., Tanino T., Tanaka: Sensitivity analysis in parametrized convex vector optimization. „Journal of Mathematical Analysis and Applications" 1996, 202, 3, pp. 511-522.
5. Steuer R.: Multiple Criteria Optimization Theory: Computation and Application. John Willey, New York 1986.
6. Thuan L.V. , Luc D.T.: On Sensitivity in Linear Multiobjective Programming. "Journal of Optimization Theory and Applications" 2000, 107, 3, pp. 615-626.
7. Vetschera R.: Volume-Based Sensitivity Analysis for Multi-Criteria Decision Models. In: Methods of Multicriteria Decision Theory. Eds. Göpfert A., Seeländer J., Tammer Chr. Hänsel-Hohenhausen, Egelsbach 1997.
8. Yildirim E.A.: Unifying Optimal Partition Approach to Sensitivity Analysis in Conic Optimization. "Journal of Optimization Theory and Applications" 2004, 122, 2, pp. 405-423.

Jerzy Michnik

MULTI-CRITERIA MODELLING OF INTEGRATED ASSET & LIABILITY MANAGEMENT IN A COMMERICAL BANK

Abstract

One of the most important category of risk banks face is the financial risk. Asset & Liability Management (ALM) is a set of techniques used to manage financial risk. Growing instability in the financial world made ALM a great challenge for both researchers and practitioners.

A basic structure of the ALM model, based on the anticipated cash flows, is constructed. It comprises the main financial risks: interest rate, foreign exchange, liquidity and capital risk. The illustration models which are set up in a framework of the linear programming, deterministic or stochastic, are presented. The simplified cases with simulated data, illustrating the activity of a commercial bank in Poland, are solved with the aid of interactive goal programming.

Keywords

Financial risk management, asset & liability management, commercial banking, multiple criteria decision making, interactive goal programming, stochastic programming.

INTRODUCTION

Banks are the major part of the financial markets and the financial risk is a serious matter for them. This is why the Asset & Liability Management (ALM) is evolving rapidly in the banking industry. Another important factor contributing to the development of risk management was the rapid advance in the state of information technology.

The aim of this paper is an analysis of multi-criteria methods which can support the decision making in ALM process. The general background for the decision making support model is presented.

The assumptions for a satisfactory ALM model are:

- the realization of an integrated asset and liability management,
- the openness and flexibility (easy modifications according to variable external conditions and changing of the bank management preferences),
- the adequacy to real processes in a bank.

1. ASSET & LIABILITY MANAGEMENT IN A BANK

ALM is the set of techniques used to manage important issues of financial risk. ALM also deals with the structure of the balance sheet, given funding constraints, regulatory and profitability targets [1].

The main important factors contributing to the rapid evolvement of the Asset & Liability Management (ALM) are:

- Banks are the major part of the financial markets and the financial risk is a serious matter for them.
- The growth of instability in the financial markets.
- The growth of trading activity.
- The development in derivatives and growth of derivatives activity.
- The rapid advance of the state of information technology.

The enormous development of both theory and practice of risk management in last thirty years deserves the name of the 'risk management revolution' [3].

There are several risks, several possible targets, several measures of each dimension of risk, several types of tools and techniques [13]. This is the reason why the ALM is a complicated, multi-step and multidimensional process.

The main steps of ALM process are:

1. Recognizing the main types of financial risk and their sources.
2. Choosing the appropriate risks' measures which constitute a base for the risk management system.
3. Establishing the set of financial instruments which are used for hedging purposes.

Of course there are many other technical problems which arise during the implementation phase, but they are not the subject of this paper. In the literature different views of the scope of the ALM can be found. They depend in part on the author opinions. The same situation is in banks, where the concrete ALM process depends strongly on the view of bank's authority. Despite of these differences there exists the core content of the ALM. It comprises the interest rate risk, foreign exchange risk, liquidity and capital management.

2. BASIC FRAMEWORK

The foundations of a model are as follows:

- The risk is a result of the uncertainty of future cash-flows.
- We consider a finite period of time divided into finite number of periods.
- All cash-flows take place only at the end of any period (simplification).

We need an index set to distinguish various financial instruments used for risk management. It is convenient to use three finite sets of indices. As a result every cash flow is identified by time, currency and the additional features as follows:

- $t \in T$– indicates the time period,
- $v \in V$ – indicates the currency of the instrument,
- $z \in Z$ – indicates definite instrument in given currency,

where T, V, Z – finite sets.

Then every financial instrument is characterized by the set of its cashflows

$$x = [x_t(v, z)], \quad t \in T \tag{1}$$

For given v and z the instrument becomes:

- a vector – deterministic case,
- a random vector – stochastic case.

With each financial instrument we associate a decision variable – a single non-negative number

$$x_t(v, z) \tag{2}$$

which determines all cash-flows of the instrument. The detailed description of the financial instruments with their cash-flows will be presented later, in the section describing the illustrative model. Below we provide two examples examples illustrating the way of deriving cash-flow structure from the generic characteristics of the instrument.

Let us start with the purchase of the instrument with coupon payments. A decision variable $x_t(v, z)$ is a volume which is purchased. At the time of purchase, there is the negative cash-flow $- x_t(v, z)$. In next consecutive periods $t' > t$, there is the series of coupon payments of the form $r_{t''}(v, z)x_{t''}(v, z)$, where $r_{t''}$ are the interest rates. At the maturity we have interest payment plus capital return: $x_{t''}(v, z) + r_{t''}(v, z)x_{t''}(v, z)$, $t'' > t'$.

In the case of purchase of interest rate call option, the decision variable $x_t(v, z)$ is the principal amount of the option and $t > 0$. At the beginning period we pay for the option: $- r_0(v, z)x_0(v, z)$, where $r_0(v, z)$ is the unit price of the option (expressed as percentage). At the exercise date, the conditional flow $x_t(v, z)[r_t(v, z) - \hat{r}_t(v, z)]+$, $t > 0$ takes place, where $\hat{r}_t(v, z)$ is the exercise rate of the option (we use the notation $[u]+ = \min\{0, u\}$).

3. MODEL ASSUMPTIONS

The division of the model constituents into hard constraints and criteria is, to the high degree, the matter of the analyst's choice. So the presented model can

be regarded as one of many other versions equally possible. Here we restrict ourselves to the linear model. Consequently we use the linear forms of general expressions presented below.

3.1 Constraints

From the modelling point of view there are two types of constraints:
1. deterministic

$$Bx \leq b \tag{3}$$

where matrix B and vector b are deterministic.
2. stochastic

$$H(\xi)x \leq h(\xi) \tag{4}$$

where both matrix H and vector h generally depend on random parameters.

According to their origin the constraints can be divided into three groups:

i. Market and technical limits.
ii. Constraints which are imposed by the legal system of the country. (e.g. the bottom limits for the capital adequacy ratio and for cash reserves.)
iii. Internal risk management constraints.

It should be noted that the above division has a rather formal character. In practice all three groups have much in common – the sources of concrete restrictions can lie in all of the above groups. For instance capital adequacy ratio limit and cash reserves limits definitely belong to the second group. However they can be regarded as the part of internal risk management constraints as well. It is common that the management of a bank imposes more strict conditions on cash reserves and sometimes on capital adequacy ratio.

3.2. Criteria

When the problem of risk management is of concern, the decision maker should:

- minimize the several types of risk,
- maximize profitability or worth of a bank.

4. MODEL DETAILS

4.1. Market and Technical Limits

These limits have to be proposed by the bank specialists and confirmed by the management of a bank. Most of them are upper limits for bank's dealers transactions and open positions. They have the deterministic form (3).

4.2. Legal System Constraints

4.2.1. Capital Adequacy Ratio

We set the bottom limit for the capital adequacy ratio. This is a deterministic constraint which can be written as the following inequality

$$A_0 - \sum_{v \in \mathcal{V}} \sum_{z \in \mathcal{Z}_a} w(v,z) q_0(v,v_0) x_0(v,z) \leq \frac{C}{\omega} \tag{5}$$

where A_0 initial value of the weighted sum of risk assets from the balance sheet, $Z \supset Z_a$ – contains only asset transactions, $q_0(v, v_0)$ translates transaction in foreign currency v into local currency v_0, $w(v, z)$ – risk factors, C – equity of the bank, ω – the minimal level of the capital adequacy ratio.

4.2.2. Cash Reserves

These arise as the recurrence series of inequalities

$$\sum_{z \in \mathcal{Z}} \sum_{t'=0}^{t} [x_{t'}(v,z) + p_{t'}(l,v)] \geq l_t(l,v), \quad t \in \mathcal{T},\ v \in \mathcal{V} \tag{6}$$

where $p_t(l, v)$ – the initial level of cash reserves of currency v at time t, $l_t(l, v)$ – the minimal requirement of cash at time t.

4.3. Criteria

An explanation should be given for the reasons for choosing the specific measures for criteria in the model. Among many possibilities, sensitivity measures were chosen. Thanks to their simplicity and convenience, they are widely used in bank practice, as it was proven by the investigation made in US and foreign banks [10]. In the last years a more sophisticated measure achieved a great success: the downside risk measure – Value at Risk [3; 8]. However it should be noted that it has its own disadvantages. It does not satisfy the conditions attributed to proper risk measures [12]. Moreover, such a risk measure introduced into banking supervisory regulations can even deepen the market crises [2]. The additional advantage of sensitivity measures chosen here is that they are linear in decision variables (under the special assumptions, the minimization of VaR can also be done by linear programming [9]).

4.3.1. Interest Rate Risk

As the criterion we take the minimization of change of market value of bank's equity due to interest rate change. Let's define

$$D_X^n = D_X PV_X = \sum_t t\gamma_t \; cf_{X,t}(x) \tag{7}$$

where $X = A$ for assets, $X = L$ for liabilities,
D_X, PV_X and $cf_{X,t}(x)$ – duration, present value and sum of all positive (negative) cash-flows at the moment t of assets (liabilities), respectively, γt – discount factor. As the risk is minimal for zero gap between assets and liabilities, we can write down

$$D_L^n - D_A^n \to 0 \tag{8}$$

4.3.2. Foreign Exchange Risk

As a measure of the foreign exchange risk with we propose foreign currency position. It is defined for each foreign currency as follows:

$$P_t(v) = p_t(v) + \sum_{z \in \mathcal{Z}_f} x_t(v, z), \quad t \in \mathcal{T} \tag{9}$$

where $P_t(v)$ – position for period t and for foreign currency v, $p_t(v)$ – initial position for period t and currency v, $Z \supset Z_f$ – transactions changing position in a given foreign currency.

The minimal risk occurs when the position is zero (closed), so similarly to the interest rate case, we write down

$$P_t(v) \to 0 \tag{10}$$

4.3.3. Profitability

We can express the profit-loss of all the transactions as

$$\sum_{v \in \mathcal{V}} \sum_{z \in \mathcal{Z}} \sum_{t \in \mathcal{T}} \gamma_t \, R[r_t(v, z)] \, x_t(v, z) \, q_t(v, v_0) \to \max \tag{11}$$

where

$$R[r_t(v, z)] = \begin{cases} \pm r(v, z), & \text{for interest rate transactions} \\ \pm (r - \hat{r})^+ & \text{for options} \\ 1, & \text{for FX transactions} \end{cases} \tag{12}$$

The sign ± in above expressions depends on the meaning of specific transaction (profit or cost in interest rate transaction, purchase/sell of call or put option). $q_t(v, v_0)$ translates transaction in foreign currency v into local currency v_0; γ_t is a discount factor.

4.4. Uncertainty in the model

We need some kind of realistic model, closely describing reality on the one hand and not too complicated on the other. Let's assume that the uncertainty is introduced to the model by the randomness of market parameters: interest rates and foreign exchange rates. If we want to limit ourselves to the linear programming we have to restrict the distributions of random variables to the discrete finite case. The deterministic model can be obtained as a special case under the assumption that all random parameters have single-valued distributions.

4.5. General Form of the Model

As the result of the above assumptions we obtain the multicreterial linear programming model. We are going to use the interactive goal programming procedure for solving it, so we present it in the form

$$\begin{array}{rcl} \min_{x,y}[cy] & & \\ Bx & \leq & b \\ H(r,q)x & \leq & h(r,q) \\ G(r,q)x - y^+ + y^- & = & g \\ x, y & \geq & 0 \end{array} \tag{13}$$

where we use auxiliary constraints for goals. y^+ and y^- are over- and underachievement variables, vector g represents the aspiration levels of goals. In deterministic case the above model can be solved as it stands. In stochastic case we need to find its deterministic equivalent [6].

4.5.1. Deterministic equivalent of stochastic model

It has been shown that stochastic goal programming model is a particular case of stochastic linear programming with recourse [5]. In our case it is convenient to formulate it as the multistage recourse program.

We need additional auxiliary variables y′ which serve to compensate the violation of stochastic constraints in some realizations of random parameters. The number of variables y′ is equal to the number of stochastic inequalities in (13).

The deterministic equivalent looks as follows

$$\begin{aligned} \min_x E_{r,q}\{cx + Q_1(x,r,q) + Q_2(x,y',r,q)\} & \\ Bx &\leq b \\ H(r,q)x - Wy' &\leq h(r,q) \\ G(r,q)x + W'y' - y^+ + y^- &= g \\ x, y, y' &\geq 0 \end{aligned} \tag{14}$$

where $E_{r,q}$ stands for expectation value with respect to the distributions of r, q. $Q_2(x, y', r, q)$ is the recursion function of the third stage, defined as follows

$$Q_2(x,y',r,q) = \min_y\{c^T y \mid y \geq 0\} \tag{15}$$

$Q_1(x, r, q)$ is the recursion function of the second stage, given by the formula

$$Q_1(x,r,q) = \min_{y'}\{d^T y' \mid y' \geq [H(r,q)x - h(r,q)]^+\} \tag{16}$$

5. ILLUSTRATIVE MODEL

For the illustrative model two time periods $T = \{0, 1, 2\}$ and two currencies $V = \{1, 2\}$ ($v = 1$ for local currency) were chosen. The detailed presentation of all decision variables with their cash flows is given below.

5.1. Decision Variables and Cash-Flows

5.1.1. Short term fixed interest rate transactions (e.g. treasury bills)

Purchase The first two variables ($z = 1$) describe the purchase of the bills which mature at the first period, the next two ($z = 2$) – of those which mature at the second period.

$$x_t(v,1) = \begin{bmatrix} -x_0(v,1) \\ [1 + r_0(v,1)]\, x_0(v,1) \\ 0 \end{bmatrix} \tag{17}$$

$$x_t(v,2) = \begin{bmatrix} -x_0(v,2) \\ 0 \\ [1 + r_0(v,2)]\, x_0(v,2) \end{bmatrix} \tag{18}$$

Sell Similar to the purchase case, we have 4 variables for selling:

$$x_t(v,3) = \begin{bmatrix} x_0(v,1) \\ -\left[1 + r_0(v,1)\right] x_0(v,1) \\ 0 \end{bmatrix} \tag{19}$$

$$x_t(v,4) = \begin{bmatrix} x_0(v,4) \\ 0 \\ -\left[1 + r_0(v,4)\right] x_0(v,4) \end{bmatrix} \tag{20}$$

As the interbanking transactions look similar, we can easily include them in the model (here we skip them so as not to expand the size of the model).

5.1.2. Coupon bonds

Again 4 variables are assigned to purchase and 4 to the selling of instruments with coupons.

Purchase

$$x_t(v,5) = \begin{bmatrix} -x_0(v,5) \\ \left[1 + r_0(v,5)\right] x_0(v,5) \\ 0 \end{bmatrix} \tag{21}$$

$$x_t(v,6) = \begin{bmatrix} -x_0(v,6) \\ r_0(v,6) x_0(v,6) \\ \left[1 + r_1(v,6)\right] x_0(v,6) \end{bmatrix} \tag{22}$$

Sell

$$x_t(v,7) = \begin{bmatrix} x_0(v,7) \\ -\left[1 + r_0(v,7)\right] x_0(v,7) \\ 0 \end{bmatrix} \tag{23}$$

$$x_t(v,8) = \begin{bmatrix} x_0(v,8) \\ -r_0(v,8) x_0(v,8) \\ -\left[1 + r_1(v,8)\right] x_0(v,8) \end{bmatrix} \tag{24}$$

5.1.3. Foreign Exchange Transactions

These transactions occur in pairs. The inflow of one currency is accompanied by the outflow of the other. Here we have 2 spot transactions and 4 forward transactions.

Spot FX

$$x_t(1,9) = \begin{bmatrix} x_0(1,9) \\ 0 \\ 0 \end{bmatrix}, \quad x_t(2,9) = \begin{bmatrix} -q_0(1,2)x_0(1,9) \\ 0 \\ 0 \end{bmatrix} \tag{25}$$

$$x_t(2,10) = \begin{bmatrix} x_0(2,10) \\ 0 \\ 0 \end{bmatrix}, \quad x_t(1,10) = \begin{bmatrix} -q_0(2,1)x_0(2,10) \\ 0 \\ 0 \end{bmatrix} \tag{26}$$

Forward FX

$$x_t(1,11) = \begin{bmatrix} 0 \\ x_1(1,11) \\ 0 \end{bmatrix}, \quad x_t(2,11) = \begin{bmatrix} 0 \\ -q_1(1,2)x_1(1,11) \\ 0 \end{bmatrix} \tag{27}$$

$$x_t(1,12) = \begin{bmatrix} 0 \\ 0 \\ x_2(1,12) \end{bmatrix}, \quad x_t(2,12) = \begin{bmatrix} 0 \\ 0 \\ -q_2(1,2)x_2(1,12) \end{bmatrix} \tag{28}$$

$$x_t(2,13) = \begin{bmatrix} 0 \\ x_1(2,13) \\ 0 \end{bmatrix}, \quad x_t(1,13) = \begin{bmatrix} 0 \\ -q_1(2,1)x_1(2,13) \\ 0 \end{bmatrix} \tag{29}$$

$$x_t(2,14) = \begin{bmatrix} 0 \\ 0 \\ x_2(2,14) \end{bmatrix}, \quad x_t(1,14) = \begin{bmatrix} 0 \\ 0 \\ -q_2(2,1)x_2(2,14) \end{bmatrix} \tag{30}$$

5.1.4. Interest Rate Options

In these transactions the conditional cash-flows appear which depend on the difference between option's exercise rate and the market rate in some future period. In our model we use the purchase of call and put options. The construction of these cash-flows and the notation are explained by the examples in the section 2 (Basic Framework).

Table 1

The list of goals in a model

No.	v	t	meaning
1	1	–	interest rate risk
2	2	–	interest rate risk
3	2	0	FX risk
4	2	1	FX risk
5	2	2	FX risk
6	1	–	profit

Call

$$x_t(v,15) = \begin{bmatrix} -r_0(v,15)x_0(v,15) \\ [r_1(v,15) - \hat{r}_1(v,15)]^+ x_0(v,15) \\ 0 \end{bmatrix} \tag{31}$$

$$x_t(v,16) = \begin{bmatrix} -r_0(v,16)x_0(v,16) \\ 0 \\ [r_2(v,16) - \hat{r}_2(v,16)]^+ x_0(v,16) \end{bmatrix} \tag{32}$$

Put

$$x_t(v,17) = \begin{bmatrix} -r_0(v,17)x_0(v,17) \\ [\hat{r}_1(v,15) - r_1(v,17)]^+ x_0(v,17) \\ 0 \end{bmatrix} \tag{33}$$

$$x_t(v,18) = \begin{bmatrix} -r_0(v,18)x_0(v,18) \\ 0 \\ [\hat{r}_2(v,18) - r_2(v,18)]^+ x_0(v,18) \end{bmatrix} \tag{34}$$

It makes the total of 30 decision variables.

5.2. Goals

The first 2 goals represent the interest rate risk for 2 currencies, as it was described in Section 4.3.1. Next 3 goals deal with foreign exchange risk (Section 4.3.2). The last one is the profitability of all transactions (see Section 4.3.3). The numbers and meanings of goals are summarized in Table 1.

5.3. Numerical tests

Distributions of random parameters are limited to three scenarios with probabilities: $p_1 = 0.25$, $p_2 = 0.5$ and $p_3 = 0.25$. For the deterministic model the same data were used with probabilities: $p_1 = p_3 = 0$, $p_2 = 1$.

Table 2

Interest rates (p.a.) for balance sheet transactions

		p1		p2		p3	
v	z	1	2	1	2	1	2
1	1	0,05450		0,05450		0,05450	
2	1	0,02070		0,02070		0,02070	
1	2		0,05650		0,05650		0,05650
2	2		0,02080		0,02080		0,02080
1	3	0,05450		0,05450		0,05450	
2	3	0,02070		0,02070		0,02070	
1	4		0,05650		0,05650		0,05650
2	4		0,02080		0,02080		0,02080
1	5	0,05450		0,05460		0,05470	
2	5	0,02070		0,02075		0,02079	
1	6		0,05650	0,05455	0,05657	0,0545	0,05660
2	6		0,02080	0,02073	0,02088	0,02078	0,02093
1	7	0,05450		0,05454		0,05458	
2	7	0,02070		0,02074		0,02077	
1	8		0,05650	0,05455	0,05659	0,05458	0,05664
2	8		0,02080	0,02074	0,02086	0,02077	0,02090

In the technical constraints (3), we put $B = I$ and b = [20]. In constraint for capital adequacy ratio (5) we put $C = 100$, $A_0 = 1220$, and $\omega = 0.08$.

In liquidity constraints for all t, t' and scenarios: $p_{tt'}(l, v) = 5.3$, $l_t(l, v) = 5$ for $v = 1$ and $p_{tt'}(l, v) = 2.7$, $l_t(l, v) = 2$ for $v = 2$.

The interest, exchange rates and other parameters are chosen to simulate the Polish market in the middle of 2004. They are presented in Tables 2, 3 and 4.

In both cases: deteministic and stochastic, the interactive goal programming was used as a solving procedure [11]. Calculations were performed with Microsoft Excel and with its accompanying optimization procedure Solver.

In interactive goal programming procedure, at every stage, several optimizations with single criterion are performed. The deterministic equivalent of stochastic programming model (14) contains the equality constraints for every goal. In the case of interest risk and FX forward positions, they are duplicated for every realization of random parameters. Consequently, for them, we need to solve the optimization problem with the third stage recursion function (15), for every realization of random parameters separately.

Table 3

Exchange rates for spot and forward FX transactions

			p1		p2		p3	
	t	0	1	2	1	2	1	2
v	z							
1	9	0,21552						
2	10	4,66						
1	11		0,2153		0,21552		0,2158	
1	12			0,2154		0,21552		0,2158
2	13		4,665		4,66		4,655	
2	14			4,666		4,66		4,657

Table 4

Exercise rates for interest rate options (in column t = 0 option prices)

			p1		p2		p3	
	t	0	1	2	1	2	1	2
v	z							
1	15	0,00000015	0,05450		0,05460		0,05470	
2	15	0,00000015	0,02070		0,02075		0,02079	
1	16	0,00000015		0,05650		0,05657		0,05660
2	16	0,00000015		0,02080		0,02088		0,02093
1	17	0,00000015	0,05450		0,05454		0,05458	
2	17	0,00000015	0,02070		0,02074		0,02077	
1	18	0,00000015		0,05650		0,05659		0,05664
2	18	0,00000015		0,02080		0,02086		0,02090

Table 5

Deterministic model. The complete first iteration of multiple interactive goal procedure and potency matrices for the next two iterations (B-best, W-worst). Aditional constrains (ac): goals 1-5 – upper limit for absolute value; goal 6 – lower limit

	1	2	3	4	5	6
1	0,0000	-0,2379	0,6316	-0,5688	2,0000	-0,0027
2	2,0000	-0,2265	-3,3104	-1,0000	6,3986	0,0517
3	-2,9172	-0,2418	0,0000	0,0601	2,0000	-0,0229
4	-2,6385	-0,2414	0,0603	0,0000	2,0000	-0,0212
5	-0,0000	-0,2566	2,6060	-0,5688	0,0000	-0,0878
6	5,5844	-1,0000	-3,3104	-1,0000	5,6223	0,0616
B	0,0000	-0,2265	0,0000	0,0000	0,0000	0,0616
W	5,5844	-1,0000	-1,0000	-1,0000	6,3986	-0,0878
ac	2	1	1	1	2	-0,1
B	0,0000	-0,2330	0,0000	0,0000	0,0000	0,0249
W	2,0000	-0,9978	0,6316	-1,0000	2,0000	-0,0715
ac	1	1	1	1	1	-0,1
B	0,0000	-0,2414	0,0000	0,0000	0,0000	-0,0110
W	1,0000	-1,0000	1,0000	-0,7035	1,0000	-0,0715

Tables 5 and 6 contain the result of calculations performed for deterministic model.

Tables 7, 8 and 9 present results of calculations for stochastic model.

6. CONCLUDING REMARKS

Presented model:
offers integrated approach to ALM management,
is founded on cash-flows basis,
is general and easy to modify,
takes into account the random nature of market parameters.

It is an attempt to develop the auxiliary tool for a complicated management process such as ALM. In the light of the analysis performed it becomes clear that relatively simple methods can considerably improve the procedure of ALM in a bank. Previous experience shows that the problem is not only academic. However not many, the real-life implementations of optimization methods in financial management in a bank were reported. Let us mention two of them. The goal programming model was implemented in large Greek commercial bank [4]. The two-stage linear programming model [7] was a step towards the stochastic programming as it used a number of alternative scenarios and the expected values for the goals.

Although the models mentioned have some similarities with the model presented here, both were designed to support the financial planning process rather than risk management. They are based on balances rather than on cash-flows and they are designed for longer time periods (years).

The model proposed in this article is a further step in the development of optimization methods designed for ALM. In natural way it suggests further issues and research directions:

- incorporating other types of risk into the model (e.g. credit risk),
- taking into account other risk measures (e.g. non-linear measures),
- developing the stochastic content of the model.

Table 6

Deterministic model. Nonzero variables of the last solution with profit optimization

v	z	x	v	z	x
1	2	18,814	1	8	20,000
1	3	3,045	1	12	4,640
2	3	20,000	2	13	0,296
2	4	18,969	1	16	20,000
2	5	19,669	2	16	20,000
2	6	20,000			

Table 7

Stochastic model. The initial solution of multiple interactive goal procedure with the potency matrix (best, worst)

goal no.		1	2	3	4	5	6	best	worst
int. rate									
	p1	-0,000	-0,024	-2,917	-0,010	-0,024	5,554	0,000	5,554
v=1	p2	-0,000	-0,012	-2,902	-0,005	-0,012	5,576	0,000	5,576
	p3	-0,000	0,000	-2,887	-0,000	-0,000	5,587	0,000	5,587
	p1	-0,249	0,000	-0,258	-0,243	-0,254	-1,000	0,000	-1,000
v=2	p2	-0,250	-0,000	-0,259	-0,243	-0,255	-1,000	0,000	-1,000
	p3	-0,252	-0,000	-0,262	-0,244	-0,258	-1,000	0,000	-1,000
FX pos.									
t=0		0,632	0,626	-0,000	0,629	0,626	-3,310	0,000	-3,310
	p1	1,416	1,416	2,039	0,000	1,416	-1,000	0,000	2,039
t=1	p2	1,416	1,416	2,039	0,000	1,416	-1,000	0,000	2,039
	p3	1,416	1,416	2,039	0,000	1,416	-1,000	0,000	2,039
	p1	0,001	0,001	0,001	1,427	0,001	5,623	0,001	5,623
t=2	p2	0,000	0,000	0,000	1,427	0,000	5,623	0,000	5,623
	p3	-0,003	-0,003	-0,003	1,426	-0,003	5,623	-0,003	5,623
profit		-0,088	-0,094	-0,016	-0,027	-0,087	0,176	0,176	-0,106

Table 8

Stochastic model. Potency matrices for the next 2 iterations with additional constraints (ac)

		2			3		
		ac	B	W	ac	B	W
int. rate							
	p1	0,700	0,000	0,693	0,600	0,000	0,595
v=1	p2	0,700	0,000	0,700	0,600	0,000	0,600
	p3	0,700	0,000	0,707	0,600	0,000	0,606
	p1	0,400	0,000	-1,000	0,200	0,000	0,000
v=2	p2	0,400	0,000	-0,102	0,200	0,000	0,000
	p3	0,400	0,000	-0,700	0,200	0,000	-0,800
FX pos.							
t=0		0,400	0,000	0,400	0,300	0,000	0,300
	p1	0,400	0,000	0,400	0,300	0,000	0,300
t=1	p2	0,400	0,000	0,400	0,300	0,000	0,300
	p3	0,400	0,000	0,400	0,300	0,000	0,300
	p1	0,800	0,001	0,801	0,700	0,001	0,700
t=2	p2	0,800	0,000	0,800	0,700	0,000	0,700
	p3	0,800	-0,003	0,799	0,700	-0,003	0,700
profit		-0,030	0,086	-0,030	-0,20	0,070	-0,020

Table 9

Stochastic model. Nonzero variables of the last solution with profit optimization

v	z	x	v	z	x	v	t		y'
1	2	20,000	2	6	20,000	1	1	p1	9,318
1	3	20,000	1	8	13,296	1	1	p2	9,311
2	3	20,000	1	9	3,248	1	1	p2	9,305
2	4	18,668	1	12	6,031	2	2	p1	1,000
1	5	11,948	2	13	1,300	2	2	p2	1,001
2	5	18,668	2	18	20,000	2	2	p3	0,199

REFERENCES

1. Bessis J.: Risk Management in Banking. J. Wiley & Sons, Chichester 1998.
2. Danielsson J.: The Emperor Has No Clothes: Limits to Risk Modelling. "Journal of Banking & Finance" 2002, 26, pp. 1273-1296.
3. Dowd K.: Beyond Value at Risk. The New Science of Risk Management. J. Wiley & Sons, Chichester 1998.
4. Giokas D., Vassiloglou M.: A goal programming model for bank assets and liabilities management. "European Journal of Operational Research" 1991, 50, pp. 48-60.
5. Heras A., Aguado A.G.: Stochastic Goal Programming. "Central European Journal of Operations Research" 1999, 7, No 3, pp. 139-158.
6. Kall P., Wallace S.W.: Stochastic Programming. J. Wiley & Sons, Chichester 1994.
7. Korhonen A.: A dynamic bank portfolio planning model with multiple scenarios, multiple goals and changing priorities. "European Journal of Operational Research" 1987, 30, pp. 13-23.
8. RiskMetrics-Technical Document (J.P. Morgan/Reuters). Fourth Edition, New York 1996.
9. Rockafellar R.T., Uryasev S.: Optimization of conditional value-at-risk. "Journal of Banking & Finance" 2002, 26.
10. Santomero A.: Commercial Bank Risk Management: An Analysis of the Process. "Journal of Financial Services Research" 1997, 12, 2/3, pp. 83-115.
11. Spronk J.: Interactive Multiple Goal Programming for Capital Budgeting and Financial Planning. Martinus Nijhoff, Boston 1981.
12. Szegö G.: Measures of risk. "Journal of Banking & Finance" 2002, 26, pp. 1253-1272.
13. Uyemura D.G., Van Deventer D.R.: Risk Management in Banking - The Theory & Application of Asset & Liability Management, Bankers Publishing Company, Probus Publishing Company: Chicago 1993.

Dorota Miszczyńska

Marek Miszczyński

APPLICATION OF DEA METHOD TO THE EVALUATION OF THE EFFICIENCY OF POLISH OPEN PENSION FUNDS IN THE YEARS 2004-2006

Abstract

The subject of this analysis is Open Pension Funds (OPF) in the period from 2004 to 2006. The purpose of this analysis is the measurement of technical efficiency of OPF. What was applied in this evaluation was the Data Envelopment Analysis (DEA). Data related to the volume of inputs (number of members, operating costs per capita) and the size of outputs (net assets, result of investments, accounting unit's values) allowed to construct basic models of DEA (output oriented); the CCR model (constant returns to scale), BCC model (variable returns to scale) and NIRS model (non-increasing returns to scale). In order to evaluate changes in efficiency of each OPF (in the years 2004-2006) the distances of Shephard being the basis for the Malmquist indexes were calculated.

Keywords

Open Pension Funds, efficiency evaluation, data envelopment analysis (DEA), Malmquist indexes.

INTRODUCTION

This paper compares presently existing Open Pension Funds (OPF) with relation to their efficiency. What is applied to evaluate this efficiency is DEA (Data Envelopment Analysis), the method whose objects of analysis are defined as Decision Making Units (in short referred to as DMU). In this work the role of DMU is performed by particular OPFs. The inputs of OPF are determined by the number of members and operating costs per capita of a given OPF, while its returns are determined by its net assets, result of investments and accounting unit's values.

Presently about 15 OPFs operate in the market. They are as follows (next to their names the names used in this paper are presented):

AIG OPF (AIG), Allianz Polska OPF (Allianz), Bankowy OPF (Bankowy), Commercial Union OPF BPH CU WBK (Commercial), OPF „Dom" (DOM), OPF Ergo-Hestia (Ergo), Generali OPF; former name Zurich OPF (Generali), ING Nationale Nederlanden Polska OPF (NNeder), Nordea OPF; former name Sampo (Nordea), Pekao OPF (PeKaO), OPF Pocztylion (Pocztylion), OPF Polstat (Polsat), OPF PZU „Złota Jesień" (PZU), OPF Skarbiec-Emerytura (Skarbiec) and Winterthur OPF; former Credit Suisse (Winterthur).

Tables 1 and 2 presents some selected data concerning the funds mentioned above (number of members and operating costs per capita of a given OPF, net assets, result of investments and accounting unit's values) in the years 2004-06 (the first quarters).

The presented data can lead to the conclusion that the market is definitely dominated by four funds, i.e. AIG, Commercial, Nationale Nederlanden and PZU Złota Jesień. These funds comprise 71% of the market measured by the share of net assets (65% measured by the share of members).

The specific character of the institutions represented by OPF imposes special care for proper evaluation of these funds as the objective of OPF is gathering funds and then investing them with the purpose of payment to members of funds when they achieve their pension age. Therefore finding effective methods of monitoring and evaluation related to activity of a particular OPF seems to be of great importance. In her hitherto existing works, where she used multi-criteria methods and forecasts of the rankings, the author mostly focuses on the rankings of OPFs [6; 7; 8]. This paper though concentrates on the evaluation of OPF efficiency. The method applied here (i.e. DEA) was used for the first time in 1978 by Charnes, Cooper and Rhodes [2]. In Polish literature this method is mostly known from the works related to banks' efficiency evaluation [4; 10].

When using the DEA models for the evaluation of OPF efficiency, treated here as Decision Making Units, definitions of all the factors influencing OPF efficiency should be an important stage of this analysis which later are to be transferred onto defined outlays and returns (effects).

Table 1

Selected features of OPF (inputs) in the years 2004-06 (fixed prices 2004)

No.	*OPF*	Number of participants [thousand persons]			Operating costs per capita [PLN (2004) / person]		
		2004	2005	2006	2004	2005	2006
1.	AIG	969,01	960,33	998,50	6,59	9,01	8,53
2.	Allianz	250,89	272,77	295,32	8,08	9,54	11,30
3.	Bankowy	401,77	406,38	435,73	5,82	7,74	9,86
4.	Commercial	2579,83	2557,20	2572,65	8,36	9,83	12,35
5.	DOM	239,38	227,22	257,15	6,59	6,99	8,06
6.	Ergo	402,78	351,56	375,33	3,97	6,15	8,16
7.	Generali	387,25	391,78	434,95	6,48	8,99	10,29
8.	NNeder	2046,00	2122,51	2277,09	8,08	10,67	12,28
9.	Nordea	539,87	587,43	647,13	4,73	5,84	7,51
10.	PeKaO	292,12	237,18	242,57	4,42	6,86	8,73
11.	Pocztylion	457,38	362,20	356,84	3,77	6,46	8,31
12.	Polsat	261,27	225,58	249,78	3,19	3,82	4,63
13.	PZU	1902,44	1773,89	1846,40	5,55	7,61	9,47
14.	Skarbiec	601,27	491,57	453,55	4,44	6,30	8,14
15.	Winterthur	380,60	407,67	484,60	5,19	7,76	10,40

Source: Quarterly Bulletin KNUiFE, www.knuife.gov.pl

Table 2

Selected features of OPF (outputs) in the years 2004-06 (fixed prices 2004)

No.	*OPF*	Net assets [mln PLN (2004)]			Result of investments [mln PLN (2004)]			Accounting unit's values [PLN (2004)]		
		2004	2005	2006	2004	2005	2006	2004	2005	2006
1.	AIG	4282,39	5539,83	7400,97	32,14	36,12	45,04	17,16	18,63	21,09
2.	Allianz	1351,89	1725,17	2224,12	12,25	15,09	12,53	17,73	18,71	20,58
3.	Bankowy	1551,96	2059,37	2778,36	9,20	12,30	15,92	18,51	19,74	21,76
4.	Commercial	14071,95	17988,79	23724,58	130,13	140,62	146,72	18,39	19,88	22,35
5.	DOM	829,36	1041,70	1379,16	3,80	6,10	6,40	18,78	20,42	23,22
6.	Ergo	1028,01	1414,13	2072,57	8,05	10,63	12,92	18,25	19,72	22,05
7.	Generali	1663,26	2193,51	3007,24	13,18	15,32	16,15	18,53	20,09	22,72
8.	NNeder	11195,24	14717,85	20193,51	92,47	126,30	148,85	19,50	21,18	24,02
9.	Nordea	1571,26	2297,20	3207,21	13,71	18,33	15,32	19,11	20,44	22,71
10.	PeKaO	809,82	1048,96	1396,89	6,78	7,66	7,22	16,94	18,54	20,86
11.	Pocztylion	1052,66	1350,26	1796,58	7,10	10,56	9,78	16,99	18,27	20,96
12.	Polsat	484,98	571,25	776,96	3,85	4,92	3,30	19,87	21,06	24,07
13.	PZU	6983,78	8989,54	11992,08	50,69	61,85	64,67	18,45	19,93	22,28
14.	Skarbiec	1752,04	2021,80	2430,55	10,44	12,87	15,18	17,11	18,43	20,85
15.	Winterthur	1327,84	2141,12	3409,91	9,77	14,29	18,47	17,86	19,58	21,85

Source: Quarterly Bulletin KNUiFE, www.knuife.gov.pl

1. METODOLOGY OF THE OPF EFFICIENCY EVALUATION

To evaluate the efficiency of each OPF in a selected moment in time (t) the DEA method was used and to evaluate changes in returns of each OPF between two moments in time (t and t+1) the Malmquist productivity indexes were applied.

1.1. Effectiveness evaluation – DEA method

The DEA method allows to evaluate efficiency solely on the basis of data on values of inputs and outputs. It does not require any knowledge of function form defining the relation between the two categories.

The guidelines of the DEA model are as follows: there are n objects operating in a given branch, each of them makes use of m varied inputs in order to obtain s different outputs(effects). Additionally, it is assumed that the value of inputs and outputs (effects) is either bigger or equals zero; however, there is at least one input and one output bigger than zero.

The efficiency of an economic subject o is defined as the relation between the sum of weighted inputs and the sum of weighted outputs.

$$\theta = h_o(\mathbf{u}, \mathbf{v}) = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \tag{1}$$

where:
y_{ro} – r-th output of the object o,
x_{io} – i-th input of the object o,
u_r – weight defining the importance of the output r-th,
v_i – weight defining the importance of the input i-th,
s – number of outputs for the object o,
m – number of inputs for the object o.

Thus outputs and inputs are reduced to single values of a "synthetic output" (sum of weighted outputs) and a "synthetic input" (sum of weighted inputs) and their relation is the function of purpose that should be maximized. In the numerator of the expression (1) there is a "complete output" of the object o while the denominator includes a "complete input" of this object.

The DEA method does not require the knowledge of the weights u and v as for each object o the weights maximizing its efficiency h_o are searched for. The process of seeking maximizing value h_o would, however, lead to achieving

incomplete solution and therefore what should be done is the introduction of additional restrictions (2) and (3) thanks to which it is possible to find the best completed solution. According to the restriction (2) for each object the quotient of the "complete output" and "complete input" is to be smaller or equal to 1. While the restriction (3) is a classic boundary restriction.

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \quad for \quad j = 1,2,...,n \tag{2}$$

$$\begin{aligned} u_r &\geq 0 \quad dla \quad r = 1,2,...,s \\ v_i &\geq 0 \quad dla \quad i = 1,2,...,m \end{aligned} \tag{3}$$

Thus evaluating the efficiency related to the object *o* means solving a problem of quotient programming with the maximization function of purpose (1) and restrictions (2)-(3).

Model (1)-(3) can be transformed into a linear form by applying the transformation method of Charnes-Cooper and by making use of simple methods of linear optimization [7].

Due to the values of the transferred model (1)-(3) it is convenient to solve this problem that is dual to a particular one.

What is an advantage of a DEA method is the fact that it does not require the knowledge of function relation between expenditures and outputs. Efficiency curve is estimated on the grounds of empirical data on values of inputs and outputs (effects) in the form of segments of linear curve and thus highly recommended everywhere where it is impossible to fix the objective function relation between inputs and outputs (effects) or by finding corresponding weights.

Some economic objects for which the optimal value of the function of purpose (1) is placed in the curve of efficiency are efficient ($1/\theta = 1$), while the ones whose value lies below the curve of efficiency ($1/\theta < 1$) are as a rule inefficient, and their inefficiency amounts to ($1-\theta = 1$).

In this paper there were applied the following, outputs oriented DEA models:[*]: a model with constant returns to scale (CCR[**]), model with variable returns to scale (BCC[***]) and model with non increasing returns to scale (NIRS[****]). Each model should be solved *n* times separately for each economic object.

---

[*] The models presented here are in compliance with the optimization theory of dual models. In literature DEA are though called primary models. Such a reverse convention is commonly encountered.

[**] CCR; after **C**harnes, **C**ooper and **R**hodes, see [2, p. 3].

[***] BCC; after **B**anker, **C**harnes and **C**ooper, see [2, p. 23-47].

[****] NIRS; **N**on **I**ncreasing **R**etuns to **S**cale model.

The CCR model

$$\begin{aligned} \theta^* &= \max \theta \\ \sum_{j=1}^{n} x_{ij}\lambda_j &\le x_{io} \quad for \quad i = 1,2,...,m \\ \sum_{j=1}^{n} y_{rj}\lambda_j &\ge \theta\, y_{ro} \quad for \quad r = 1,2,...,s \\ \lambda_j &\ge 0 \quad for \quad j = 1,2,...,n \end{aligned} \tag{4}$$

The solution of the problem (4) consists in finding a maximum value θ which allows to maximize outputs in such a way so as not to exceed the inputs*. The efficiency ($1/\theta^*$) calculated on the grounds of the CCR model is called a *technical efficiency (e_crs)*. If $1/\theta^* = 1$ object *o* is efficient while if $1/\theta^* < 1$ object *o* is non-efficient.

The BCC model

$$\begin{aligned} \theta^* &= \max \theta \\ \sum_{j=1}^{n} x_{ij}\lambda_j &\le x_{io} \quad for \quad i = 1,2,...,m \\ \sum_{j=1}^{n} y_{rj}\lambda_j &\ge \theta\, y_{ro} \quad for \quad r = 1,2,...,s \\ \sum_{j=1}^{n} \lambda_j &= 1 \\ \lambda_j &\ge 0 \quad for \quad j = 1,2,...,n \end{aligned} \tag{5}$$

The measure of efficiency ($1/\theta^*$) calculated on the grounds of the BCC model is marked as *e_vrs*. It is so-called *pure technical efficiency* (*e_vrs*) which defines how many more outputs (effects) could be achieved with the same volume of inputs.

The NIRS model

$$\begin{aligned} \theta^* &= \max \theta \\ \sum_{j=1}^{n} x_{ij}\lambda_j &\le x_{io} \quad for \quad i = 1,2,...,m \\ \sum_{j=1}^{n} y_{rj}\lambda_j &\ge \theta\, y_{ro} \quad for \quad r = 1,2,...,s \\ \sum_{j=1}^{n} \lambda_j &\le 1 \\ \lambda_j &\ge 0 \quad for \quad j = 1,2,...,n \end{aligned} \tag{6}$$

* Contrary to the outlays oriented model whose purpose is to minimize inputs while maintaining constant level of outputs.

The NIRS model differs from the BCC model by a „less strict" condition concerning the factors of linear combination (λ_j).The measure of efficiency ($1/\theta^*$) calculated on the grounds of the NIRS model is denoted by *e_nirs*. It includes some information on the types of returns to scale, i.e. it answers the question whether an economic object (o) functions within increasing or decreasing returns of scale.

When solving the CCR model what is received is the information on complete technical efficiency of a given economic object, while the BCC model provides some information on pure technical efficiency, i.e. the one which considers variable returns to scale. If there is a considerable difference between the calculated values of efficiency in case of constant and variable returns of scale, then by comparing the two measures one can assume the existence of the returns of scale in a given group of objects. The measure of return to scale efficiency is defined in (7).

$$e\_s\_vrs = \frac{e\_crs}{e\_vrs} \tag{7}$$

The efficiency of scale (*e_s_vrs*),connected with scale (volume) of production informs, how many fewer inputs could be used if the volume of outputs were optimal. The efficiency of scale calculated in this way tells us nothing though about types of returns to scale, i.e. whether an object functions within increasing or decreasing returns to scale. Only when there is no statistically considerable difference between complete technical efficiency and pure technical efficiency it can be supposed that a given decision making unit is efficient as far as scale of engaged productivity factors are concerned. However, if the compared volumes are different we receive no answer regarding the productivity scale a given decision making unit operates in. In other words if $0 < e\_crs < e\_vrs < 1$ the obtained measure is smaller than 1 and a decision making unit is inefficient in respect to scale of the engaged productivity factors, however, the region a given decision making unit operates in is unknown. To define this aspect another measure of the scale efficiency is applied (8):

$$e\_s\_nirs = \frac{e\_crs}{e\_nirs} \tag{8}$$

Comparing the efficiency measure obtained in the model CCR with the efficiency measure obtained in the NIRS model allows to define the types of returns to scale (*e_s_nirs*). And thus if $e\_s\_nirs = 1$ the decision making unit operates within the region of increasing returns to scale, but if $e\_s\_nirs < 1$ the decision making unit operates in the region of decreasing returns to scale. In other words if $e\_crs = e\_nirs$ the object is in the region of increasing returns to scale, but if $e\_crs < e\_nirs$, the object is in the region of decreasing returns to scale.

1.2. Evaluation of efficiency changes in time – the Malmquist index

To evaluate the efficiency changes in time what is used is the index grounded to a large extent on Farrell's efficiency that was used for the first time by Malmquist [5]. He suggests that the levels of efficiency $F^t(x^t, y^t)$ and $F^{t+1}(x^{t+1} y^{t+1})$ should be compared in two different moments in time t and t+1. In this paper the function of distance by Shephard [3] $D^{\bullet}(x^{\bullet}, y^{\bullet})$ was used in place of the levels of efficiency $F^{\bullet}(x^{\bullet} y^{\bullet})$. The Malmquist index for the year t assumes the form of (9), and for the year t+1 it is (10). Index (9) compares efficiency of the period t+1 to the efficiency of the period t by using as a point of reference the technology of the t period. While index (10) compares efficiency of the period t+1to the efficiency of the period t by using as a point of reference the technology of the t+1 period.

$$M^t\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \frac{D^t\left(y^{t+1}, x^{t+1}\right)}{D^t\left(y^t, x^t\right)} \tag{9}$$

$$M^{t+1}\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \frac{D^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D^{t+1}\left(y^t, x^t\right)} \tag{10}$$

Measures $D^{\bullet}(x^{\bullet}, y^{\bullet})$ are technical effectivenesses $1/\theta^*$ obtained from solving the CCR model (4).The parameters of the left-hand sides of first two limits (*LHS*) and parameters of right-hand side limits (*RHS*) are changed in compliance with the rules presented below in the table 3.

Table 3

Rules for constructing the CCR model (4) in defining volumes $D^{\bullet}(x^{\bullet}, y^{\bullet})$

	$D^{\bullet}(x^{\bullet}, y^{\bullet})$	*LHS* (technology)	*RHS* (evaluated object)
1.	$D^t(x^t, y^t)$	From period t	From period t
2.	$D^{t+1}(x^{t+1} y^{t+1})$	From period t+1	From period t+1
3.	$D^t(x^{t+1} y^{t+1})$	From period t	From period t+1
4.	$D^{t+1}(x^t, y^t)$	From period t+1	From period t

In practice the formula (11) of the Malmquist index is applied which is the geometrical* of both indexes (9) and (10).

$$M^{t,t+1}\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \sqrt{\frac{D^t\left(y^{t+1}, x^{t+1}\right)}{D^t\left(y^t, x^t\right)} \cdot \frac{D^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D^{t+1}\left(y^t, x^t\right)}} \tag{11}$$

* Suggested by R. Färe, S. Grasskopf, B. Lindgren, P. Ross in [2, chapter 13].

After the transformations the Malmquist index can be presented in the form of a ratio (12) dividing the Malmquist index (11) into two terms (13) and (14).

$$M^{t,t+1}\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = TE\left(x^t, y^t, x^{t+1}, y^{t+1}\right) \times TC\left(x^t, y^t, x^{t+1}, y^{t+1}\right) \quad (12)$$

The first term (13) stands for the *change in technical efficiency* which defines a relative change in efficiency of a given object between two periods t and $t + 1$ but without including any changes in the curve of efficiency (as efficiency is measured in respect to a curve from a proper time period either t or $t + 1$).

The other term (14) stands for the *technical change* (connected with technological progress), which defines relative change in technology (presented in the change in curve of efficiency), measured separately in relation to the technologies from two different periods of time, i.e. the efficiency of a given object in the period t is measured with respect to the technology of the period $t + 1$ and the efficiency of an object in the period $t + 1$ is measured with respect to the technology of the period t.

$$TE\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \frac{D^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D^t\left(y^t, x^t\right)} \quad (13)$$

$$TC\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \sqrt{\frac{D^t\left(y^{t+1}, x^{t+1}\right)}{D^{t+1}\left(y^{t+1}, x^{t+1}\right)} \cdot \frac{D^t\left(y^t, x^t\right)}{D^{t+1}\left(y^t, x^t\right)}} \quad (14)$$

The coefficient of the change in efficiency of an object is the result of the Malmquist index calculation. It is assumed that for the value of the index bigger than 1 in the period of time in question, a relative increase in efficiency took place, while the value smaller than 1 means decrease in efficiency, finally the value that equals 1 means maintaining the same level of efficiency.

1.3. Evaluation of efficiency in time - other indexes of dynamics

In the DEA analysis, apart from the Malmquist indexes, there are also used indexes of dynamics of efficiency based on the models taking into account variable returns to scale (the BCC models). To differentiate the measures of Shephard obtained by the solution of the CCR models (constant returns to scale) and BCC models (variable returns to scale) we define the distances in the following way:

- $D^{\bullet}_{CRS}\left(x^{\bullet}, y^{\bullet}\right)$ for the CCR models (4) and
- $D^{\bullet}_{VRS}\left(x^{\bullet}, y^{\bullet}\right)$ for the BCC models (5). We define and calculate these measures in the way described in 2.2. We apply here the BCC models (instead of CCR).

Here are the signaled indexes:
- *PTE* – the index of *change of clear technical efficiency* and
- *SE* – the index of *change of scale of efficiency*.

$$PTE\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \frac{D_{VRS}^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D_{VRS}^{t}\left(y^t, x^t\right)} \tag{15}$$

$$SE\left(x^t, y^t, x^{t+1}, y^{t+1}\right) = \frac{D_{CRS}^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D_{CRS}^{t}\left(y^t, x^t\right)} \Big/ \frac{D_{VRS}^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D_{VRS}^{t}\left(y^t, x^t\right)} \tag{16}$$

3. OPF EFFICIENCY EVALUATION (2004-06)

The results of DEA calculations for the years 2004-2006 are presented in the tables 4-6. With 15 OPFs it was required the solution* 3×3×15+2x2×15=195 of linear optimization models (each year – by 15 models: CCR, BCC and NIRS and additionally by 2x2x15 of modified CCR while calculating Shephard measures).

The efficiency based on the CCR model is called a technical efficiency (*e_crs*) or in other words complete efficiency. The value of the coefficient $1/\theta^*$ (*e_crs*) is in between $\langle 0,1 \rangle$. While the value $1/\theta^* = 1$ says that a given OPF is efficient and that means complete transformation of outlays into returns. Whereas in case of all OPFs where $1/\theta^* < 1$ (efficiency index is <1) it means that they function inefficiently against the others. In other words, with no increased outlays they should improve their returns by $(1 - 1/\theta^*)\times 100\%$.

What can be concluded from the calculations included in the table 4 is the fact that in 2004 the Allianz, Commercial, DOM, Nationale Nederlanden, Polsat and the PZU funds represented the optimal (model) efficiency. In the year 2004 the Generali could have increased its returns by 4,3% and the Bankowy by 7,3%. As far as the aspect of efficiency is concerned other funds functioned much worse. The increase of their efficiency in relation to the outlays incurred should have amounted from 9,7% (in case of Wintherthur) to 25,1% (in case of Skarbiec). In the years 2005-06 the Allianz, Commercial, DOM, Nationale Nederlanden and Polsat represented also optimal efficiency. The remaining funds were marked with greater inefficiency. The worst results were achieved by Nordea whose increase in returns should have amounted to almost 25,5% in the first quarter of 2006 (in 2005 by 23,3%).). PZU (the third one as far as the value of its assets is concerned) should have increased its returns by 12,9% (in 2005)

* To solve the CCR models, BCC and NIRS Solver tools of Excel were used.

and by 13,1% (in 2006) at the same assets. At this point it should be also mentioned that in the years 2003-2006 Polsat achieved the highest rate of return (60%) right before the DOM Fund (59,4%).

As a result of the BCC model solution there was achieved a so-called pure technical efficiency (*e_vrs*). The relation of two efficiencies (*e_s_vrs*), i.e. complete one to pure one allows to evaluate whether a given unit is within the scope of the returns to scale. The index *e_s_vrs*=1 says about constant benefits of scale. When the relation between the complete technical efficiency and pure technical efficiency is<1 the funds are within the variable efficiency of scale. As far as the researched OPFs are concerned constant benefits of scale were noted in case of Allianz, Commercial, DOM, Nationale Nederlanden, PeKaO, Polsat and PZU in 2004-06.

To decide whether in case of variable returns to scale the type is either increasing or decreasing the NIRS model may be applied. For the Allianz, Commercial, DOM, Nationale Nederlanden, Polsat and the PZU in the years 2004-06 (PZU only in 2004) funds the index of efficiency type *e_s_nirs* calculated as the relation of complete efficiency to the efficiency calculated by the use of NIRS model equals one. It means that the funds mentioned above operated in the period in question in the area of growing returns to scale. The remaining funds operated in the area of decreasing returns to scale in the years 2004-06 (PZU in 2005-06).

To compare the changes in efficiency of OPF in time the Malmquist productivity index was applied. The comparison was carried out for some successive years of the period 2004-06. The determination of Malmquist index per each OPF required prior calculation of so-called Shaphard measurements (compare table 5).

When comparing so-called Malmquist indexes (the relation of returns to outlays) in different periods of time ($M_{t,t+1}$) we can see the increase in efficiency most of all pension funds (compare table 6). Considering the modified Malmquist index being a geometrical average of the index for the year t and t+1, we can see that the increase in efficiency most of all results from the changes in the so-called technological efficiency ($TC_{t,t+1}$) that takes into consideration the change of the curve of efficiency location. The technical efficiency, measuring the change in relative efficiency in the periods 2004-05 and 2005-06 ($TE_{t,t+1}$) is constans or decreased in the majority of funds.

Table 4

Measures of OPF efficiency in the years 2004-06 obtained by means of DEA

l.p.	*OPF*	*e_crs*			*e_vrs*			*e_nirs*			*e_s_vrs*			*e_s_nirs*		
		2004	2005	2006	2004	2005	2006	2004	2005	2006	2004	2005	2006	2004	2005	2006
1.	AIG	0,865	0,865	0,889	0,883	0,892	0,898	0,883	0,892	0,898	0,980	0,970	0,990	0,980	0,970	0,990
2.	Allianz	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**
3.	Bankowy	0,928	0,897	0,890	0,964	0,952	0,930	0,964	0,952	0,930	0,963	0,942	0,957	0,963	0,942	0,957
4.	Commercial	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**
5.	DOM	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**
6.	Ergo	0,831	0,864	0,917	0,919	0,936	0,935	0,919	0,936	0,935	0,904	0,923	0,981	0,904	0,923	0,981
7.	Generali	0,957	0,935	0,943	0,985	0,981	0,979	0,985	0,981	0,979	0,972	0,953	0,963	0,972	0,953	0,963
8.	NNeder	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**
9.	Nordea	0,801	0,767	0,745	0,963	0,970	0,943	0,963	0,970	0,943	0,832	0,791	0,790	0,832	0,791	0,790
10.	PeKaO	0,890	0,938	**1,000**	0,913	0,973	**1,000**	0,890	0,938	**1,000**	0,975	0,964	**1,000**	**1,000**	**1,000**	**1,000**
11.	Pocztylion	0,791	0,788	0,833	0,856	0,867	0,884	0,856	0,867	0,884	0,924	0,909	0,942	0,924	0,909	0,942
12.	Polsat	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**	**1,000**
13.	PZU	**1,000**	0,871	0,869	**1,000**	0,955	0,927	**1,000**	0,955	0,927	**1,000**	0,912	0,937	**1,000**	0,912	0,937
14.	Skarbiec	0,749	0,779	0,818	0,864	0,874	0,870	0,864	0,874	0,870	0,867	0,891	0,940	0,867	0,891	0,940
15.	Winterthur	0,903	0,914	0,922	0,925	0,950	0,949	0,925	0,950	0,949	0,976	0,962	0,972	0,976	0,962	0,972

To obtain better evaluation of the situation of the pension fund market the efficiency *e_crs* was compared with the yield index. To do so the shown in picture 1 was divided into four parts by means of straight lines corresponding to the average level of efficiency and average level of yield. Thus the created areas are called as follows: "stars", "sleeping ones", "question marks" and "poor dogs". The correlogram in this form is called a x BCG matrix[*]. The pension funds that are within the area of „stars" belong to well-managed funds (very good financial strategy and level of productivity). "The stars" of the year 2006 are Nationale Nederlanden, Commercial, DOM, Generali, PeKaO, and Allianz. "The sleeping ones" are Pocztylion, PZU, AIG and Bankowy, which despite the high yield could use their potentials better (too little efficiency). "The question marks" (Ergo, Skarbiec, Nordea) have quite a chance to increase their efficiency at the proper strategy of development being applied. "The poor dogs" are in the worst situation as their chances to improve the yield despite great efficiency are slight. In 2006 Winterthur and Polsat faced the same situation.

[*] The name *BCG matrix* comes from the name of the firm *Boston Consulting Group*, which was the first one to propose such a division of a correlogram at the evaluation of economic subjects. Quote from [1].

Table 5

Changes in OPF efficiency 2004-06 – distance measures according to Shephard

l.p.	*OPF*	$D_t(x_t,y_t)$		$D_t(x_{t+1},y_{t+1})$		$D_{t+1}(x_{t+1},y_{t+1})$		$D_{t+1}(x_t,y_t)$	
		2005/04	2006/05	2005/04	2006/05	2005/04	2006/05	2005/04	2006/05
1.	AIG	0,865	0,865	1,057	1,093	0,865	0,889	0,714	0,716
2.	Allianz	1,000	1,000	1,170	1,171	1,000	1,000	0,975	1,179
3.	Bankowy	0,928	0,897	1,013	1,012	0,897	0,890	0,802	0,801
4.	Commercial	1,000	1,000	1,287	1,311	1,000	1,000	1,088	1,181
5.	DOM	1,000	1,000	1,172	1,073	1,000	1,000	0,864	1,022
6.	Ergo	0,831	0,864	0,983	0,988	0,864	0,917	0,882	0,870
7.	Generali	0,957	0,935	1,033	1,073	0,935	0,943	0,842	0,904
8.	NNeder	1,000	1,000	1,269	1,263	1,000	1,000	0,935	0,990
9.	Nordea	0,801	0,767	0,911	0,871	0,767	0,745	0,822	0,813
10.	PeKaO	0,890	0,938	1,045	1,072	0,938	1,000	0,771	1,005
11.	Pocztylion	0,791	0,788	0,892	0,921	0,788	0,833	0,873	0,805
12.	Polsat	1,000	1,000	1,227	1,062	1,000	1,000	1,130	1,094
13.	PZU	1,000	0,871	0,947	0,949	0,871	0,869	1,019	0,920
14.	Skarbiec	0,749	0,779	0,909	0,923	0,779	0,818	0,803	0,684
15.	Winterthur	0,903	0,914	1,030	1,063	0,914	0,922	0,777	0,850

Table 6

Changes in OPF efficiency 2004-06 – Malmquist and other indexes

l.p.	*OPF*	$M_{t,t+1}$		$TE_{t,t+1}$		$TC_{t,t+1}$		$PTE_{t,t+1}$		$SE_{t,t+1}$	
		2005/04	*2006/05*	*2005/04*	*2006/05*	*2005/04*	*2006/05*	*2005/04*	*2006/05*	*2005/04*	*2006/05*
1.	AIG	**1,217**	**1,253**	1,000	1,028	1,217	1,219	1,010	1,007	0,990	1,021
2.	Allianz	**1,095**	**0,997**	1,000	1,000	1,095	0,997	1,000	1,000	1,000	1,000
3.	Bankowy	**1,105**	**1,119**	0,967	0,992	1,143	1,128	0,988	0,977	0,978	1,016
4.	Commercial	**1,088**	**1,054**	1,000	1,000	1,088	1,054	1,000	1,000	1,000	1,000
5.	DOM	**1,165**	**1,025**	1,000	1,000	1,165	1,025	1,000	1,000	1,000	1,000
6.	Ergo	**1,076**	**1,097**	1,040	1,061	1,035	1,034	1,018	0,999	1,021	1,062
7.	Generali	**1,095**	**1,095**	0,977	1,009	1,121	1,085	0,996	0,998	0,981	1,011
8.	NNeder	**1,165**	**1,129**	1,000	1,000	1,165	1,129	1,000	1,000	1,000	1,000
9.	Nordea	**1,031**	**1,020**	0,958	0,971	1,076	1,050	1,007	0,972	0,951	0,999
10.	PeKaO	**1,195**	**1,066**	1,054	1,066	1,134	1,000	1,066	1,028	0,989	1,037
11.	Pocztylion	**1,009**	**1,099**	0,996	1,057	1,013	1,040	1,013	1,020	0,983	1,036
12.	Polsat	**1,042**	**0,985**	1,000	1,000	1,042	0,985	1,000	1,000	1,000	1,000
13.	PZU	**0,900**	**1,015**	0,871	0,998	1,033	1,017	0,955	0,971	0,912	1,028
14.	Skarbiec	**1,085**	**1,191**	1,040	1,050	1,043	1,134	1,012	0,995	1,028	1,055
15.	Winterthur	**1,158**	**1,123**	1,012	1,009	1,144	1,113	1,027	0,999	0,986	1,010

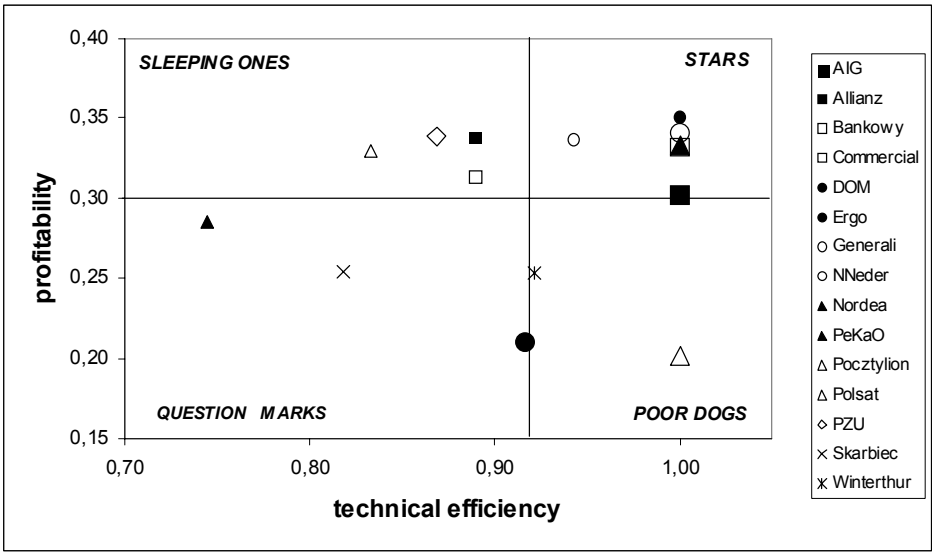


Fig. 1. OPF in the year 2006 – profitability and efficiency (BCG matrix)

SUMMARY

The findings can be summarized as follows:

1. In the period of 2004-06 almost all the funds were inefficient in case of assumed both constant and variable returns to scale (exluding Allianz, Commercial, Dom, Nationale Nederlanden, Polsat founds).
2. In the period of 2004-06 almost all the funds operated within the area of increasing returns to scale (exluding Allianz, Commercial, Dom, Nationale Nederlanden, Polsat founds).
3. In the years 2004-06 the efficiency of almost all pension funds increased period by period (growth in efficiency by 1,5-25,3%).
4. The executed research on efficiency and its changes in time confirms that two funds remain the leaders in the market of pension funds and they are: Commercial Union and Nationale Nederlanden.

REFERENCES

1. Białas P.: Zastosowanie metody DEA do analizy i oceny banków. M.A. Thesis. Faculty of Economics and Sociology, Lodz Univeristy, Łódź 2006.
2. Charnes A., Cooper W.W., Lewin A.Y., Seiford L.M.: Data Envelopment Analysis: Theory, Methodology and Applications. Kluwer Publishing, Boston 1995.

3. D. Caves D., Christiansen L., Diewert E.W. : The economic Theory of Index Numbers and the Measurement of Input, Output and Productivity. "Econometrica" 1982, Vol. 50.
4. Gospodarowicz M.: Procedury analizy i oceny banków. Materiały i Studia Zeszyt no. 103, Warsaw 2000.
5. Malmquist S. : Index Numbers and Indifferences Surfaces. "Trabajos de Estatistica" 1953, Vol. 4.
6. Miszczyńska D.: Wielokryteriowy ranking Otwartych Funduszy Emerytalnych, Modelowanie preferencji a ryzyko'01. AE, Katowice 2002.
7. Miszczyńska D.: Wielokryteriowy ranking Otwartych Funduszy Emerytalnych metodami AHP i PROMETHEE. Forecasting of Financial Markets, Lodz University, Lodz 2003.
8. Miszczyńska D.: Ranking Otwartych Funduszy Emerytalnych - Prognozy 2004-2005. Forecasting of Financial Markets. Lodz University, Lodz 2004.
9. Miszczyński M.: Programowanie liniowe, series: Metody badań operacyjnych. Absolwent, Łódź 1997.
10. Rogowski G.: Metody analizy i oceny banku na potrzeby zarządzania strategicznego. Scientific Works of Wrocław Economy Acadamy, Wroclaw 1996.
11. Websites (2004), (2005), (2006): www.money.pl, www.knuife.gov.pl

Andrzej M.J. Skulimowski

ON MULTICRITERIA PROBLEMS WITH MODIFICATION OF ATTRIBUTES*

Abstract

In this paper we propose a mathematical model for multicriteria decision problems with alternatives which may change their properties in a direct response to external actions. We assume that the change of attributes may be controlled by the decision-maker taking into account that an improvement of the criteria values bears certain cost. Thus we get a bi-level multicriteria optimisation problem: an optimal allocation of resources at the lower level, and finding the related nondominated outputs surpassing a reference point q at the higher level. A concrete problem of this type, motivated by technological, ecological and socio-economical applications, will be discussed in more detail, namely optimising the structure of a finite population X by assuring that after a fixed time T a maximal number of its elements is characterised by nondominated values of criteria. Assuming that X consists of N elements, the solution to this problem is equivalent to solving in parallel N discrete dynamic programming problems sharing the same resources.

Keywords

Multicriteria optimisation, decision theory, dynamic programming, discrete-event systems, discrete-time control systems.

INTRODUCTION

Real-life decision-making is a dynamic process, even if time is not expressed explicitly in the usually simplified mathematical problem formulation.

Applying standard static methods of multicriteria decision-making one assumes that the alternatives are characterised by fixed attributes, whereas the main difficulty consists in finding and accepting the nondominated compromise values. In dynamical multicriteria decision models based on optimal control the evolution of criteria values is described over certain interval of time, however, usually only the values of criteria evaluated *a posteriori* at the end of the control

* Research part-financed by the Polish Ministry of Science and Higher Education, carried out within the Multi-Year Programme "Development of innovativeness systems of manufacturing and maintenance 2004-2008". Contract No. PW-004/01/2006/5/UW-2006.

period are taken into account for the decision-making purposes. At last, some dynamical programming and trajectory optimisation models allow to consider the intermediate criteria values, but the control principle consists rather in switching between alternatives than in changing their properties.

Therefore there exists a need for appropriate mathematical models for the decision problems with alternatives which may change their properties as a direct response to the external actions. As an example of such problem may serve e.g. the situation where the crew is to be completed from among a set of candidates based on several criteria related to the knowledge, abilities etc., and the selection committee considers for each non-perfect candidate a possibility of investing some amount for the additional education, internal training, etc. to achieve the desired virtues after a period of time. To apply a quantitative analysis method, in the above example one has to evaluate the estimates of cost and duration of the additional training, as well as to elaborate a model of evolution of the attributes. Another type of problems, which may be treated within the framework here proposed are investment problems, where the initial decision determines the scope of future actions. An example of a problem of this kind, referring to the choice of the computer system, is shown in Sec.3.

In this paper we propose a family of such models which may be regarded as a generalisation of discrete choice methods to the situations where the time evolution plays an essential role. The changes of the values of criteria may occur spontaneously, as well as they may be systematically influenced by the decision-maker. It will also be assumed that the decision-maker's actions resulting in a desired improvement of the criteria values may bear certain cost and may not be immediate, i.e. the lapse of time necessary to realize the desired change may be considered as an auxiliary criterion. Thus we get a bi-level multicriteria optimisation problem consisting of the optimal allocation of resources at a lower level, and selecting the related Pareto optimal outputs of the original problem in a minimal time at the higher-level. An important feature of this approach is a proper description of possible transitions between the attributes of each criterion, which will be accomplished by introducing in Sec.1 so called transition patterns.

To make the presentation of the above ideas maximal comprehensible, the scope of this paper will be confined to the problems with a finite alternative set Ω, and the performance criterion $F = (F_1,...,F_N)$ admitting values from a partially ordered finite set V.

Let $V_i := \{v_{i,1}, v_{i,2}, ..., v_{i,c(i)}\}$ denote the set of values of the criterion F_i ordered from the least to the most preferred one by an order "$\prec_I$". Then V is the Cartesian product of V_i, i.e. $V = V_1 * ... * V_N$ with the coordinate-wise partial or-

der "$\prec$". The changes of values of F for a fixed $\omega \in \Omega$ are results of the control actions $u(t) \in U$ for t from a discrete time interval $[t_0,T]$:

$$\begin{gathered}((F_1(\tau),\ldots,F_N(\tau)): \Omega \to V) \to opt, \text{ for a } \tau \in [t_0,T] \\ F(t+1)(\omega) = \varphi(F(t)(\omega),u(t),t) \text{ for each } \omega \in \Omega,\ t \in [t,T-1]\end{gathered} \tag{1}$$

The optimality principle "opt" in (1) should model adequately the decision situation concerned. As a representative example, in Sec. 2 we will investigate a reference point problem which may be formulated as:

"$(F_1(\tau),\ldots,F_N(\tau))(\omega)$ is Pareto-optimal in $F(\tau)(\Omega)$ and exceeds given $q \in V$ for the minimal $\tau \in [t_0,T]$ and at a minimal cost of control".

The triple (Ω,F,ϕ) will be called the *decision process*.

A control u(t) will be identified with a transition $(v_i \to v_j)$ on the time interval (t,t+1], which is the result of an external action undertaken by the supervisor of the decision process. Besides of controlled transitions we will distinguish the deterministic uncontrolled ones which may not be influenced, as e.g. passing to the following age classes, random transitions occurring spontaneously, and non-admissible transitions. The classes of controllable and random transitions need not be disjoint, although in this paper we will be concerned with deterministic control only. Thus, the evolution of attributes may be modelled in a manner similar to the discrete-event systems described in [8], whereby the values of F play the role of system's states.

Since the set of alternatives Ω, the set of values of criteria V, and time are all discrete, such decision processes will be called a *D-D-D-system*. The particular relevance of D-D-D-systems consists in the fact that they constitute a natural extension of the discrete choice and outranking methods. Moreover, there exist close relations to the multicriteria optimal stopping problem, and to the multicriteria problems with variable constraints described in [7; 2]. In the final section we will provide several illustrative examples and point out the further research problems related to D-D-D-systems.

After a suitable discretisation of the criteria values, the solution methods here presented may also be used for the discrete-continuous processes (D-D-C), where merely the criteria values are arbitrary real numbers. The analysis of continuous processes (D-C-C or C-C-C) can in most cases be accomplished within the framework of the multicriteria optimal control problems with the criteria included in the state-space vector.

1. EVOLUTION OF ATTRIBUTES: THE TRANSITION AND COST PATTERNS

Even under the above simplified assumptions the number N of the admissible criteria values may be very large. However, the task of supplying all necessary information concerning the transitions between the values of F might be considerably simplified if it were possible to find a convenient description of the transfer function ϕ, and to identify the non-admissible transitions before starting the numerical solution process. A further reason for introducing here the transition and cost patterns is to reduce the computational complexity of the general problem by decomposing it into several subproblems, each one of them referring to the single criterion F_i, i = 1,..N. This would be possible if the characterisation of transitions between the values of $(F_1,...F_N)$, including their admissibility, could be derived from the properties of the single criteria, $F_1(t)$ through $F_N(t)$, considered separately. Below we will show that this goal can be achieved under some additional assumptions concerning the set Ω and the criterion F.

Let us fix the moment of time $t \in [t_0,T]$ and let $V_i := \{v_{i,1},v_{i,2},...v_{i,(c(i))}\}$ denote the set of values of the criterion F_i ordered from the least to the most preferred one. If we know which transitions between the values of the criterion F for an $\omega \in \Omega$ are at all possible on the time interval (t,t+1], we could define for F_i and ω the *transition pattern* as a quadratic 0-1 matrix

$$P(F_i)(\omega) = [p^i_{jk}(\omega)]$$

with the following coefficients:

$$p^i_{jk}(\omega) = \{1 \text{ iff } \omega \in \Omega \text{ may change its classification in one time step from j-th to the k-th attribute of } F_i$$
$$0 \text{ otherwise}\} \qquad (2)$$
$$\text{for } j,k=1,..c(i),\ i = 1,..N$$

Observe that the dimension of $P(F_i)(\omega)$ equals to the number of elements of V_i , c(i), and its columns indicate the admissible transitions from an appropriate fixed starting value of F. Remark that to each transition pattern $P(F_i)$ one can associate the digraph $G(F_i)$ such that $P(F_i)$ is its structural matrix. In general, the transition patterns may vary on the interval $[t_0,T]$, being thus functions of both, ω and t.

Transitions from v to w on the time interval (s,t] may be regarded as pairs (v,w) and will be denoted by v->w. By a *superposition of the transitions* $\xi_1 := v_1 \to v_2$ on the interval $(t_1,t_2]$, and $\xi_2 := v_2 \to v_3$ on the interval $(t_2, t_3]$ we will mean the transition $\xi := v_1 \to v_3$ on $(t_1,t_3]$, and denote it by $\xi = \xi_1 \circ \xi_2$.

Suppose now that $v_1 = (v_{11},v_{12})$. By the *composition of the transitions* $\xi_1 := v_{11} \to v_2$ on $(t_1,t_2]$ and $\xi_2 := v_{12} \to v_3$ on the same interval $(t_1,t_2]$, we will mean the transition $v_1 \to (v_2 ,v_3)$ on $(t_1, t_2]$. To denote compositions we will use the notation ξ_1 c ξ_2.

Let

$$\xi_1 := (v_{11} ,v_{12}) \to (v_2, v_{12}) \text{ and } \xi_2 := (v_{11} ,v_{12}) \to (v_{11} ,v_3)$$

Then, formally, $\xi_1 \circ \xi_2 = \xi_1$ c ξ_2, the diversity between composition and superposition being expressed by the associated time intervals. Observe that the superpositions and compositions describe sequential and parallel processing of transitions, respectively. By definition, the compositions are always admissible in one time step, the superpositions may, but need not necessarily have this property. To assure a minimal number of non-zero coefficients in $P(F)(\omega,t)$, it is convenient to include in the transition patterns only the transitions which may not be represented as compositions of other admissible transitions.

Transitions lasting several time steps may often be represented as superposition of one-step transitions. If it is not so, they can still be considered within the same framework by introducing the intermediate or wait values of F. This question will be considered in more detail further in this section.

Now we will introduce several properties of the decision process (Ω,F,ϕ), which will be used in the further analysis of the initial decision-making problem.

Definition 1.1. We will say that the set of alternatives Ω is *homogeneous* with respect to F at the moment $t \in [t_0 ,T]$, iff

$$\forall 1 \le i \le N \ \forall x,y \in \Omega : P(F_i)(x) = P(F_j)(y) \qquad (3)$$

If (3) is satisfied for all $t \in [t_0 ,T]$, we will call Ω homogeneous.

If Ω is not homogeneous but handling a separate transition pattern for all alternatives would be computationally inefficient then one may consider instead the Boolean product of transition patterns for all $x \in \Omega$. Generally, in models of real-life discrete dynamical systems the transition patterns depend on discretisation of time, which should be suitably chosen. Moreover, as we already noted, they may depend on time itself. In the sequel we will usually admit the assumption that the decision process (Ω,F,ϕ) is stationary, according to the following definition:

Definition 1.2. If for each $\omega \in \Omega$ and $1 \le i \le N$, $P(F_{\cdot i})(\omega)$ remains constant on the whole interval $[t_0, T]$ then the decision process (Ω, F, ϕ) will be called *stationary*.

Observe that the stationarity assumption is equivalent to the fact that the function φ from (1) does not depend on time t.

Another important set of properties concerns the independence of criteria $F_1, \ldots, F_N$.

Definition 1.3. The criteria $F_1, \ldots F_N$ are *evolution-independent* at $x \in \Omega$ and $t \in [t_0, T]$, by definition it means that any transition

$$F(t)(x) := (v_{1,i(1)}, \ldots v_{N,i(N)}) \to (v_{1,j(1)}, \ldots v_{N,j(N)}) = F(t+1)(x)$$

is admissible iff for each $1 \le k \le N$ the transitions $v_{k,i(k)} \to v_{k,j(k)}$ are admissible, i.e. iff $p^{k}_{i(k),j(k)}(x)(t) = 1$.

The criteria $F_1, \ldots F_N$ will be called globally evolution independent, or simply evolution-independent iff the above holds for all $x \in \Omega$ and $t \in [t_0, T]$. Roughly speaking, the criteria are evolution independent iff the admissibility of transitions between the values of F_k, $1 \le k \le N$, is not affected by the present values of all remaining criteria.

It is easy to observe that the following fact is true :

Proposition 1.1. If the decision process (Ω, F, ϕ) is homogeneous and stationary, and for an $x_0 \in \Omega$, $t \in [t_0, T]$, $F_1, \ldots F_N$ are evolution independent at (x_0, t) then $F_1, \ldots F_N$ are globally evolution independent.

In the sequel we will always assume that the criteria concerned are evolution independent.

Example 1.1. Suppose that Ω is the population of citizens of a city and one of the objectives s is the age scale with the attributes $v_1 = [0,20]$, $v_2 = (20,40]$, $v_3 = (40,60]$, and $v_4 = (60,\infty)$, denoting the age in years of a single individual. If all time discretisation steps in the process (1) are less than 20 years, which is usually the case, then the transition pattern $P(s) = [p_{jk}]$ is the matrix

$$P(s) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

No transitions are controllable, unless we dispose a relativistic vehicle to force remaining within the same age class (elements on the main diagonal would then correspond to controllable transitions).

An important feature of the transition patterns for evolution independent criteria consists in the fact that it is sufficient to determine the patterns for the single criteria only, while the transition patterns for the vector criterion may be calculated basing on the following

Proposition 1.2. Assume that the evolution independent optimisation criteria F_i and F_j are defined on a homogeneous population Ω with the transition patterns

$$P_i = [p^{(i)}_{k,l}] := P(F_i)(t) \in M_{c(i),c(i)} \text{ and } P_j = [p^{(j)}_{m,n}] := P(F_j)(t)\ M_{c(j),c(j)}$$

for certain fixed $t \in [t_0,T]$, respectively. Then the transition pattern $P_{ij} := P(F_i, F_j)(t)$ of the vector criterion $F_{ij} := (F_i, F_j)$, is the block matrix

$$P_{ij} = [p^{(i)}_{k,l}\ P_j]^{c(i)}_{k,l=1} \in M_{c(i)c(j),c(i)c(j)} \tag{4}$$

(the block product of P_i and P_j), where c(i) and c(j) denote the number of admissible values of the criteria F_i and F_j, respectively. The values of F_{ij}, $v_{kl} = (v^{(i)}_k, v^{(j)}_l)$, labelling the rows and columns in P_{ij}, are ordered lexicographically with the first coordinate more relevant than the second.

Proof: Suppose first that the transitions $(v_{i,k(i)} \to v_{i,l(i)})$ and $(v_{j,k(j)} \to v_{j,l(j)})$ are both admissible, i.e. $p^{(i)}_{k(i),l(i)} = 1$ and $p^{(j)}_{k(j),l(j)} = 1$. Then from the evolution independency assumption (Def. 1.3) it follows that the transition

$$\xi := (v_{i,k(i)}, v_{j,k(j)}) \to (v_{i,l(i)}, v_{j,l(j)})$$

is admissible. According to the construction of P_{ij} (cf. (4)), the element corresponding to ξ in P_{ij}, $p^{(ij)}_{m,n}$ with m = k(i)c(j)+k(j) and n = l(i)c(j)+l(j), is the (k(j),l(j))-th coefficient of the block $p^{(i)}_{k(i),l(i)}\ P_j$, which is equal to 1 since $p^{(i)}_{k(i),l(i)} = 1$, and $p^{(i)}_{k(i),l(i)} P_j = P_j$.

If at least one from the above simple transitions is non-admissible then $p^{(i)}_{k(i),l(i)} = 0$ or $p^{(j)}_{k(j),l(j)} = 0$. In the first case the block $p^{(i)}_{k(i),l(i)} P_j$ of P_{ij} contains only zero elements, in the second, its (k(j),l(j))-th coefficient, $p^{(ij)}_{m,n}$, is equal to zero. However, from the definition of the evolution independent criteria it follows that any transition between values of (F_i, F_j) must be represented as a su-

perposition of simple admissible transitions cannot be admissible, therefore the transition ξ corresponding to the zero coefficient $p^{(ij)}_{m,n}$ is not admissible.

Now, let us fix a coefficient $p^{(ij)}_{m,n}$ of P_{ij}. Then there exist k1(m,n),l1(m,n)∈[1,c(i)] and k2(m,n),l2(m,n)∈[1,c(j)] such that $p^{(ij)}_{mn}$ is the (k2(m,n),l2(m,n))-th coefficient of the block $p^{(i)}_{k1(m,n),l1(m,n)} P_j$ of P_{ij} , i.e. to $p^{(ij)}_{m,n}$ there can be associated the transition

$$(v_{i,k1(m,n)} , v_{j,k2(m,n)}) \rightarrow (v_{i,l1(m,n)} , v_{i,l2(m,n)})$$

and $p^{(ij)}_{m,n}$ determines its admissibility, as shown in the first part of the proof. Hence we conclude that the above characterisation of the of P_{ij} as transition pattern (2) for (F_i , F_j) is complete.

Corollary 1.1. If the transition patterns of two evolution independent criteria s and q, P(s) and P(q), have p(s) and p(q) non-zero elements, respectively, then the transition pattern for (s,q), P(s,q) contains at most p(s)p(q) non-zero elements.

Corollary 1.2. Consequently, the transition patterns for any finite number of evolution independent criteria, $F_1,\ldots,F_n$, are sparse block matrices which can be constructed making recursively use of Proposition 1.2.

1.1. Assignment of controls and costs to admissible transitions

Let us fix w∈Ω, ω∈Ω, and t∈ $[t_0,T]$, and assume that the coefficients of the transition pattern P(F)(ω,t) are ordered lexicographically. Then to each admissible transition v→w we can associate a control $u_m(t)$ and its cost J(v,w, ω ,t) := $J(u_m, \omega ,t)$, where the integer m:= m(v,w) is the ordinal number of the appropriate control from the list U:= $\{u_1,...,u_M\}$ responsible for the change from v to w.

If J(v,w, ω ,t) does not depend on the past transitions then, analogously to the transition patterns, for each ω∈Ω, t∈ $[t_0 ,T]$, and each criterion F_i, one can define the cost pattern $J(F_i)(\omega,t)$, as a function associating to each transition $v_{i,k}$ →$v_{i,l}$ that is feasible between t and t+1, the cost of applying the control u_m := u(i,k,l, ω ,t) that causes the change from $v_{i,k}$ to $v_{i,l}$,

$$J(F_i , v_{i,k} , v_{i,l} , \omega ,t) := J_i(u_m , \omega ,t)$$

Hence, the cost pattern for F_i can be represented as the c(i)*c(i) real matrix defined as follows:

$$J^{i}_{kl}\ (\omega,t) := \{J_i(u(i,k,l,\ \omega\ ,t),\ \omega,t) \text{ iff } v_{i,k} \text{ is admissible} \\ \infty - \text{otherwise} \tag{6}$$

Consequently, the cost pattern for F_1 ,...,F_N , $J(F_1$,..F_N $)(\omega,t)$, is the c(1)...c(N) x c(1)...c(N) real matrix storing the costs of transitions between the values of the vector criterion F:= (F_1,...F_N), i.e. $J_{i(1),...,i(N),\ j(1),...,j(N)}(t,\ \omega)$ is the cost of changing the value ($v_{1,i(1)}$,...,$v_{N,i(N)}$) of F to ($v_{1,j(1)}$,...$v_{N,j(N)}$), or it is undefined iff such transition is non-admissible. Hence it follows that the structure of the cost patterns is closely related to the transitions patterns whereby only those coefficients of J which correspond to a "1" in P are finite. Thus, in a machine implementation of the above decision process, the transition patterns may serve as addresses of those elements of $J(F_1$,..F_N) which has to be stored in the memory. Moreover, observe that the zero coefficients of J correspond usually to non-controllable transitions.

For the evolution-independent criteria F_1 ,...F_N an important role is played by the following condition:

Definition 1.4. The cost function $J(F_1$,..F_N) satisfies the *cost-additivity* condition iff for any $\omega\in\Omega$, $t\in [t_0$,T], the cost of any admissible transition v->w, where v:= ($v_{1,i(1)}$,...$v_{N,i(N)}$) and w:= ($v_{1,j(1)}$,...$v_{N,j(N)}$), is the sum of changing the single criteria values, i.e.:

$$J(v,w,\ \omega\ ,t) = \sum_{k=1}^{N} J(v_{k,i(k)},v_{k,j(k)},\ \omega,t) \tag{7}$$

The above condition lets us consider each transition as a composition of simple transitions during the computation of the optimal improvement strategy for the values of F. This, in turn, allows to omit the operations on J(F), using instead $J(F_i)$, for I = 1,…,N.

1.2. Handling the transitions distributed over time

Let us start this subsection from the following definition :

Definition 1.5. A transition v:= ($v_{1,k(1)}$,...$v_{N,k(N)}$)→ ($v_{1,l(1),...}v_{N,l(N)}$), which:
(i) cannot be represented as a superposition or composition of other admissible transitions,
(ii) may be realised after θ, θ>1, time units at the soonest,
(iii) while realising v, F does not admit any other admissible values,
will be called *irreducible*.

To consider the irreducible transitions within the uniform decision model, one can proceed as follows:

Algorithm 1.1.
Repeat for all irreducible transitions v with the realisation time $\theta:=\theta(v)$:

Case 1:
If k(j) = l(j) for j∈{1,...N}\{i} then

Step 1.1. Define the wait values $v_{i,k,1}$,...$v_{i,k,}$ θ_{-1} and attach them to V_i .
Set $V_i := V_i \cup \{v_{i,k,1}$,...$v_{i,k,}$ $\theta_{-1}\}$.
Step 1.2. Order the wait values in any way, according to their real-life interpretation, but without affecting the existing partial order "$\prec_I$", i.e.,
if $v_{k(i)} \prec_i v_{l(i)}$, or $v_{l(i)} \prec_i v_{k(i)}$ then $v_{k(i)} \prec_i v_{i,k,1} \prec_i ... \prec_i v_{i,k,}$ $\theta_{-1} \prec_i v_{l(i)}$, or
$v_{l(i)} \prec_i v_{i,k,}$ $\theta_{-1} \prec_i ... \prec_i v_{i,k,1} \prec_i v_{l(i)}$, respectively.
Step 1.3. Update the transition pattern $P(F_i)$, respectively.
If n is the last irreducible transition STOP else proceed with the next v.

In particular, one proceeds in this manner in a single-criteria model with the objective F_i if a transition from $v_{j,k}$ to $v_{i,l}$ is irreducible.

Case 2:
If $v:= (v_{1,(k(1)}$,...,$v_{N,k(N)})$, $w:= (v_{1,l(1)}$,...,$v_{N,l(N)}$), $v:=(v \rightarrow w)$, and $k(i) \neq l(i)$ for at least two different values of $i \in \{1,...N\}$, then

Step 2.1. Represent the transition $v:= (v_{i,k} \rightarrow v_{j,l})$ in the form

$$v = v_{1i} \, o \ldots o \, v_{i} \rho \tag{8}$$

where v_{ij} is an admissible transition or a non-reducible simple transition, taking into account as far as possible the real-life behaviour of the system during the transition.
If the representation in form (8) is impossible go to Step 2.3.
Step 2.2. Perform for each irreducible simple transition the Steps 1.1-1.3 from the Case 1.
If v is the last irreducible transition STOP else proceed with the next v.
Step 2.3. Introduce the wait values v $v_{,1}$,...v $v_{,}$ θ_{-1} , directly as elements of V, similarly as in the Steps 1.1-1.2, but do not associate them with any V_i.
If n is the last irreducible transition STOP, else proceed with the next v.

Observe that in the case dealt with in the Step 2.3, the criteria F_1 ,...F_N are not evolution-independent and the further analysis of the problem cannot be based only on the simple transition and cost patterns $P(F_i)$.

Same intermediate values may be shared by different irreducible transitions, and one can show that a minimal set of such values may be found.

2. A NEW CLASS OF MULTICRITERIA DECISION-MAKING PROBLEMS

As an example of decision models applying the above sketched family of quantitative structures, we will discuss in more detail the following basic problem:

Problem 2.1. Find the alternative $\omega \in \Omega$ and an optimal allocation of resources $u(1),...,u(\tau)$, to achieve or surpass by the value of $F(\tau)(\omega)$, in the minimal time τ and at minimal cost, one of the reference points $q \in \theta$ defined in the space of criteria values.

In a more rigorous setting, let Ω be the finite set of admissible alternatives at time $t_0 := 1$, and F_1 ,..F_N the criteria functions defined on Ω with values in the discrete sets V_i with the partial order "$\prec_I$" for I = 1,..N. Similarly as in the previous sections denote by F the vector criterion F:= (F_1 ,..F_N), F: $\Omega \to V$, valued in the Cartesian product V:= V *...*V with the coordinatewise partial order "$\prec$". For a fixed $\omega \in \Omega$, the values of criteria on w may vary according to (1), i.e.

$$F(t+1)(\omega) = \phi(F(t)(\omega),u(t),t), \text{ for } t \in [t_0,T-1]$$

Our task consists in finding an $\omega \in \Omega$, a $\tau \in [t_0,T]$, and a sequence of controls $u(t_0),...,u(\tau)$, so that :

(i) $((F_1(t_0),..F_N(t_0))(\omega)$ is nondominated in V and

$$((F_1(\tau),..F_N(\tau))(\omega) \prec q \tag{9}$$

for certain reference point $q \in \theta$;
(ii) τ with the property (9) is minimal in $[t_0,T]$ for a fixed ω; such value will be denoted $\tau(\omega)$;
(iii) $\sum_{t0 \le t \le \tau(\omega)} J(u(t), \omega,t)$ is minimal on the set Λ defined as follows:

$$\Lambda := \bigcup_{w\in\Omega} \{w\} * \{y\in U^S : \tau(w) \text{ and } F(\tau(w))(w) \text{ satisfy (i) and (ii)}\} \qquad (10)$$

where $s := \tau(w) - t_0$, and $y := (u(t_0),...,u(\tau(w)))$

Observe that, according to (1), $F(\tau(w))(w)$ and $\tau(w)$ are indirect functions of y. The minimal value of (iii) on Λ will be denoted by $J_{min}(\omega)$.
(iv) $((F_1(\tau(\omega)),..F_N(\tau(\omega)))(\omega), \tau(\omega), J_{min}(\omega))$ is nondominated in the set $VxIR^2$ with the coordinatewise partial order.

The general Problem 2.1 consists of two tasks : finding an optimal alternative w, and a sequence of controls y assuring the achievement of q at a minimal cost. Each alternative $\omega\in\Omega$ is characterised by the minimal time $\tau(\omega)$ and the minimal cost $J_{min}(\omega)$ of achieving or surpassing q. Consequently, if for each $\omega\in\Omega$ one knows these minimal parameters, then the final choice of ω is a bicriteria trade-off between the cost and time, which can be made using one of well-known interactive decision-making methods applied for bicriteria problems.
We will present a solution to the above problem for the decision processes satisfying the following assumptions:
(i) the decision process (Ω,F,ϕ) is stationary and homogeneous;
(ii) the criteria F are evolution independent;
(iii) all transitions between the values of F are deterministic;
(iv) the costs of transitions satisfy the cost-additivity condition;
(v) the reference set $Q\subset V$ can be represented in the form $Q := \{p\in V: p\prec q\}$ for certain $q\in V$.

As the first step of the solution, below we will show how can one determine $\tau(\omega)$ and $J_{min}(\omega)$ for a fixed ω.

2.1. Solving single-object evolution problems

Let us admit all above assumptions (i)-(iv), let us fix an $\omega\in\Omega$, and let $f_0 := F(t_0)(\omega)$. Further, let us consider a directed network G=(V,E), where the nodes V can be identified with the set $V = V_1 x...V_N$ of potential values of F, while the edges $e\in E\subset V^2$ are determined by the transition patterns P_1 ,...,P_N, $P_i := P(F_i)(\omega)$, for i = 1,..N, in the following way:

$$\begin{aligned} e = (f,g) \in E \Leftrightarrow f\neq g \text{ and} \exists!\, j\in\{1,..N\}\ f_i = g_i \text{for } i\in\{1,..N\}\setminus\{j\} \text{ and} \\ P_j(f_j,g_j) = 1 \text{ or } f = g \text{ and } \forall\, i\in\{1,..N\}\ P_i(f_i,gi) = 1 \end{aligned} \qquad (11)$$

Thus, the edges of G correspond to the simple transitions between the values of F or may be loops. Additionally, the edges of G are equipped with quantitative labels describing the time θ_i and the cost of transition J_i, and qualitative labels c_i indicating whether the corresponding transition is forced or controllable. Hence, the following observation is straightforward :

Proposition 2.1. The transition between two values of criteria, f and g, is possible iff the nodes corresponding to f and g in G can be connected by a path.

As a corollary from Prop. 2.2 we get

Proposition 2.2. The graph G is the Cartesian product of graphs $G_1,...G_N$, which correspond to the single criteria $F_1,...F_N$ and their transition patterns $P_1,...P_N$, respectively. Its structural matrix is given as the block product of $P_1,...P_N$, $P_{1,...,N}$.

Hence it follows

Theorem 2.1. The solution to the Problem 2.1 for a single alternative $x \in \Omega$ can be found as a bicriteria shortest path in G between f_0 and the reference set $Q := \{v \in V : q\{v\}$.

The solution algorithm which can be derived from the above Prop. 2.1. and 2.2 and Thm. 2.1 may be presented as follows:

Algorithm 2.1.
The input data:
The transition and cost patterns for $F_1 ,...F_N$, $P_1 ,...P_N$, and $J_1 ,...J_N$, respectively. The time horizon T, the starting value f_0 , the reference point $q \in V$, the reference set Q.
Step 1. Augment the transition patterns by the time-distributed transitions, applying the procedure presented in Sec. 1.3.
Step 2. Check whether the criteria are evolution independent.
If yes :
– construct the network G using the Prop. 2.2 and 2.2;
otherwise :
– set manually all edges of G.
Step 3. Check whether it exists a path joining f_0 and q, or any other $p \in Q$
(i.e. check whether Q is attainable from f_0).
If not, return to the communication shell to let the decision-maker, define new reference point or to undertake another modification of the decision-maker's preference structure.

Step 4. Determine the set D containing all bicriteria shortest paths between f_0 and all $p \in Q$, using the bicriteria shortest path algorithm. Find the set of nondominated points of D, P(D).
Step 5. Select a compromise strategy from P(D) using any bicriteria trade-off procedure.

2.2. The selection problem from among multiple evolving alternatives

In the present setting we assume that at the moment t_0 the decision-maker should choose that alternative $\omega_0 \in \Omega$ which gives the best chances to be improved till the time T so as it were not worse than q. After simulating the evolution of $F(t)(\omega)$ over time $t \in [t_0,T]$, one chooses ω_0, which will be called prospective alternative, and starts investing in its development, by undertaking the actions $u(t_0),...u(T-1)$, without taking care what happens with all remaining alternatives. This solution procedure implies the following :

Theorem 2.2. To select the prospective alternative and the best strategy in Problem 2.1 for stationary homogeneous processes with evolution independent criteria, it is necessary to solve the simultaneous bicriteria shortest path problem for the set of starting points $V_0:=\{f \in V: f = F(t_0)(\omega)$ for all $\omega \in \Omega\}$ and Q as the set of terminal points.

As the simultaneous shortest path algorithm one can apply a combination of the well-known Dijkstra algorithm and the bicriteria shortest path method (cf. e.g. Henig, 1985ab).

The above presented procedure will be illustrated by the following example.

Example 2.1. The choice of a computer system.
Suppose that a company is offered k different computer systems, each of them satisfies its present needs. The computer differ in price, reliability, service quality, and technical characteristics such as the processor type and its clock speed, RAM, hard disk capacity and average access time, and a possibility of attaching additional equipment and peripherals like 3D video accelerators, sensors, or control devices. All above characteristics (except reliability and, perhaps, price) may be regarded as performance criteria $F_1,...,F_N$ with discrete attribute sets. Assume that from the technical and financial purposes the buyer decides to do not buy the full system configuration at once, but prefers to extend it successively according to the future needs. Thus the choice of a system at time t_0 should

not be made taking into account just the values of $F(t_0)$, but one should apply a model of the future acquisition process within the decision process. According to the scheme presented in this paper, a system configuration would be modelled as a node in the graph G, while to an extension of the system would correspond an edge labelled by the expected price of this system extension. The final objective may consist in getting a fully configured and ready-to-use system at time T at possibly minimal price. The "full configuration" mentioned may be interpreted as a reference point in the sense of Algorithm 2.1 which allows to apply the solution methods specified in Thms. 3.1 and 2.2 and the Algorithms 1.1 and 2.1.

Finally, let us discuss the solution methods for decision processes which do not satisfy the stationarity, homogeneity, or evolution independence assumptions.

(i) For non-stationary homogeneous processes the graph G will be a function of time t. To solve the Problem 2.1 one has to apply a bicriteria shortest path algorithm for variable-structure networks [8].

(ii) If the decision process is not homogeneous, one cannot apply the simultaneous shortest path algorithm for all alternatives $\omega \in \Omega$ in G, since the structure of the graph G depends on ω. A solution to the Problem 2.1 may be found by solving the single evolution problem described in Sec. 2.1 for each graph $G(\omega)$, and aggregate the solutions as in Step 4 of Algorithm 2.1.

(iii) If the criteria are not evolution independent then the edges of the graph G may not be associated to any combination of edges in the graphs $G_1,...,G_N$. Manual editing of the transition pattern $P_{1,...N}$ is required.

(iv) Processes which are neither homogeneous nor stationary, nor the criteria are evolution independent may be analysed applying simultaneously the appropriate combination of procedures outlined in (i) and (ii).

3. OPTIMISING THE POPULATION STRUCTURE

The above presented framework may be applied to solve a variety of decision problems. Here, we will formulate the problem of optimising the structure of a finite population Ω by assuring that after a fixed time T a maximal number of elements of Ω is characterised by nondominated values of criteria.

We assume that the elements of a finite population Ω are classified according to N ordered classification criteria $F_1,...F_N$. Each element $\omega \in \Omega$ may pass to another class on the time interval [t,t+1] if according to (1) a control $u_\alpha(t) \in U(t)$ has been applied to ω_α. All transitions can be described by M:= #Ω equations of type (1) sharing the same resources:

$$\sum_{1 \le \alpha \le M} b_\alpha(t)\, J(u_\alpha(t), \omega_\alpha, t) \le u_t \tag{12}$$

where $b_\alpha(t) \ge 0$ for all α. Let us note that the case $b_\alpha(t) := 1/x_{ti}$ with

$$x := \#\{\omega \in \Omega: F(t)(w) = F(t)(\omega)\}$$

corresponds to the situation where the same control u(t) acts simultaneously on all elements of Ω characterised by the same values of criteria.

At the macroscopic level, the evolution of Ω may be described by the following discrete-time controlled dynamical system:

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t w_t + \eta_t \\ z_{t+1} &= C_t x_t + D_t w_t + \zeta_t \end{aligned} \tag{13}$$

for $t=t_0,...,T-1$, subject to the constraints

$$\sum_{t_0 \le t \le T} (u_t + w_t) \le \xi(t_0,T) \tag{14}$$

where x_t, w_t, and z_t are the state, macroscopic control, and observation vectors, respectively, A_t, B_t, C_t, and D_t are real matrices, and η_t and ζ_t are random factors perturbing the growth/migration and observation processes for $t = t_0,..T$. The matrices A_t and B_t may be derived by aggregating the equations (1) for all $\omega \in \Omega$ and $t \in [t_0,T]$. The macroscopic controls w_t allow an "external" migration by attaching to (or removing from) Ω elements independently from the "internal" transitions controlled at the lower level (1), and may bear certain additional costs.

The state vectors $x_t = (x_{1t},...,x_{nt})$ contain the numbers of elements of Ω characterised by the same values of $F_1,...F_N$, for $t = t_0,...T$ (cf. Skulimowski and Schmid, 1992). Thus, there is a one-to-one correspondence *I* between the indices of the state variables and the elements of *V*, so for each $t \in [t_0,T]$ one may order the state variables $x_{1t},...,x_{nt}$ by the partial order generated from V. The values of F corresponding to the state variables are called the interpretation vector (cf. the above quoted paper of Skulimowski and Schmid) and denoted by $I(F,x_t)$. Assum-

ing that a population Ω is characterised by certain distribution of attributes at an initial moment t_0 represented by the state vector $x^0 := x(t_0)$, the aim of control is to achieve an optimal distribution of elements of Ω at time T, using for that a minimal quantity of resources represented by u_t and w_t. Below we propose two of a variety of possible optimisation problem statements. According to an initial remark, the first one of them is related to the nondominated values of F.

Let

$$K := \{v \in V : \exists\, \omega \in \Omega \text{ such that } v = F(T)(\omega)\}$$

and let $P(K)$ be the set of nondominated elements of K. Denote by $\Pi(K)$ the set of nondominated indices of the corresponding state variables, i.e. $\Pi(K) := I^{-1}(P(K))$. By definition, the relative population structure will be optimal iff

$$\Big(\sum_{j \in \Pi(T)} x_{Tj}\Big) / \Big(\sum_{1 \le k \le T} x_{Tj}\Big) \text{ is maximal} \tag{15}$$

$$\sum_{t_0 \le t \le T} (q_t u_t + r_t w_t) \text{ is minimal} \tag{16}$$

and

$$m_0 \le \sum_{1 \le k \le n} x_{Tk} \le m_1 \tag{17}$$

where q_t and r_t are positive real coefficients.

The above problem formulation may have a ecological, sociological or economical motivation, namely, assuming that a population Ω remains stable if under a classification F a maximal number of its members cannot get in touch with another individuals which are better (in the partial order in V) than themselves in all relevant aspects (represented as the criteria $F_1,...F_N$). In this setting, it is less important what is the shape of K and where it is situated at time T.

Introducing a loss function $\psi: V \times V \to IR_+$, which is right strictly order increasing, i.e.

$$v_1 \prec v_2,\ v_2 \prec v_3, \text{ and } v_1 \neq v_3 \Rightarrow \psi(v_1, v_2) \le \psi(v_1, v_3)$$

as e.g. a strictly convex distance function, we can evaluate the deviations from the ideal value $v^* := (v_{1,c(1)}, v_{2,c(2)}, ... v_{N,c(N)})$ at time T for each $\omega \in \Omega$. Consequently, the deviation of the whole set Ω can be characterised by the following criterion σ:

$$\sigma(u(t_0)(\Omega), ...u(T\text{-}1)(\Omega), w_{t0}, ...w_{T\text{-}1}) := \sum_{1 \le \alpha \le M} \psi(v^*, F(T)(\omega_\alpha)) \to \min \tag{18}$$

which may be more suitable for economical applications such as e.g. balancing the portfolio structure than (15)-(17). While optimising σ, we strive to approach the most preferred element of V for a possibly maximal number of elements of Ω. As the result, the set of alternatives actually characterised by nondominated values of F need not be numerous, but in average, their values are better approximating the ideal value v^* than in case of optimising the criterion (15). Let us note that always

$$\sigma = \sum_{1 \le k \le n} x_{Tk} \, \psi(v^*, I(F, x_{Ti})) \tag{19}$$

From a computational point of view a solution to the above problems consists in solving parallel N discrete optimal control problems coupled by the common resource or expense limitation (12). Thus, this problem requires nonstandard solution algorithms based on dynamical programming which have been proposed in Sec. 2.2. Roughly speaking, if the decision process is homogeneous and the criteria are evolution independent, one can construct the network G presented in Secs. 2.1 and 2.2, assigning additionally the varying labels x_{ti} to the nodes $v \in V$ determined by the interpretation vector $I(F,x_t)$.

The further procedure consists in finding shortest paths (in terms of the cost function J) to the nondominated values of $F(T)(\Omega)$, calculating the values of the macroscopic criterion σ, and choosing a subset $\Omega_1 \subset \Omega$ of elements which values are to be improved.

Let us remark that the above specified class of systems requires a state-space description with a usually large number of state variables representing the quantities of elements of each class, or other characteristics as functions of time.

The above description and assumptions reflects a complicated nature of certain real-life systems, where the growth coefficients may be derived *a posteriori* from empirical experience. As examples of such systems may serve e.g. the populations of concurrent technologies or innovations, inhabitants of a town, portfolio of a company, or a wildlife reservation.

4. CONCLUDING REMARKS

The motivation for introducing the above theory originates from the real-life multicriteria decision problems, such as portfolio management, technology transfer and foresight (cf. Skulimowski, 2006) or personnel choice, where the classical decision support methods do not allow to include the time aspects into the problem analysis. The new theoretical issues should constitute a basis for a more frequent including the dynamics in the analysis of multicriteria choice

problems, allowing thus a more efficient use of all preference information available which is a basis for an adequate modelling of real-life decision situations.

The approach proposed can be applied even in the simplest problems with discrete set of alternatives, and finite sets of admissible attribute values of each criterion. Methodologically, there is no conflict with other preference information since we merely extend the set of criteria values at t_0, $F(\Omega)$, to the set

$$\Psi := F(\Omega)x\{t_0\} \cup F(\Omega)(t_1)x\{t_1\} * J(U_1) \cup ... \cup F(\Omega)(T)x\{T\} * J(U_T) \qquad (20)$$

where $F(\Omega)(t_i)$ is the set of reachable values of F at time t_i, and U_i is the set of all sequences of strategies $(u_{j(1)}, ...u_{j(i)})$ on $[t_0, t_1]$ applicable to one of ω. This converts the initial problem $(F:X \rightarrow V) \rightarrow \min$ into a new problem with set-valued objective $\phi:\Omega \rightarrow 2^{\Psi}$, each alternative ω being characterised by the set of values of F, cost function J, and time :

$$\phi(\omega) := \{F(\omega)(t_0)\} \cup F(\omega)(t_1)x\{t_1\}xJ(U_1(\omega)) \cup \ldots \cup F(\omega)(T)x\{T\}xJ(U_T(\omega)) \quad (21)$$

where $U_i(\omega)$ contains the sequences of controls from U applicable to ω on $[t_0, t_1]$. Observe that $\psi = U \cup \{\phi(\omega) : \omega \in \Omega\}$.

However, the choice problem, as applied to the set Ψ, remains the same as in the classical multicriteria discrete choice problem, i.e. one has to select a compromise value $\psi_c := (v_c, J_c, \tau)$ from Ψ. Once this is done, one has to find the set

$$\phi^{-1}(\psi_c) := \{\omega \in \Omega : F(\omega)(\tau) = v_c \text{ and } J(F(\omega), u_1, ...u\tau) = J_c\} \qquad (22)$$

If the set (22) contains more than one alternative, all they are equivalent with respect to the choice criteria admitted. Therefore we expect that the above presented issues might be implemented as direct extensions of well known discrete choice and outranking decision models. Moreover, in most problems, J can be aggregated with one or more of the criteria $F_1, ..., F_N$, which simplifies the formulation of the problem (21)-(22).

The solution method applied to the Problem 2.1 shows a remarkable coincidence with the approach to the multicriteria optimal control of discrete-event systems presented in [8]. On the other hand, however, the approach to simultaneously control the evolution of a population Ω outlined in Sec. 3, resulting in a discrete-time control system model (13) could be applied to control large-scale discrete-event systems which allow an appropriate decomposition of the state-space.

In the present paper we concentrated our attention on deterministic processes, although in real-life situations some of the transitions may be stochastic. The analysis of such systems which involves the optimal control of discrete Markov processes (cf. e.g. [7]) may be considered as generalisation of the methods here presented and needs further investigation.

REFERENCES

1. Bertsekas D.P.: Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Inc., Englewood Cliffs 1987, pp. 376.
2. Decision Support Systems Based on Reference Sets; AGH University Scientific Publishers, Series: Monographs 1996, No. 40, p. 165.
3. Henig M.I.: The shortest path problem with two objective functions. European J. Oper. Res. 1985, 25, pp. 281-291.
4. Henig M.I.: The Principle of Optimality in Dynamic Programming with Returns in Partially Ordered Sets. Mathematics of Operations Research. 1985, 10, No 3, pp. 462-470.
5. Optimal Control of a Class of Asynchronous Discrete-Event Systems. Proceedings of the 11th IFAC World Congress. Tallinn (Estonia), August 1990. Vol. 3, p. 489-495. Pergamon Press, London 1991.
6. Optimising the Structure of a Partitioned Population. In: J. Henry and J.-P. Yvon. System Modelling and Optimisation; Lecture Notes in Control and Information Sciences. Vol. 197, Springer-Verlag, Berlin-Heidelberg-New York 1994, pp. 771-782.
7. Skulimowski A.M.J.: An Interactive Modification of the Decision Set to Attain a Target Point in Vector Optimisation Problems. In: Toward Interactive and Intelligent Decision Support Systems. Eds. Sawaragi Y., Inoue K., Nakayama H. Vol. 1, Lecture Notes in Economics and Mathematical Systems 1987, 285, Springer, pp.142-153.
8. Skulimowski A.M.J., Schmid B.F.: Redundance-free description of partitioned complex systems. Mathl. Comput. Modelling. 1992, 16, No 10, pp. 71-92.
9. Technology Transfer in Computer Science and Automatic Control (in Polish). Progress & Business Publishers, Kraków 2006.

Jaroslav Ramík

MULTI-CRITERIA APPROACHES TO DIAGNOSTICS OF ENTERPRISES USING ANALYTICAL NETWORK PROCESS*

Abstract

In this paper we investigate an economic diagnostic system in the situation of lack of data. We propose a diagnostic model working both with statistical and expert data. In case there exists statistical data, the diagnostic model should supply the results based on the well known Bayes' approaches, otherwise, the model should combine statistical and expert data by a generalized approach. Hence, this model is a generalization both the classical statistical approach and also expert one, which is allowed by Analytic Network Process.

Keywords

Multi-criteria decision making, analytic hierarchy process (AHP), analytic network process (ANP), pair-wise comparisons, subjective probability, Bayes' theory, diagnostics of enterprises.

INTRODUCTION

Solving problems of diagnostics of economical systems, particularly enterprises, we meet usually difficulties with interdependences among individual symptoms, i.e. the symptoms of economical systems and also causes of these symptoms. By the diagnostics we understand here a test (or system of tests) for predicting a state of the system in the future.

Recently, many diagnostic approaches are based on artificial intelligence, e.g. neural networks, see [3], the classical statistical approach based on Bayes theory is, however, still attractive, see e.g. [1], [2], [7]. This approach is focused on the assumption that the decision under uncertainty should utilize information about the decision environment, i.e. information about the history of the solution of the problem, expert knowledge etc.

Subjective probabilities in Bayes' theory allow for revision of the original prior information acquired from a large sample of population by means of the results of experiments, i.e. by so called posterior probabilities, see [7].

* This research has been supported by GACR Project No. 402/06/0431.

In economical diagnostics we often meet a situation with the lack of statistical data, the data are either out of reach or the sample is too small due to significant changes during the time and the dynamics of the process. That is why diagnostic systems utilizing both statistical and at the same time expert data are needed. In case there exists statistical data, the diagnostic model should supply the results based on the well known Bayes' approaches, otherwise, the model should combine statistical and expert data by a generalized approach, see [6]. In this paper we shall deal with the case of expert data i.e. a situation where no statistical data exists. This case is based on multi-criteria approach, particularly analytic hierarchy process (AHP), see [5] – [7].

1. BAYES' THEORY

Consider two-stage decision system: On the first stage we consider n disjoint events – *states of the system*: $S_1, S_2,...,S_n$, such that $S_i \cap S_j = \varnothing$ for $i \neq j$ and $\sum_{i=1}^{n} P(S_i) = 1$, $P(S_i) > 0$, $i=1,2,...,n$, is a probability of state S_i , see Fig.1. On the second stage consider m outcomes of the experiment $E_1, E_2,...,E_m$ such that $E_r \cap E_s = \varnothing$ for $r \neq s$ and $\sum_{r=1}^{m} P(E_r \mid S_i) = 1$, where $P(E_r \mid S_i)$, $i=1,2,...,n$, is a subjective probability of E_r on condition the existence of state S_i.

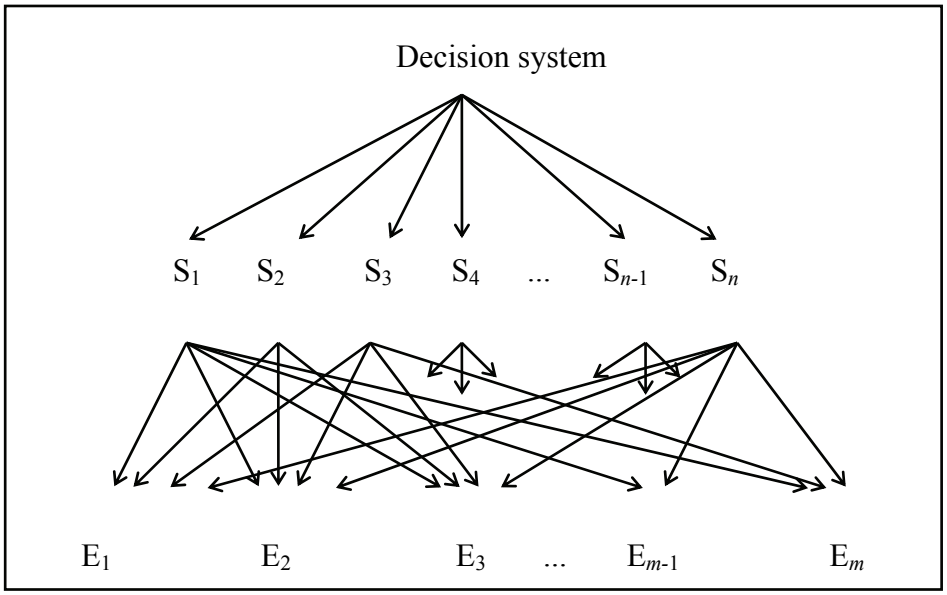


Fig. 1. Two-stage experiment – Diagnostic system

Subjective probability can be calculated by the well known *Bayes' formula*:

$$P(\mathrm{E}_r \mid \mathrm{S}_i) = \frac{P(\mathrm{E}_r \cap \mathrm{S}_i)}{P(\mathrm{S}_i)} \tag{1}$$

By disjointness of states S_i and properties of probabilities we get:

$$P(\mathrm{E}_r) = \sum_{i=1}^{n} P(\mathrm{E}_r \cap \mathrm{S}_i) \tag{2}$$

Substituting from (1) to (2) we obtain for $r = 1,2,...,m$:

$$P(\mathrm{E}_r) = \sum_{i=1}^{n} P(\mathrm{E}_r|\mathrm{S}_i)P(\mathrm{S}_i) \tag{3}$$

Further, we denote:

$$P(\mathrm{S}) = \begin{bmatrix} P(\mathrm{S}_1) \\ P(\mathrm{S}_2) \\ \vdots \\ P(\mathrm{S}_n) \end{bmatrix}, \; P(\mathrm{E}) = \begin{bmatrix} P(\mathrm{E}_1) \\ P(\mathrm{E}_2) \\ \vdots \\ P(\mathrm{E}_m) \end{bmatrix} \tag{4}$$

$$P(\mathrm{E} \mid \mathrm{S}) = \begin{bmatrix} P(\mathrm{E}_1|\mathrm{S}_1) & P(\mathrm{E}_2|\mathrm{S}_1) & \cdots & P(\mathrm{E}_n|\mathrm{S}_1) \\ P(\mathrm{E}_1|\mathrm{S}_2) & P(\mathrm{E}_2|\mathrm{S}_2) & \cdots & P(\mathrm{E}_n|\mathrm{S}_2) \\ \vdots & \vdots & \vdots & \vdots \\ P(\mathrm{E}_1|\mathrm{S}_m) & P(\mathrm{E}_2|\mathrm{S}_m) & \cdots & P(\mathrm{E}_n|\mathrm{S}_m) \end{bmatrix} \tag{5}$$

$$P(\mathrm{S} \mid \mathrm{E}) = \begin{bmatrix} P(\mathrm{S}_1|\mathrm{E}_1) & P(\mathrm{S}_1|\mathrm{E}_2) & \cdots & P(\mathrm{S}_1|\mathrm{E}_m) \\ P(\mathrm{S}_2|\mathrm{E}_1) & P(\mathrm{S}_2|\mathrm{E}_2) & \cdots & P(\mathrm{S}_2|\mathrm{E}_m) \\ \vdots & \vdots & \vdots & \vdots \\ P(\mathrm{S}_n|\mathrm{E}_1) & P(\mathrm{S}_n|\mathrm{E}_2) & \cdots & P(\mathrm{S}_n|\mathrm{E}_m) \end{bmatrix} \tag{6}$$

Then (3) can be expressed as follows:

$$P(\mathrm{E}) = P(\mathrm{E} \mid \mathrm{S})\, P(\mathrm{S}) \tag{7}$$

$P(\mathrm{S}_i)$ are called *prior probabilities*, they are known in advance - "a priori", usually as relative frequencies of populations. Also $P(\mathrm{E}_r \mid \mathrm{S}_i)$ are usually known in advance as statistical characteristics of the experiment. *Bayes' theorem*, the essence of the theory with the same name answers the question what is the probability of state S_i assuming that the outcome of the experiment is E_r. We look for *posterior probability* $P(\mathrm{S}_i \mid \mathrm{E}_r)$. Using the above defined notation the posterior probabilities are given by the following formula (called Bayes' theorem):

$$P(\mathrm{S}_i \mid \mathrm{E}_r) = \frac{P(\mathrm{E}_r \mid \mathrm{S}_i)\, P(\mathrm{S}_i)}{\sum_{k=1}^{n} P(\mathrm{E}_r \mid \mathrm{S}_k)\, P(\mathrm{S}_k)} \tag{8}$$

Let $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$ be a k-dimensional vector, then $diag\,(\mathbf{c}) = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & c_k \end{bmatrix}$

is called the *diagonal matrix to vector* **c**. Then Bayes' formula (8) can be also expressed in the following matrix form:

$$P(\mathrm{S} \mid \mathrm{E}) = diag(P(\mathrm{S}))\cdot P(\mathrm{E} \mid \mathrm{S})^{\mathrm{T}}\cdot [diag(P(\mathrm{E} \mid \mathrm{S})\cdot P(\mathrm{S}))]^{-1} \tag{9}$$

2. MULTI-CRITERIA DECISIONS AND AHP/ANP

Consider a decision system with three hierarchical levels:

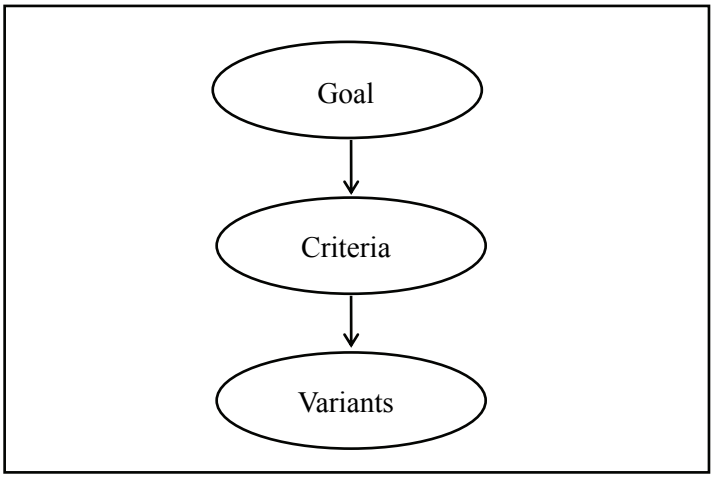


Fig. 2. Hierarchical system with 3 levels

This system is characterized by the *supermatrix* (see [7]):

$$\mathbf{W} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W_{21}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W_{32}} & \mathbf{I} \end{bmatrix} \tag{10}$$

here $\mathbf{W_{21}}$ is the $n\times1$ matrix (weighting vector of the criteria), $\mathbf{W_{32}}$ is the $m\times n$ matrix (the columns of this matrix are evaluations of variants by the criteria), I is the unit $m\times m$ matrix. The limit matrix $\mathbf{W}^{\infty} = \lim_{k\to+\infty} \mathbf{W}^{k}$ (see [6]) is given as follows:

$$\mathbf{W}^{\infty} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W_{32}W_{21}} & \mathbf{W_{32}} & \mathbf{I} \end{bmatrix} \tag{11}$$

Here $\mathbf{Z} = \mathbf{W_{32}W_{21}}$ is the $m\times1$ matrix, i.e. the resulting priority vector of the variants. The variants can be ordered according to these priorities.

In real decision systems with 3 levels there exist typical interdependences among individual elements, e.g. criteria. Consider now the dependences among the criteria, see Fig. 3. Such a system can be solved by the method named Analytical Network Process (ANP), an extension of AHP, see [6].

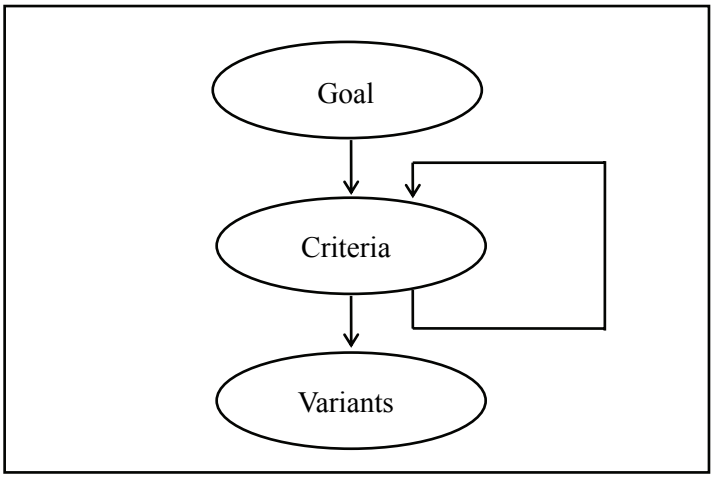


Fig. 3. Dependencies amongst criteria

This system is given by the supermatrix:

$$\mathbf{W} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W_{21}} & \mathbf{W_{22}} & \mathbf{0} \\ \mathbf{0} & \mathbf{W_{32}} & \mathbf{I} \end{bmatrix} \tag{12}$$

where the interdependences are characterized by $n\times n$ matrix $\mathbf{W_{22}}$. It is clear that matrix (12) need not be column-stochastic, i.e. sum of the elements in each column is equal to one, hence in general the limiting matrix does not exist. Stochasticity of this matrix can be saved by additional normalization of the col-

umns of the submatrix $\begin{bmatrix} W_{22} \\ W_{32} \end{bmatrix}$, by applying e.g. the Saaty's pairwise comparison method. Then there exists a limiting matrix $\mathbf{W}^{\infty}$ such that

$$\mathbf{W}^{\infty} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W}_{32}\left(\mathbf{I} - \mathbf{W}_{22}\right)^{-1}\mathbf{W}_{21} & \mathbf{W}_{32}\left(\mathbf{I} - \mathbf{W}_{22}\right)^{-1} & \mathbf{I} \end{bmatrix} \tag{13}$$

Hence the vector

$$\mathbf{Z} = \mathbf{W}_{32}\left(\mathbf{I} - \mathbf{W}_{22}\right)^{-1}\mathbf{W}_{21}$$

is used for ordering the variants i.e. for the decision making process.

In the systems with 3 levels there are usually interdependences among criteria and variants, see Fig. 4.

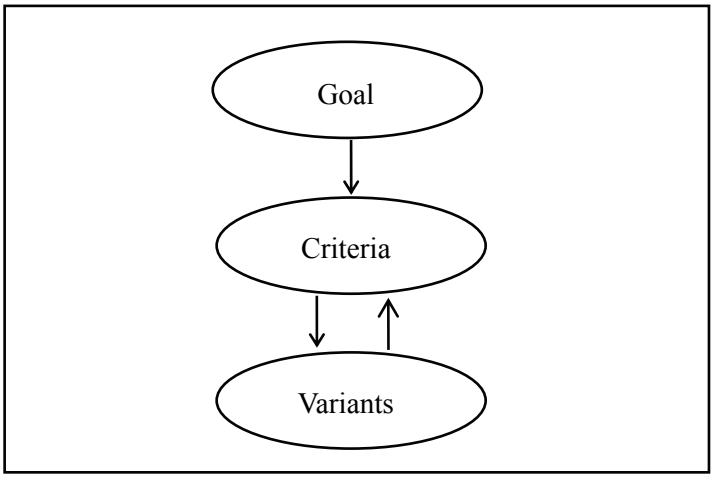


Fig. 4. Dependences amongst criteria and variants

This system is characterized by the supermatrix:

$$\mathbf{W} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W}_{21} & \mathbf{0} & \mathbf{W}_{23} \\ \mathbf{0} & \mathbf{W}_{32} & \mathbf{0} \end{bmatrix} \tag{14}$$

where the dependences are given by the $m \times n$ matrix $\mathbf{W}_{32}$, resp. by $n \times m$ matrix $\mathbf{W}_{23}$. Evidently, matrix (14) is stochastic, however, it is neither primitive nor irreducible, hence for the limiting matrix we apply Perron-Frobenius theorem, see [4].

Let $\mathbf{W_{21}}$, $\mathbf{W_{32}}$, $\mathbf{W_{23}}$ be column stochastic matrices with positive elements. Then for the limiting matrix $\mathbf{W}^{\infty}$ of the supermatrix $\mathbf{W}$ it holds:

$$\mathbf{W}^{\infty} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 \\ \mathbf{W}_{23}\mathbf{B}\mathbf{W}_{32}\mathbf{W}_{21} & \mathbf{A} & \mathbf{W}_{23}\mathbf{B} \\ \mathbf{B}\mathbf{W}_{32}\mathbf{W}_{21} & \mathbf{W}_{32}\mathbf{A} & \mathbf{B} \end{bmatrix} \tag{15}$$

where

$$\mathbf{A} = \lim_{k\to+\infty}\left[\mathbf{W}_{23}\mathbf{W}_{32}\right]^k, \ \mathbf{B} = \lim_{k\to+\infty}\left[\mathbf{W}_{32}\mathbf{W}_{23}\right]^k \tag{16}$$

Remark

Matrices $\mathbf{W_{32}}$ a $\mathbf{W_{23}}$ are supposed to be stochastic with positive elements, consequently they are primitive. The same holds for $\mathbf{W_{32}W_{23}}$ and $\mathbf{W_{23}W_{32}}$. Then there exist limit matrices (16), uniquely defined by positive vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$, such that $\mathbf{A} = \mathbf{a}\ \mathbf{e}_n^{\mathrm{T}}$, $\mathbf{B} = \mathbf{b}\ \mathbf{e}_m^{\mathrm{T}}$ and $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j = 1$, where $\mathbf{e}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, resp. $\mathbf{e}_m = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, is the n-dimensional, resp. m- dimensional vector. The matrix $\mathbf{A}$ is then stochastic $n\times n$ matrix, where all its columns are identically equal to vector $\mathbf{a}$, similarly, $\mathbf{B}$ is a stochastic square $m\times m$ matrix where all its columns are identically equal to vector $\mathbf{b}$. The priority vector $\mathbf{Z}$ is located in the third row of the limit matrix $\mathbf{W}^{\infty}$, i.e.

$$\mathbf{Z} = \mathbf{B}\mathbf{W}_{32}\mathbf{W}_{21} \tag{17}$$

Consider the following matrices of prior probabilities:

$$\mathrm{W}_{21} = \begin{bmatrix} P(\mathrm{S}_1) \\ P(\mathrm{S}_2) \\ \vdots \\ P(\mathrm{S}_n) \end{bmatrix} \tag{18}$$

$$W_{32}=\begin{bmatrix} P(E_1|S_1) & P(E_2|S_1) & \cdots & P(E_n|S_1) \\ P(E_1|S_2) & P(E_2|S_2) & \cdots & P(E_n|S_2) \\ \vdots & \vdots & \vdots & \vdots \\ P(E_1|S_m) & P(E_2|S_m) & \cdots & P(E_n|S_m) \end{bmatrix} \tag{19}$$

And the matrix of posterior probabilities:

$$W_{23}=\begin{bmatrix} P(S_1|E_1) & P(S_1|E_2) & \cdots & P(S_1|E_m) \\ P(S_2|E_1) & P(S_2|E_2) & \cdots & P(S_2|E_m) \\ \vdots & \vdots & \vdots & \vdots \\ P(S_n|E_1) & P(S_n|E_2) & \cdots & P(S_n|E_m) \end{bmatrix} \tag{20}$$

Bayes' theorem (9) gives the relationship among prior and posterior probabilities as follows:

$$W_{23}= diag(W_{21})\cdot W_{32}^{T}\cdot[diag(W_{32}W_{21})]^{-1} \tag{21}$$

If in the supermatrix $\mathbf{W}$ the block W_{23} is defined by (13), then it holds:

$$W_{23}W_{32}W_{21} = W_{21} \tag{22}$$

On the other hand, if in the supermatrix $\mathbf{W}$ the block W_{23} is defined by (16), then for the limiting matrix $\mathbf{W}^{\infty}$ we get:

$$\mathbf{W}^{\infty} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 \\ \mathbf{W}_{21} & \mathbf{W}_{21}\mathbf{e}_n^{T} & \mathbf{W}_{21}\mathbf{e}_m^{T} \\ \mathbf{W}_{32}\mathbf{W}_{21} & \mathbf{W}_{32}\mathbf{W}_{21}\mathbf{e}_n^{T} & \mathbf{W}_{32}\mathbf{W}_{21}\mathbf{e}_m^{T} \end{bmatrix} \tag{23}$$

In the limiting matrix (23) the first column is important as in the second row we have the vector of prior probabilities $P(S)= W_{21}$ and in the third row we get the vector of posterior probabilities $P(E) = W_{32}W_{21}$. The form (21) of marix W_{23}, i.e. Bayes' theorem, is a sufficient condition for $\mathbf{W}^{\infty}$ of the feedback system given by $\mathbf{W}$ in (14) can be written as (23), however, this condition is not sufficient. A natural question arises whether in $\mathbf{W}$ there exists a block W_{23} different to (21), with the same limiting matrix $\mathbf{W}^{\infty}$. The question is what is a sufficient condition for W_{23}, such that limit matrix $\mathbf{W}^{\infty}$ to matrix $\mathbf{W}$ from (21) has the form (23). The following theorem gives the answer to this question, see [6].

Theorem 1

Let W_{21}, W_{32}, W_{23} be column stochastic with positive elements in blocks of **W** defined by (14). Then $\mathbf{W}^{\infty}$ is in the form (23) if and only if the following equation holds:

$$W_{23}W_{32}W_{21} = W_{21} \tag{24}$$

In a particular system (e.g. diagnostic system) the matrices W_{21} and W_{32} are given beforhand. In case statistical data are at disposition they are prior probabilities P(S) and $P(E \mid S)$, otherwise, in case of expert data the matrices of priorities might be collected by Saaty's method of pairwise comparisons. We have to find matrix W_{23} of posterior probabilities $P(S \mid E)$ (case of statistical data), or, the feedback matrix of priotrities (case of expert data).

System (24) is a reasonable model for finding matrix W_{23}, which is, however, not uniquely solvable. In the stochastic case of matrix (21) classical Bayes' approach is a suitable method for finding solution of (24). However, in case of expert data this approach need not be the unique possible solution, there exist also some other solutions, different to Bayes' one, that might also be sufficient or even more advantageous. Here ANP is a new method generalizing the classical Bayes' approach allowing for a mix of statistical and expert data. We have the following theorem.

Theorem 2

Let Q be the $(m \times n)$ column stochastic matrix with positive elements such that:

$$QW_{21} = W_{32}W_{21} \tag{25}$$

Then matrix:

$$W_{23}^{*} = diag(W_{21}) \cdot Q^{T} \cdot [diag(W_{32}W_{21})]^{-1} \tag{26}$$

is a column stochastic solution of the system:

$$W_{23}^{*}W_{32}W_{21} = W_{21} \tag{27}$$

Let W_{23}^{*}, W_{32} , W_{21} satisfy (25) – (27). Then the limiting matrix $\mathbf{W}^{\infty}$ to supermatrix

$$\mathbf{W}^{*} = \begin{bmatrix} 0 & 0 & 0 \\ W_{21} & 0 & W_{23}^{*} \\ 0 & W_{32} & 0 \end{bmatrix}$$

is written in the form (23). If $Q \neq W_{32}$, then the solution of (27) is different to W_{23} in (21). This property is illustrated in the next section.

3. APPLICATION – A DIAGNOSTIC SYSTEM

In this part we apply the model described in the previous section to particular feedback system (14) of small and medium enterprises (SMEs). In the diagnostic system we consider three states: S_1 – the enterprise will bankrupt, S_2 – the enterprise will survive, S_3 – the enterprise will succeed. The prior probabilities – relative frequencies from statistical data of about 200 SMEs in Ostrava – Karviná region – are listed in the following table:

States	$P(S)$
S_1	0.20
S_2	0.70
S_3	0.10

To find out the economic state of the enterprise we applied a special test (experiment) with 4 outcomes (results):

E_1 – very bad result, E_2 – bad result, E_3 – good result and E_4 – excellent result.

In the next table the prior subjective probabilities are listed. They are based again on the above mentioned statistical data.

$P(E \mid S)$	SS_1	S_2	S_3
E_1	0.70	0.30	0.15
E_2	0.20	0.40	0.20
E_3	0.07	0.20	0.25
E_4	0.03	0.10	0.40

Probabilities of the symptoms are calculated as: $P(E) = P(E \mid S) \cdot P(S)$, the results is in the next table:

Symptoms	$P(E)$
E_1	0.365
E_2	0.340
E_3	0.179
E_4	0.116

The posterior probabilities are calculated from (9) as:

$$P(S \mid E) = diag(P(S)) \cdot P(E \mid S)^T \cdot [diag(P(E \mid S) \cdot P(S))]^{-1}.$$

The results are summarized in the following table:

$P(S \mid E)$	E_1	E_2	E_3	E_4
S_1	0.38	0.12	0.08	0.05
S_2	0.58	0.82	0.78	0.60
S_3	0.04	0.06	0.14	0.35

The values e.g. from the first column of the previous table can be interpreted as follows: If the outcome of the test of an enterprise is very bad (symptom E_1), then the probability that the enterprise would bankrupt (state S_1) is equal to 0.38, the probability that this enterprise would survive (state S_2) is 0.58 and probability the same enterprise would be successful (state S_3) is only 0.04. Analogically we could interprete the other three columns of the table, i.e. the other outcomes of the test. Now, let $W_{21} = P(E)$, $W_{32} = P(E \mid S)$, $W_{23} = P(S \mid E)$.

As an example consider the matrix Q defined below which satisfies (25) and (27), hence:

$$W_{23}^* \cdot W_{32} \cdot W_{21} = W_{21},$$
$$Q \cdot W_{21} = W_{32} \cdot W_{21},$$

with the following matrices:

$$W_{23}^* = diag(W_{21}) \cdot Q^T \cdot [diag(W_{32} W_{21})]^{-1}.$$

Q	S_1	S_2	S_3
E_1	0.52	0.33	0.33
E_2	0.38	0.37	0.02
E_3	0.03	0.20	0.33
E_4	0.07	0.10	0.32

W_{32}	S_1	S_2	S_3
E_1	0.70	0.30	0.15
E_2	0.20	0.40	0.20
E_3	0.07	0.20	0.25
E_4	0.03	0.10	0.40

W_{23}^*	E_1	E_2	E_3	E_4
S_1	0.29	0.22	0.03	0.12
S_2	0.62	0.77	0.78	0.60
S_3	0.09	0.01	0.18	0.28

By Theorem 1 and 2 the limiting matrices to the following matrices **W** and **W***:

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 \\ W_{21} & 0 & W_{23} \\ 0 & W_{32} & 0 \end{bmatrix}, \mathbf{W}^* = \begin{bmatrix} 0 & 0 & 0 \\ W_{21} & 0 & W_{23}\,^* \\ 0 & W_{32} & 0 \end{bmatrix}$$

are identical, in spite of $Q \neq W_{32}$.

CONCLUSION

In this paper we have investigated an economic diagnostic system in the situation of lack of data. We have proposed a diagnostic model working both with statistical and expert data. In case there exists statistical data, the diagnostic model should supply the results based on the well known Bayes´ approach, otherwise, the model should combine statistical and expert data by a generalized approach. Hence, this model is a generalization both the classical statistical approach and also expert one, which is allowed by Analytic Network Process.

REFERENCES

1. Berger J.: Statistical decision theory and Bayesian analysis. Springer-Verlag, Berlin 1985.
2. Blair A., Nachtmann, R. Olson J., Saaty T.L.: Forecasting foreign exchange rates: An expert judgement approach. Socio-Economic Planning Sci. 21, pp. 363-369.
3. Charniak E., McDermott D.: Introduction to artificial intelligence. Addison Wesley, New York 1995.
4. Gantmacher F.R.: Teorija matric (in Russian). Nauka, Moskva, 1966.
5. Saaty T.L.: Multicriteria decision making - the Analytical Hierarchy Process. RWS Publications, Pittsburgh 1991, Vol. I, pp. 502.
6. Saaty T.L.: Decision Making with Dependence and Feedback – The Analytic Network Process. RWS Publications, Pittsburgh 1996, pp. 370.
7. Saaty T.L., Vargas T.L.: Diagnosis with dependent symptoms: Bayes theorem and the Analytic Hierarchy Process. Operations Research 1998, Vol. 46, No 4, pp. 491-501.

Tadeusz Trzaskalik

Sylwia Zawadzka

AHP APPLICATION TO RAW MATERIALS STOCK MANAGEMENT

Abstract

In the paper the problem of choosing logistics methods for control the level of stock is considered. The aim of the paper is to show how the AHP method can be applied for each sort of raw material stock management. For numerical illustration we use numerical data from a ceramic factory.

Keywords

Raw materials, stock management, analytic hierarchy process (AHP).

INTRODUCTION

One of the key problems for production firms is raw material stock management. In many small and medium-size Polish firms the problem of raw materials storage occurs. Firms very often apply former experience and signed agreements with suppliers. Logistic methods are used to determine the time of and quantity of ordering [2; 7; 8].

The problem of choosing the logistic method to control the optimal level of stock for each sort of raw material separately is multicriterial. It can be formulated as the problem of choosing the best alternative which can be solved by means of AHP method. Application of such methods seems to be attractive for decision makers [5; 6; 1].

The aim of this paper is to show how the AHP method can be applied for each sort of raw material stock management. For numerical illustration we use numerical data from a ceramic factory.

The paper consists of five chapters. In Chapter 2 logistics of raw materials is discussed. In Chapter 3 raw material stock management in a ceramic factory is shown. In Chapter 4 an application of AHP method for raw material stock management in the ceramic factory is proposed. The summary is given in Chapter 5.

1. LOGISTICS OF RAW MATERIAL STOCK

The logistic system in a firm can be divided into three parts:
- delivery logistics,
- production logistics,
- sale logistics.

Raw material stock management is an integral part of delivery logistics. The ability of keeping a low level of raw material stock is an important factor determining the competitiveness of the firm.

Raw materials are delivered by suppliers and do not require any technological operations. They are the basic materials bought for the production purposes. In a ceramic factory important raw materials are clay, chalk and sand.

There are several reasons to keep the raw materials stock in a firm:
- necessity to compensate for differences in intensity flows,
- protection against the uncertainty. Forecasts of demand and supply can be inexact. Raw materials stock protects the company against such forecasts and random disturbances as well,
- protection against an increase of demand,
- covering the shortages caused by delay in delivery,
- discounts connected with greater orders.

ABC classification can be applied for rational raw material stock management [2]. Raw materials utilized by a firm can be classified into class A, class B and class C.

Class A

The most important raw materials in production process. They influence the production output very much, thus they should be efficiently ordered and stored. They comprise 5-10% of the total quantity and 75-80% of the row materials total value.

Class B

Raw materials with stabilized characteristics. They comprise about 20% of the total quantity of row materials.

Class C

Mass raw materials. Their value is not significant.

Storing raw materials involves lock-up capital. Raw materials should be thus purchased in quantities needed for production only. It is important with respect to ordering the materials from suppliers and determining the rate of using them. It can be noticed that there are raw materials regularly used, raw materials

used at changeable rate, and raw materials rarely used. It is important for the decision maker to know the characteristics mentioned above for all the raw materials used in the production process.

If raw materials are used regularly, the synchronization of the demand and orders is required. In the case of raw materials used with random deviation, storing them is the best solution. If raw materials are used rarely, the determination of supplies and storing should be done on individual basis.

The appropriate raw materials stock management policy depends mainly on the rapidity of using them. The following questions should be answered before making a decision about storing raw materials:

1. Which materials should be stored in the storehouse?
2. What is the size of an optimal order?
3. When should the order be placed?
4. Which raw materials stock control system should be applied?

We will consider five most frequently applied raw materials stock management models [3; 8].

Model M1

This is a model of stock level determining the time of ordering. The company that uses this model defines the alarm stock level w that indicates the time of supply. The size of this supply should be fixed at the level Q^*, which is an optimal size of supply that depends on the average level of using raw materials, storing costs and fixed costs of supplies. The optimal size of supply ensures the minimal total cost (the original formula of the total costs includes stock creation costs and stock maintaining costs) (see Fig.1).

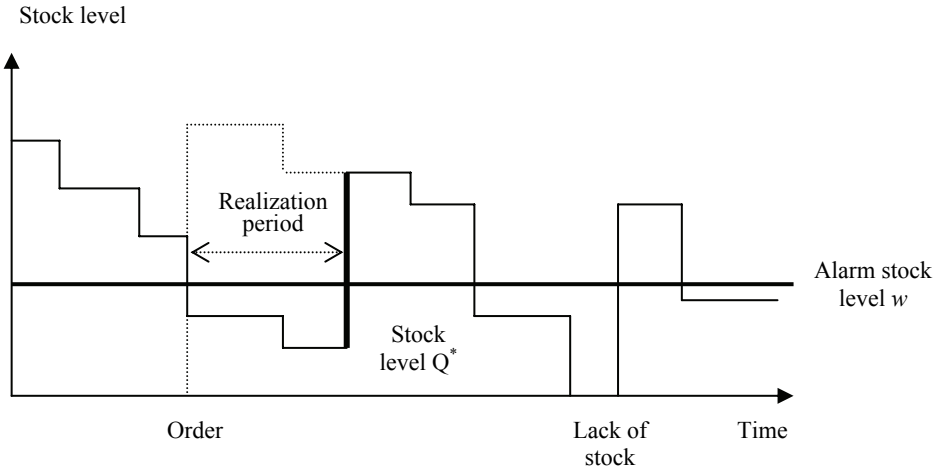


Fig. 1. Model M1

Model M2

This is a model of regular cycle of ordering. It takes into consideration the regularity of supplies and controls the stock at fixed and regular intervals. The main factor considered by the model is the size of order that increases the stock up to the fixed level *W*. The size of order is the result of subtraction of the current level of stock from the fixed level *W*. The supply is bigger if the current level of stock is low and it is smaller if the current level of stock is high. This method requires the determination of two factors: the level of resources *W* and the ordering cycle (see Fig. 2).

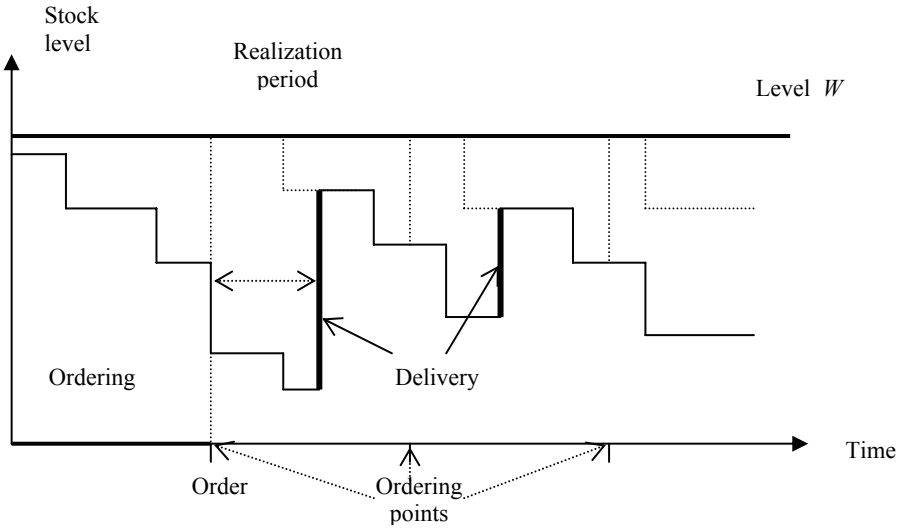


Fig. 2. Model M2

Model M3

This model determines the time of ordering in regular cycles. It is assumed that the stock level of a considered raw material is checked at fixed and regular periods. An optimal cycle of ordering *C*, which determines the time of checking the stock level, is identified. The order is placed when the current level of resources is lower than or equal to the ordering level *w* at the time of checking. This model is similar to the model M1, but does not require regular stock level checking. All the parameters are calculated in the same way as in the model M1 (see Fig. 3).

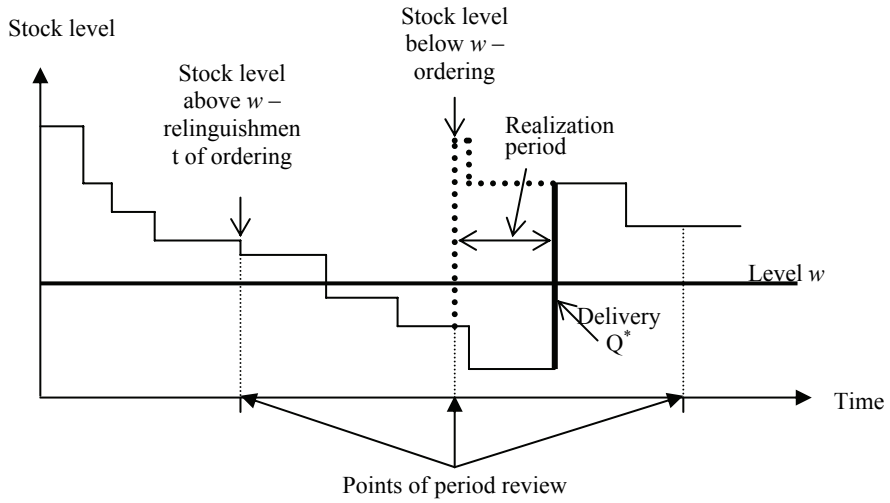


Fig. 3. Model M3

Model M4

This model is called a minimum-maximum model ("w,W"). Periodical monitoring of stock level is required. Orders are placed when the stock level is lower than the alarm level *w*. The size of order is the result of subtraction of the current stock level from the maximum level *W*. The additional factor *w*, called the alarm level ($w<W$) is determined. The order is not placed when the current level is between *w* and *W* (see Fig. 4).

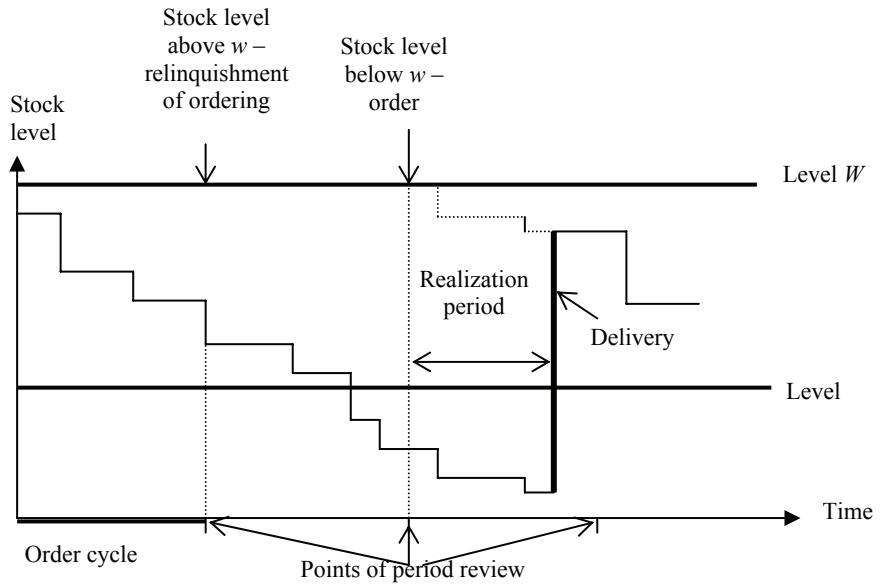


Fig. 4. Model M4

Model 5

In this model the time of ordering and the regular cycle of ordering are determined. There is a double protection against the stock shortage. The order that increases the stock up to the level W is placed if the current level drops below the alarm level w or in fixed periods. The scale of order is the result of subtraction of the current level of stock from the level W (see Fig. 5).

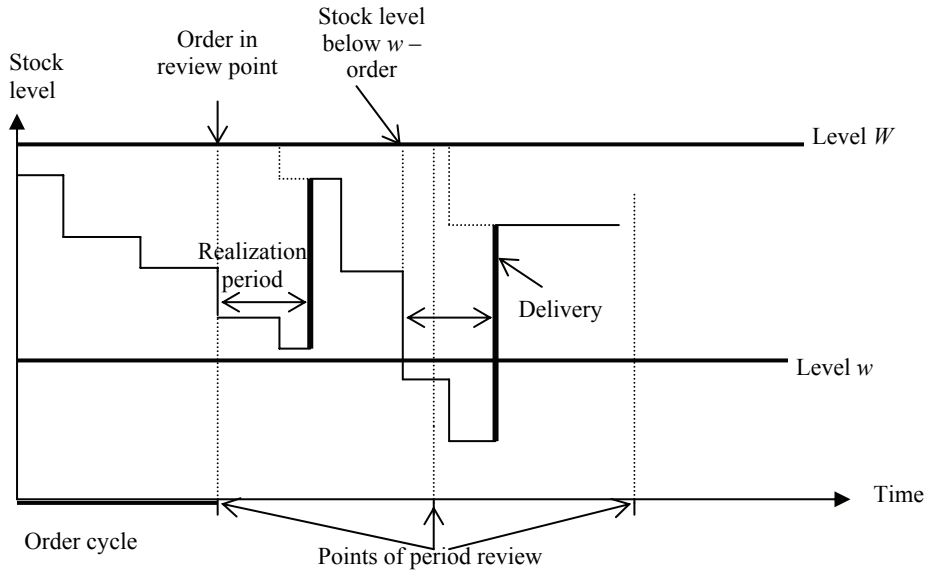


Fig. 5. Model M5

3. RAW MATERIALS STOCK MANAGEMENT MODELS IN A CERAMIC FACTORY

The production process in a ceramic factory can be divided into four stages:

1. Dry mass production – the grinding of the raw materials and mass liquidation.
2. Forming of products.
3. Baking.
4. Decorating – hand- or machine-made.

The dry mass production is the basic part of the production process because dry mass is the main component of every single product. The list of raw materials is given in Table 1. The quantities of raw materials given above are sufficient to produce two tones of dry mass.

Table 1

Raw materials for dry mass

No	Raw material	Symbol	Quantity [kg]	%
1	Clay TC1W	S1	300	12,7
2	Kaolinite Grudzeń	S2	1200	50,6
3	Clay Bełchatów	S3	180	7,6
4	Waste mass	S4	210	8,9
5	Sand	S5	130	5,5
6	Chalk	S6	150	6,3
7	Talc	S7	25	1,1
8	Glass capsules	S8	125	5,3
9	Feldspar	S9	50	2,1

Source: Database in considered ceramic factory

We will apply ABC classification method in our ceramic factory. The results are given in Table 2.

Table 2

ABC classification of raw materials used for dry mass production

Raw material	Value	Cumulated value	%	ABC Classification, A=75%,B=90%
S2	6507600,00	6507600,00	33,60	A
S1	4291650,00	10799250,00	55,77	A
S3	4256307,00	15055557,00	77,74	B
S6	1609368,75	16664925,75	86,06	B
S9	1028654,00	17693579,75	91,37	C
S4	883575,00	18577154,75	95,93	C
S8	584410,00	19161564,75	98,95	C
S5	201135,00	19362699,75	99,99	C
S7	2710,69	19365410,44	100,00	C
Sum	19365410,44			

Source: Databese in considered ceramic factory

The percentage share of raw materials in subsequent classes in total value of raw materials is given in Table 3.

Table 3

ABC Classification

Class	Number of raw materials	Quantity %	Value %
A	2	22,22	55,77
B	2	22,22	30,29
C	5	55,56	13,94
Sum	9		

In further considerations we will analyze raw materials of classes A and B.

Information about daily stock levels of raw materials of classes A and B and details of the production process in the period under consideration (July-September 2001) made it possible to conduct simulations by means of the models M1-M5.

The results of managing the level of stock of raw material S2 under application of the model M1 are given in Table 4. In the example shown below we apply daily utilization of kaolinite (S2), starting from the initial level of 20 000 kg. The days of placing the order are identified. The length of the arrow in Table 4 corresponds to the period of the order realization.

Table 4

Model M1 for kaolinite Grudzeń (S2)

No	S2	Model M1									
0		**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**
1	18800	14	20364	27	↓ 1164	**40**	**19891**	53	↓ 291	66	↓ 54
2	17600	15	19164	28	18327	41	18691	54	18054	67	17818
3	16400	16	17564	29	17127	42	17491	55	16854	68	16618
4	14600	17	15964	30	15927	43	16291	56	15654	69	15418
5	12800	18	14364	31	14727	44	15091	57	14454	70	14218
6	11000	19	12764	32	13527	45	13491	58	13254	71	13018
7	9800	20	11164	33	12327	46	11891	59	12054	72	11818
8	8600	21	9564	34	10527	47	10291	60	10854	73	10618
9	7400	22	8364	35	8727	48	8691	61	9054	74	9018
10	6200	23	7164	36	6927	49	7091	62	7254	75	7418
11	**5000**	24	5964	**37**	**5327**	50	5491	63	5454	76	5818
12	↓ 3800	**25**	**4764**	38	↓ 3727	**51**	**3691**	**64**	**3654**	**77**	**4218**
13	↓ 2600	26	2964	39	↓ 2127	52	1891	65	1854	78	2618

The results of applying the other models for material S2 are given in Tables 5-8. In a similar way the remaining raw materials from classes A and B are considered.

Table 5

Model M2 for kaolinit Grudzeń (S2)

No	**S2**	Model M2									
0	**20000**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**
1	18800	14	1400	27	3902	40	3302	53	3502	66	3102
2	17600	15	200	28	2102	41	2102	54	1902	67	1902
3	16400	16	20302	29	19502	42	20702	55	19902	68	20102
4	14600	17	18702	30	18302	43	19502	56	18702	69	18902
5	12800	18	17102	31	17102	44	18302	57	17502	70	17702
6	11000	19	15502	32	15902	45	16702	58	16302	71	16502
7	9800	20	13902	33	14702	46	15102	59	15102	72	15302
8	8600	21	12302	34	12902	47	13502	60	13902	73	14102
9	7400	22	11102	35	11102	48	11902	61	12102	74	12502
10	6200	23	9902	36	9302	49	10302	62	10302	75	10902
11	5000	24	8702	37	7702	50	8702	63	8502	76	9302
12	3800	25	7502	38	6102	51	6902	64	6702	77	7702
13	***2600***	***26***	***5702***	***39***	***4502***	***52***	***5102***	***65***	***4902***	***78***	***6102***

Tabela 6

Model M3 for kaolinite Grudzeń (S2)

No	**S2**	Model M2									
0	**20000**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**
1	18800	14	1400	27	1164	40	927	53	291	66	54
2	17600	15	19164	28	-636	41	18691	54	-909	67	17818
3	16400	16	17564	29	17127	42	17491	55	16854	68	16618
4	14600	17	15964	30	15927	43	16291	56	15654	69	15418
5	12800	18	14364	31	14727	44	15091	57	14454	70	14218
6	11000	19	12764	32	13527	45	13491	58	13254	71	13018
7	9800	20	11164	33	12327	46	11891	59	12054	72	11818
8	8600	21	9564	34	10527	47	10291	60	10854	73	10618
9	7400	22	8364	35	8727	48	8691	61	9054	74	9018
10	6200	23	7164	36	6927	49	7091	62	7254	75	7418
11	5000	24	5964	37	5327	50	5491	63	5454	76	5818
12	3800	25	4764	38	3727	51	3691	64	3654	77	4218
13	***2600***	***26***	***2964***	***39***	***2127***	***52***	***1891***	***65***	***1854***	***78***	***2618***

Table 7

Model M4 for kaolinite Grudzeń (S2)

No	S2	Model M4									
0	**20000**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**
1	18800	*14*	*1400*	27	2702	***40***	***3302***	53	4702	*66*	*6302*
2	*17600*	15	20702	*28*	*902*	41	2102	***54***	***3502***	67	5102
3	16400	*16*	*19102*	29	19502	*42*	*902*	55	2302	***68***	***3902***
4	*14600*	17	17502	*30*	*18302*	43	20702	*56*	*1102*	69	2702
5	12800	*18*	*15902*	31	17102	*44*	*19502*	57	20702	*70*	*1502*
6	*11000*	19	14302	*32*	*15902*	45	17902	*58*	*19502*	71	20702
7	9800	*20*	*12702*	33	14702	*46*	*16302*	59	18302	*72*	*19502*
8	*8600*	21	11102	*34*	*12902*	47	14702	*60*	*17102*	73	18302
9	7400	*22*	*9902*	35	11102	*48*	*13102*	61	15302	*74*	*16702*
10	*6200*	23	8702	*36*	*9302*	49	11502	*62*	*13502*	75	15102
11	5000	*24*	*7502*	37	7702	*50*	*9902*	63	11702	*76*	*13502*
12	***3800***	25	6302	*38*	*6102*	51	8102	*64*	*9902*	77	11902
13	2600	***26***	***4502***	39	4502	*52*	*6302*	65	8102	*78*	*10302*

Table 8

Model M5 for kaolinite Grudzeń (S2)

No	S2	Model M5									
0	**20000**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**	**No**	**S2**
1	18800	14	20702	27	1502	40	1502	53	1902	66	1702
2	17600	15	19502	28	18902	41	20302	54	19702	67	19502
3	16400	16	17902	29	17702	42	19102	55	18502	68	18302
4	14600	17	16302	30	16502	43	17902	56	17302	69	17102
5	12800	18	14702	31	15302	44	16702	57	16102	70	15902
6	11000	19	13102	32	14102	45	15102	58	14902	71	14702
7	9800	20	11502	33	12902	46	13502	59	13702	72	13502
8	8600	21	9902	34	11102	47	11902	60	12502	73	12302
9	7400	22	8702	35	9302	48	10302	61	10702	74	10702
10	6200	23	7502	36	7502	49	8702	62	8902	75	9102
11	**5000**	24	6302	37	5902	50	7102	63	7102	76	7502
12	3800	**25**	**5102**	**38**	**4302**	**51**	**5302**	**64**	**5302**	77	5902
13	*2600*	*26*	*3302*	*39*	*2702*	*52*	*3502*	*65*	*3502*	***78***	***4302***

3. APPLICATION OF AHP METHOD TO MODEL SELECTION

We assume that the decision maker wants to choose the best model for raw material stock level management. We will consider each sort of raw materials of classes A and B separately. The decision alternatives are as follows:

M1 – the model of stock level determining the time of ordering,
M2 – the model of constant cycle of order,
M3 – the model of stock level determining time of order in regular cycle of ordering,
M4 – the model minimum-maximum "w-W",
M5 – the model in which time of ordering and regular cycle of ordering are determined.

It seems reasonable to consider the following four criteria:
K1 – total cost of placing orders in the considered period,
K2 – total number of inspections,
K3 – unit price of order,
K4 – scale of the shortage (if applicable).

A hierarchical model created according to AHP rules is given in Fig. 6.

Applying AHP method for the kaolinite Grudzen (S2) raw material we obtain the ranking presented in Table 9.

Table 9

Priorities for S2

No	Model	Priority %
1	M4	31,90
2	M2	25,23
3	M5	19,37
4	M1	17,37
5	M3	6,12

A hierarchical model created according to AHP rules is given in Fig. 6.

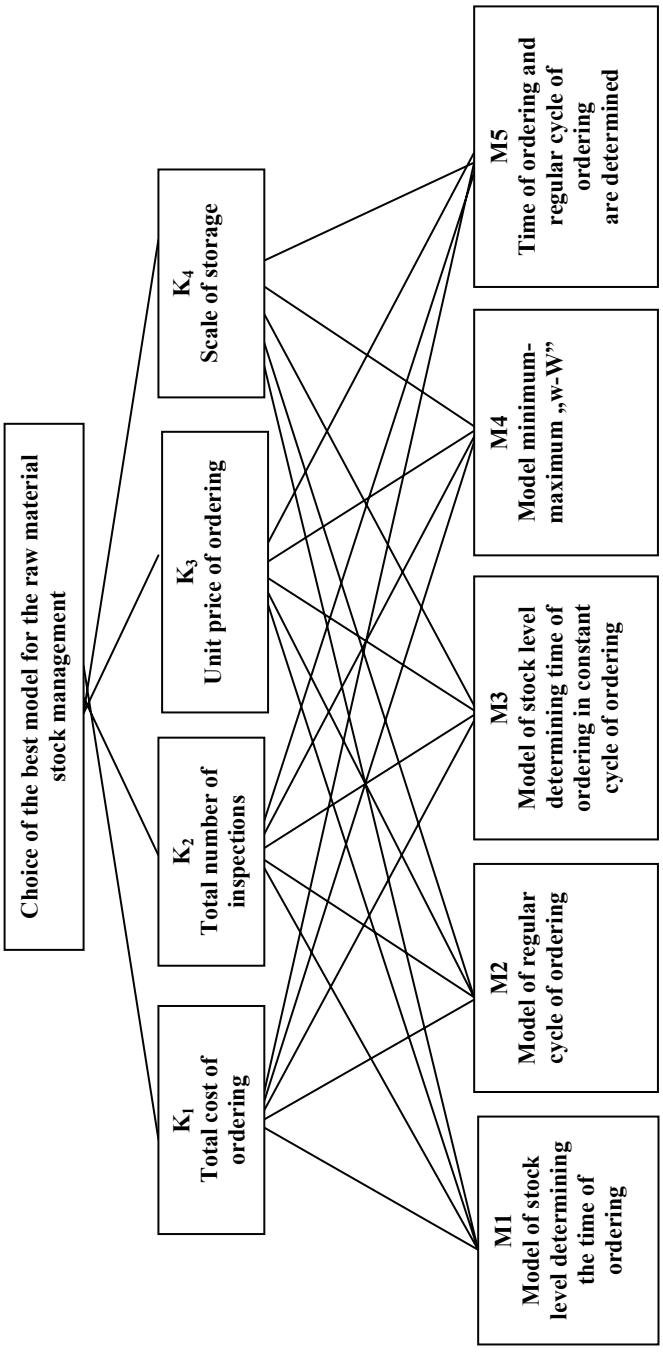


Fig. 6. AHP hierarchical model

The results show that the model M4 should be applied for stock level management of the kaolinite Grudzeń (S2) raw material. The local preferences show that the most important factor was the unit price of order. Tables 10 – 12 show the results obtained for the remaining materials.

Table 10

Priorities for S1

No	Model	Priorities %
1	M3	25,63
2	M1	24,58
3	M5	17,99
4	M2	16,67
5	M4	15,13

Table 11

Priorities for S3

No	Model	Priority %
1	M2	25,71
2	M1	25,61
3	M4	19,21
4	M5	18,24
5	M3	11,24

Table 12

Priorities for S6

No	Model	Priority %
1	M1	23,64
2	M4	22,72
3	M2	20,38
4	M5	17,67
5	M3	15,59

CONCLUSION

For the most important kaolinite Grudzen (S2) raw material the model M4 was determined by AHP method to be the most profitable one. For the clay TC1W (S1) raw material the model M3 is a little more favorable than the model

M1. Both these models can be thus applied to S1 raw material management. For the clay Bełchatów (S3) raw material the models M2 and M1 appear to be equally good. The less important chalk raw material (S6) can be managed best by the models M1 and M4.

The results of applying AHP method obtained for the presented example show that it can be successfully applied for real-world resources management problems in production companies.

REFERENCES

1. Golden B.L., Wasil E.A., Harker P.T.: The Analytic Hierarchy Process, Springer, Heidelberg 1989.
2. Kempny D.: Zapasy w systemie logistycznym firmy. AE, Katowice 1993.
3. Krawczyk S.: Metody ilościowe w logistyce (przedsiębiorstwa). C.H. Beck, Warszawa 2001.
4. Saaty T.L.: The Analytic Hierarchy Process. McGraw-Hill Co., New York 1980.
5. Saaty T.L., Kearns K.P.: Analytical Planning. The Organization of Systems. RWS Publications, Pittsbourgh 1991.
6. Saaty T.L., Vargas L.G.: The Logic of Priorities. RWS Publications, Pittsbourgh 1991.
7. Sarjusz-Wolski Z.: Definicje wybranych pojęć z dziedziny zakupu przemysłowego, gospodarki materiałowej i logistyki. „Gospodarka Materiałowa i Logistyka" 2000, No 7-8, 10, 12.
8. Sarjusz-Wolski Z.: Wybrane metody regulowania zapasów. Centralny Ośrodek Badawczo-Rozwojowy Gospodarki Materiałowej. Warszawa 1974.

CONTRIBUTING AUTHORS

Salvatore Greco
Faculty of Economics, University of Catania,
Corso Italia, 55, 95129 Catania, Italy
salgreco@unict.it

Hamdjatou Kane
Université du Québec en Outaouais,
Département des Sciences administratives,101, rue Saint-Jean-Bosco, C.P. 1250,
Succ. Hull,Gatineau, (Québec) Canada J8X 3X7,
Hamdjatou.Kane@uqo.ca

Benedetto Matarazzo
Faculty of Economics, University of Catania,
Corso Italia, 55, 95129 Catania, Italy
matarazz@unict.it

Jerzy Michnik
Department of Operations Research,
The Karol Adamiecki University of Economics in Katowice,
ul. Bogucicka 14, 40-226 Katowice, Poland
jmichnik@ae.katowice.pl

Dorota Miszczyńska
University of Łódź, Department of Operational Studies,
ul. Rewolucji 1905 r. 41, 90-214 Łódź, Poland
dmiszczynska@uni.lodz.pl

Marek Miszczyński
University of Łódź, Department of Operational Studies,
ul. Rewolucji 1905 r. 41, 90-214 Łódź, Poland
mmiszczynski@uni.lodz.pl

Maciej Nowak
Department of Operations Research, The Karol
Adamiecki University of Economics in Katowice,
ul. Bogucicka 14, 40-226 Katowice, Poland
nomac@ae.katowice.pl

Jaroslav Ramík
Silesian University Opava, School of Business Administration Karviná,
Czech Republic
ramik@opf.slu.cz

Sebastian Sitarz
State Higher Vocational School in Racibórz,
Institute of Technology and Mathematics,
ul. Słowackiego 55, 47-400 Racibórz, Poland
sebastian.sitarz@pwsz.raciborz.edu.pl

Andrzej M.J. Skulimowski
Institute of Automatic Control, AGH University of Science & Technology,
Cracow, Poland and International Centre for Decision Sciences and Forecasting,
Progress & Business Foundation, Cracow
ams@agh.edu.pl

Roman Słowiński
Institute of Computing Science,
Poznan University of Technology, 60-965 Poznan,
and Institute for Systems Research, Polish Academy of Sciences,
01-447 Warsaw, Poland
roman.slowinski@cs.put.poznan.pl

Tadeusz Trzaskalik
Department of Operations Research,
The Karol Adamiecki University of Economics in Katowice,
ul. Bogucicka 14, 40-226 Katowice, Poland
ttrzaska@ae.katowice.pl

Kazimierz Zaraś
Université du Québec en Abitibi-Témiscamingue,
Département des Sciences de la gestion, 445, boulevard de l'Université,
Rouyn-Noranda, Québec, J9X 5E4, Canada
Kazimierz.Zaras@uqat.ca

Sylwia Zawadzka
University of Łódź, Department of Operational Studies,
ul. Rewolucji 1905 r. 41, 90-214 Łódź, Poland
(master student)