

Juuso Liesiö, Pekka Mild, and Ahti Salo. 2007. Preference programming for robust portfolio modeling and project selection. *European Journal of Operational Research*, volume 181, number 3, pages 1488-1505.

© 2007 Elsevier Science

Reprinted with permission from Elsevier.



Preference programming for robust portfolio modeling and project selection [☆]

Juuso Liesiö, Pekka Mild, Ahti Salo ^{*}

Systems Analysis Laboratory, Helsinki University of Technology, P.O. Box 1100, 02015 TKK, Finland

Received 1 January 2005; accepted 1 December 2005

Available online 24 May 2006

Abstract

In decision analysis, difficulties of obtaining complete information about model parameters make it advisable to seek robust solutions that perform reasonably well across the full range of feasible parameter values. In this paper, we develop the Robust Portfolio Modeling (RPM) methodology which extends Preference Programming methods into portfolio problems where a subset of project proposals are funded in view of multiple evaluation criteria. We also develop an algorithm for computing all non-dominated portfolios, subject to incomplete information about criterion weights and project-specific performance levels. Based on these portfolios, we propose a project-level index to convey (i) which projects are robust choices (in the sense that they would be recommended even if further information were to be obtained) and (ii) how continued activities in preference elicitation should be focused. The RPM methodology is illustrated with an application using real data on road pavement projects.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Multiple criteria decision analysis; Project selection; Investment appraisal; Portfolio optimization; Incomplete information; Robustness

1. Introduction

Project portfolio selection is a strategic decision problem which is often characterized by multiple, conflicting and incommensurate objectives. The methodological literature on this problem offers a variety of approaches which have been used in public administration (e.g., Golabi et al., 1981; Kleinmuntz and Kleinmuntz, 1999) and industrial organizations (e.g., Stummer and Heidenberger, 2003). Reported experiences from these applications suggest that simple and transparent approaches which consider multiple criteria and accommodate incomplete information (e.g., parametric uncertainties) are more likely to be accepted by practitioners; they also tend to yield better decisions (Archer and Ghasemzadeh, 1999; Cooper et al., 1999; Keefer et al., 2004; Hämäläinen, 2004).

[☆] The authors have contributed equally and are listed in alphabetical order.

^{*} Corresponding author.

E-mail addresses: juuso.liesio@tkk.fi (J. Liesiö), pekka.mild@tkk.fi (P. Mild), ahti.salo@tkk.fi (A. Salo).

Methods of multi-criteria weighting – such as value tree analysis – are well-suited to the selection of a single alternative from a short-list of candidates in view of several evaluation criteria (Keeney and Raiffa, 1976). But this problem context, too, often involves incomplete information: for instance, the decision maker (DM) may be unable or reluctant to state exact preference statements, or it may be impossible to obtain complete information about how the alternatives perform with regard to the different criteria. A considerable amount of research has consequently sought to accommodate incomplete information in multi-criteria weighting models. This work has resulted in several methods – jointly referred by the term *Preference Programming* – which accommodate incomplete information by way of set inclusion (i.e., the ‘true’ parameter is contained in a feasible set characterized by the DM’s preference statements). The methods resemble each other in that they employ dominance concepts and decision rules in the development of decision recommendations; they also guide further preference elicitation efforts by providing supplementary information (see, e.g., White et al., 1984; Kirkwood and Sarin, 1985; Hazen, 1986; Weber, 1987; Rios Insua and French, 1991; Salo and Hämäläinen, 1992, 1995, 2001; Mármol et al., 1998; Dias and Clímaco, 2000).

In this paper, we develop the Robust Portfolio Modeling (RPM) methodology which extends the principles of Preference Programming to the problem of project portfolio selection. In RPM, the values of individual projects as well as project portfolios are modeled by an additive weighting model (e.g., Golabi et al., 1981; Golabi, 1987). Incomplete information about criterion weights is captured through linear inequalities (e.g., Arbel, 1989; Park and Kim, 1997; Salo and Punkka, 2005), while intervals are employed to model the performance of projects with regard to different criteria. In the presence of such incomplete information, the number of non-dominated portfolios can be very large. This leads to challenges in the computation of non-dominated portfolios, and also to questions about how such portfolios should be presented to the DM or harnessed in the development of decision recommendations.

Specifically, we first formalize the notion of dominance for project portfolio selection problems under incomplete information. Second, because the explicit enumeration of feasible portfolios becomes intractable when the number of projects increases (e.g., Stummer and Heidenberger, 2003), we develop a dynamic programming algorithm for determining all non-dominated portfolios subject to incomplete information. Third, we utilize these portfolios at the project-level by developing a core index which (i) identifies projects that are *robust* in the sense that they would be surely selected, even if additional information were to be acquired, (ii) helps the DM reject projects that are not contained in any non-dominated project portfolios, and (iii) provides guidance for further preference elicitation efforts. Fourth, we extend relevant decision rules from Preference Programming to the portfolio context and apply them as robust performance measures (Kouvelis and Yu, 1997; Salo and Hämäläinen, 2001). Taken together, these conceptual and computational results lead to an interactive decision support process for project portfolio selection.

The remainder of this paper is organized as follows. Section 2 introduces the RPM framework and defines the relevant dominance concept for project portfolios under incomplete information. An algorithm for computing all non-dominated portfolios is presented in Section 3. Section 4 considers the elicitation of additional information, proposes the core index for the development of project-level decision recommendations and describes the interactive RPM decision support process. An application of RPM to the selection of road pavement projects is presented in Section 5.

2. Preference programming for portfolio problems

2.1. Additive value

Multi-criteria weighting models are widely employed to assess and evaluate the overall value of project proposals (e.g., Henriksen and Traynor, 1999; Salo et al., 2004). In the portfolio context, several authors describe models where the recommended project portfolio is determined by maximizing the overall sum of the additive values from the individual projects, subject to specified resource constraints. For instance, Golabi et al. (1981) describe a case study on the selection of a portfolio of solar energy production projects that were evaluated with regard to several technical criteria. Kleinmuntz and Kleinmuntz (1999) use a similar approach to support capital budgeting decisions in health-care organizations where both qualitative (e.g., patient satisfaction) and quantitative (e.g., standard net present value) criteria are maximized. Memtsas (2003) presents an additive

model for the reserve site selection problem, with the aim of maximizing the (multi-criteria) conservation value of the reserve portfolio. Liberatore (1987), in turn, extends the Analytic Hierarchy Process (AHP) to assist managers in R&D portfolio selection and resource allocation.

Formally, let $X = \{x^1, \dots, x^m\}$ denote a set of m project proposals that are to be evaluated with regard to n criteria. The score of project x^j on the i th criterion is denoted by $v_i^j \geq 0$, i.e., the j th project is thus represented by the score vector $v^j = [v_1^j, \dots, v_n^j]$. These vectors form the rows of the score matrix $v \in \mathbb{R}_+^{m \times n}$ such that $[v]_{ji} = v_i^j$ (here $\mathbb{R}_+ = [0, \infty)$). The overall value of project x^j is the weighted average of its scores or, more specifically, $V(x^j) = \sum_{i=1}^n w_i v_i^j$ where the weight w_i measures the relative importance of the i th criterion. Without loss of generality, the weights $w = (w_1, \dots, w_n)^T$ and scores can be scaled so that

$$w \in S_w^0 = \left\{ w \in \mathbb{R}^n \mid w_i \geq 0, \sum_{i=1}^n w_i = 1 \right\} \tag{1}$$

and $v_i^j \in [0, 1]$. In this additive model, the weight w_i relates a unit increase in the criterion-specific score to an increase in the overall value. A project is preferred to another if it has the higher overall value of the two.

A project portfolio $p \subseteq X$ is a subset of available projects. Thus, the set of all possible portfolios is the power set $P = 2^X$. For any portfolio, the corresponding overall value is modeled as the sum of the values that are associated with the projects in the portfolio. We thus have the mapping $V : P \times S_w^0 \times \mathbb{R}_+^{m \times n} \rightarrow \mathbb{R}_+$:

$$V(p, w, v) = \sum_{x^j \in p} V(x^j) = \sum_{x^j \in p} \sum_{i=1}^n w_i v_i^j. \tag{2}$$

Rearranging the terms in (2) gives $V(p, w, v) = \sum_{i=1}^n w_i \sum_{x^j \in p} v_i^j$, where $\sum_{x^j \in p} v_i^j$ is the score of portfolio p with regard to the i th criterion. Theoretical premises of the additivity assumption (2) are discussed thoroughly by Golabi et al. (1981) and Golabi (1987).

2.2. Resource constraints and feasible portfolios

The number of all possible portfolios is $|P| = 2^m$. Typically, however, only a subset of the project proposals can be funded with available resources. These are modeled so that B_k is the available amount of the k th resource type ($k = 1, \dots, q$). The total budget vector is denoted by $B = [B_1, \dots, B_q]^T \in \mathbb{R}_+^q$.

If started, project x^j consumes $c_k^j \geq 0$ units of the k th resource. It is associated with the resource consumption vector $C(x^j) = [c_1^j, \dots, c_q^j]^T$ which is referred to as its cost. The total cost vector of a project portfolio is obtained by summing the cost vectors of its constituent projects, i.e., $C(p) = \sum_{x^j \in p} C(x^j)$ and $C(\emptyset) = \bar{0}$. The set of feasible portfolios is consequently

$$P_F = \{p \in P \mid C(p) \leq B\},$$

where the inequality holds componentwise. We also assume that (i) each project has a strictly positive cost with regard to some resource type and (ii) single projects, when taken alone, are feasible in the sense that they do not consume more resources of any type than what is available.

If complete weight and score information is available, the most preferred portfolio is the one that maximizes the overall value (2) subject to resource constraints. This optimal portfolio can be obtained as the solution to the following linear integer programming problem (ILP)

$$\max_{p \in P_F} \sum_{x^j \in p} \sum_{i=1}^n w_i v_i^j = \max_{z_1, \dots, z_m} \left\{ \sum_{j=1}^m z_j \sum_{i=1}^n w_i v_i^j \mid \sum_{j=1}^m z_j C(x^j) \leq B, z_j \in \{0, 1\} \right\}, \tag{3}$$

where $z_j = 1$ if and only if $x^j \in p$.

2.3. Incomplete information

In the RPM methodology, incomplete information is modeled by means of set inclusion. Thus, instead of using point estimates for weights and scores, the analysis is based on the consideration of sets of feasible

parameters that are consistent with the DM’s preference statements. While such sets may not lead to the identification of a single optimal portfolio, they may result in informative decision recommendations.

Incomplete information about criterion weights is modeled by the set of feasible weights, denoted by $S_w \subseteq S_w^0$ where S_w^0 is given by (1). The weight set S_w is assumed to be a convex set of weight vectors constrained by a set of linear inequalities that correspond to the DM’s preference statements. The literature on Preference Programming provides several methods for the elicitation of both complete and incomplete weight information (see Salo and Punkka, 2005, among others). At the two extremes, $S_w = S_w^0$ is the largest possible weight set which corresponds to lack of any weight information, while a point estimate in S_w^0 corresponds to complete information.

Incomplete score information about projects is modeled through score intervals which are assumed to contain the ‘true’ scores v_i^j . The lower and upper bounds of these intervals are denoted by \underline{v}_i^j and \bar{v}_i^j such that $\underline{v}_i^j \leq v_i^j \leq \bar{v}_i^j$ for all $j = 1, \dots, m$ and $i = 1, \dots, n$. The set of feasible scores is consequently $S_v = \{v \in \mathbb{R}_+^{m \times n} | v_i^j \in [\underline{v}_i^j, \bar{v}_i^j]\}$.

For a given portfolio p , the selection of different feasible scores and weights defines an interval of overall portfolio value such that for any $w \in S_w, v \in S_v$,

$$V(p, w, v) \in \left[\min_{w \in S_w} \underline{V}(p, w), \max_{w \in S_w} \bar{V}(p, w) \right], \tag{4}$$

where $V(p, w, v)$ is given by (2) and the mappings $\underline{V}, \bar{V} : P \times S_w^0 \rightarrow \mathbb{R}_+$ are given by

$$\underline{V}(p, w) = \sum_{x^j \in p} \sum_{i=1}^n w_i \underline{v}_i^j, \tag{5}$$

$$\bar{V}(p, w) = \sum_{x^j \in p} \sum_{i=1}^n w_i \bar{v}_i^j. \tag{6}$$

These upper and lower bounds on portfolio value are linear in w . Moreover, for a given weight vector w , the overall value of portfolio p ranges over the entire interval $[\underline{V}(p, w), \bar{V}(p, w)]$ when feasible scores v are allowed to vary in S_v .

The composite set of feasible weight and score parameters is denoted by the Cartesian product of S_w and S_v , i.e., $S \equiv S_w \times S_v$ and $s = (w, v) \in S$ is equivalent to $w \in S_w$ and $v \in S_v$. Such a non-empty set S of feasible weights and scores will be referred to as the *information set*.

2.4. Dominance structures

For any two portfolios p and p' , it is of interest to determine if one is preferred to the other in view of the information set S . Even if the value intervals (4) of these portfolios overlap, it is possible that (i) the overall value of p is higher than or equal to that of p' for all feasible combinations of the parameters (i.e., weights and scores) and that (ii) there exists a feasible combination of parameters such that the overall value of p is strictly higher than that of p' . If these two conditions hold, p is preferred to p' in the sense of (pairwise) dominance:

Definition 1. Let $p, p' \in P$. Portfolio p dominates p' with regard to the information set S , denoted by $p \succ_S p'$, iff

$$\begin{aligned} V(p, w, v) &\geq V(p', w, v) \quad \text{for all } (w, v) \in S \quad \text{and} \\ V(p, w, v) &> V(p', w, v) \quad \text{for some } (w, v) \in S. \end{aligned}$$

We denote $p \succ p'$ when there is no risk of confusion about the information set S . Based on Definition 1, it can be readily shown that the dominance relation is asymmetric ($p \succ p' \Rightarrow p' \not\succeq p$), irreflexive ($p \succ p' \Rightarrow p' \not\succeq p$) and transitive ($p \succ p' \wedge p' \succ p'' \Rightarrow p \succ p''$).

If the score information is complete (S_v such that $\underline{v}_i^j = \bar{v}_i^j \forall i, j$) and there is no *a priori* weight information ($S_w = S_w^0$), Definition 1 coincides with the concept of dominance as usually presented in multiple criteria optimization literature (see, e.g., Steuer, 1986). In this special case, the determination of dominance reduces to the

comparison of portfolio score vectors: portfolio p dominates p' if and only if its score with regard to each criterion is higher than or equal to that of p' , with a strict inequality on some criterion.

Dominance between two portfolios can be readily checked by using the bounds (5) and (6) and by noting that (i) projects that are included in both portfolios contribute equally to them both and (ii) project scores may vary across the full range of their respective intervals, regardless of what the other scores or weights are. All proofs are in Appendix A.

Theorem 1. For any $p, p' \in P$ and information set $S = S_w \times S_v$

$$p \succ_S p' \iff \begin{cases} \min_{w \in S_w} [\underline{V}(p \setminus p', w) - \overline{V}(p' \setminus p, w)] \geq 0, \\ \max_{w \in S_w} [\overline{V}(p \setminus p', w) - \underline{V}(p' \setminus p, w)] > 0, \end{cases}$$

where $\underline{V}(\cdot, \cdot)$ and $\overline{V}(\cdot, \cdot)$ are given by (5) and (6), respectively.

Because S_w is a convex polyhedron, the minimization and maximization in Theorem 1 are linear programming problems which obtain their solutions at extreme points of S_w .

A rational DM who seeks to maximize the overall portfolio value would not choose a dominated portfolio; hence, dominated portfolios can be discarded from further analysis. Conversely, if portfolio p is non-dominated, any other feasible portfolio p' either (i) gives a lower overall value than p for some feasible parameters $(w, v) \in S$ in the information set, or (ii) is equivalent to p , in the sense that the overall values of both portfolios are the same across the information set.

Definition 2. The set of non-dominated portfolios with regard to the information set S , denoted by $P_N(S)$, is

$$P_N(S) = \{p \in P_F \mid p' \not\succeq_S p \ \forall p' \in P_F\}.$$

We denote $P_N \equiv P_N(S)$, when there is no risk of confusion about the information set S .

Because the dominance relation is asymmetric, irreflexive and transitive, P_N cannot be empty unless the set of feasible portfolios P_F is empty. Furthermore, for each dominated portfolio $p' \in P_F \setminus P_N$ there exists at least one non-dominated portfolio $p \in P_N$ such that $p \succ p'$. The computation of all non-dominated portfolios can thus be regarded as a key step in supporting the selection of projects subject to incomplete information – it can eliminate unacceptable (dominated) portfolios from further consideration while retaining the interesting (non-dominated) ones.

3. Computation of non-dominated portfolios

In principle, the set of non-dominated portfolios can be determined by (i) enumerating all possible portfolios P , (ii) discarding the infeasible portfolios to obtain P_F and (iii) using pairwise dominance checks within P_F to obtain the set P_N . But because the number of possible portfolios with m projects is $|P| = 2^m$, this explicit enumeration procedure becomes intractable as the number of projects grows. For instance, if the generation of P with 20 projects takes one second (see, e.g., Stummer and Heidenberger, 2003), then it would take 1×2^{20} seconds (about 12 days) to generate P with 40 projects. More efficient optimization algorithms are therefore needed.

If complete score information is available ($\underline{v}_i^j = \overline{v}_i^j \ \forall i, j$), non-dominated portfolios can be computed with multi-objective multi-dimensional knapsack algorithms (for a survey, see Ehrgott and Gandibleux, 2000). To our knowledge, however, algorithms for solving the multi-objective knapsack problem with incomplete score information have not been developed to-date.

Building on the work of Villarreal and Karwan (1981), we propose a dynamic programming algorithm for computing the non-dominated portfolios. In our algorithm, projects are treated sequentially so that the k th round ($k \leq m$) of the algorithm generates portfolios that contain projects that belong to the set $\{x^1, \dots, x^k\}$. Moreover, only those portfolios which use resources efficiently are stored for use in subsequent rounds.

More specifically, for $k = 1, \dots, m$ we denote two sets of portfolios

$$P_F^k = \{p \in P_F | p \subseteq \{x^1, \dots, x^k\}\}, \tag{7}$$

$$P_N^k = \{p \in P_F^k | \nexists p' \in P_F^k \text{ such that } p' \succ p, C(p') \leq C(p)\}. \tag{8}$$

The set of non-dominated portfolios P_N can now be obtained by structuring the auxiliary sets P_N^k recursively, as stated by the following lemma.

Lemma 1. *Let $2 \leq k \leq m$. Then*

- (i) $p \in P_N^k \Rightarrow p \setminus \{x^k\} \in P_N^{k-1}$,
- (ii) $p \in P_F^k \setminus P_N^k \Rightarrow \exists p' \in P_N^{k-1}$ such that $p' \succ p, C(p') \leq C(p)$,
- (iii) $p \in P_N \Rightarrow p \in P_N^m$.

In view of the first two properties of Lemma 1, the set P_N^k can be constructed by extending the set P_N^{k-1} (rather than by examining the entire feasible set P_F^k). The third property, in turn, states that all the non-dominated portfolios can be readily obtained from the final set P_N^m . The algorithm can now be formalized as follows.

1. $P_N^1 \leftarrow \{\{\emptyset\}, \{x^1\}\}$
2. For $k = 2, \dots, m$ do
 - (a) $\tilde{P}_N^k \leftarrow \{p \in P_F^k | x^k \in p, (p \setminus \{x^k\}) \in P_N^{k-1}\}$
 - (b) $P_N^k \leftarrow \{p \in \tilde{P}_N^k | \nexists p' \in P_N^{k-1} \text{ such that } p' \succ p, C(p') \leq C(p)\} \cup \{p \in P_N^{k-1} | \nexists p' \in \tilde{P}_N^k \text{ such that } p' \succ p, C(p') \leq C(p)\}$
3. $P_N \leftarrow \{p \in P_N^m | p' \not\succeq p \ \forall p' \in P_N^m\}$

Step 1 considers only portfolios $\{\emptyset\}$ and $\{x^1\}$ for which conditions $C(\emptyset) \neq C(x^1) \leq B$ and $\{\emptyset\} \not\succeq \{x^1\}$ hold by assumption; thus $P_N^1 \leftarrow \{\{\emptyset\}, \{x^1\}\}$. The loop in Step 2 is repeated recursively by using the first property of Lemma 1, i.e., if x^k is removed from a portfolio that belongs to P_N^k , then the resulting portfolio must belong to P_N^{k-1} . In Step 2(a), project x^k is appended to all portfolios in P_N^{k-1} , and Step 2(b) removes portfolios that do not use resources efficiently, with the aim of obtaining P_N^k . In Step 2(b), pairwise dominance checks are needed only for pairs of portfolios that are taken from the sets \tilde{P}_N^k and P_N^{k-1} . Steps 2(a) and 2(b) are repeated for rounds $k = 2, \dots, m$. Finally in Step 3, dominated portfolios are removed from P_N^m to obtain P_N .

A concern with the above algorithm is that the auxiliary sets P_N^k may include many portfolios with a low value and low resource consumption. These portfolios are carried on until Step 3, where they are evaluated only based on their value and are therefore discarded.

However, the size of the set P_N^k can be reduced as follows. At round k of step 2, let $U_{k+1}(\tilde{B}, w)$ be an upper bound for the overall value of a portfolio $p \subseteq \{x^{k+1}, \dots, x^m\}$, measured at an extreme point of S_w subject to the resource constraint $C(p) \leq \tilde{B}$. Thus, $U_{k+1}(\tilde{B}, w)$ satisfies

$$U_{k+1}(\tilde{B}, w) \geq \max_p \{\bar{V}(p, w) | p \subseteq \{x^{k+1}, \dots, x^m\}, C(p) \leq \tilde{B}\}. \tag{9}$$

With regard to each resource type, the problem of maximizing the overall value with fixed (extreme point) weights is a knapsack problem, whereby an upper bound with regard to a single resource can be expressed in closed form through a continuous relaxation of the knapsack problem (Marthello and Toth, 1990). An upper bound with regard to all resources can thus be taken as the minimum of these resource-specific bounds. In view of Lemma 2, the bound $U_{k+1}(\tilde{B}, w)$ may help establish a dominance relationship already at the k th round of the algorithm.

Lemma 2. *Let $p, p' \in P_F, 1 \leq k < m$ and $a = \{x^{k+1}, \dots, x^m\}$. If*

$$\min_{w \in \text{ext}(S_w)} [\underline{V}(p \setminus (p' \setminus a), w) - \bar{V}((p' \setminus a) \setminus p, w) - U_{k+1}(B - C(p' \setminus a), w)] > 0,$$

then $p \succ p'$.

In order to discard portfolios $p' \in P_N^k$ on the basis of dominance $p \succ p'$, the value of the reference portfolio p should be as high as possible over the information set S ; preferably, it should be a non-dominated portfolio $p \in P_N$. Let $P_D \subset P_F$ now be a set of reference portfolios, which can be generated, for instance, by solving the ILP-problem (3) with arbitrary weights and scores $(w, v) \in S$.

We now modify the algorithm by adding an additional Step 2(c) which discards at the k th round some portfolios $p' \in P_N^k$ that cannot be augmented with projects x^{k+1}, \dots, x^m by expending remaining resources $B - C(p')$ so that the resulting portfolio would belong to P_N :

$$P_N^k \leftarrow \left\{ p' \in P_N^k \mid \min_{w \in \text{ext}(S_w)} [L(p \setminus p', w) - \bar{V}(p' \setminus p, w) - U_{k+1}(B - C(p'), w)] \leq 0 \ \forall p \in P_D \right\}.$$

As we shall demonstrate in the context of the example in Section 5, the above algorithm is more efficient than the explicit enumeration approach. The use of upper bounds in Step 2(c), in particular, seems very efficient in reducing the size of the auxiliary sets.

4. Development of decision recommendations

4.1. Additional information

After the computation of $P_N(S)$, the DM is typically faced with several non-dominated portfolios. If the number of such portfolios is very large, additional preference information (see Salo and Hämäläinen, 2001) can be solicited, with the aim of obtaining more conclusive decision recommendations based on fewer non-dominated portfolios.

In this context, *additional information* corresponds to further preference statements which reduce the set S , resulting in the revised information set $\tilde{S} \subset S$ (which, by assumption, is also non-empty). These statements can be elicited either as further weight constraints ($\tilde{S}_w \subset S_w$) and/or narrower score intervals ($\tilde{S}_v \subset S_v$). Because the dominance relationships between portfolios are contingent on the information set, the shift from S to \tilde{S} usually leads to a different set of non-dominated portfolios.

For the purpose of examining the impacts of additional information, we assume that $\tilde{S} \subset S$. We also assume that the ‘true’ parameter values are contained in \tilde{S} , as well as in the (relative) interior of S , defined as

$$\text{int}(S) = \{s \in S \mid \forall s' \in S \ \exists \delta > 0 \ \text{s.t.} \ s + \varepsilon(s - s') \in S \ \forall \varepsilon \in [0, \delta]\}.$$

This requirement means that the initial information set S is assumed to be balanced, in the sense that the revised set \tilde{S} is not entirely contained on the ‘border’ of S . Thus, for example, additional preference statements should not reduce a score interval to one of its original end points: rather, the initial score intervals should be wide enough so that the ‘true’ score is somewhere in the middle.

When the above assumptions hold, additional information may eliminate some portfolios from the previous set of non-dominated portfolios, but cannot add any new portfolios to it. That is, if a given portfolio is dominated subject to the information set S , it will remain dominated subject to the revised information set \tilde{S} , too.

Theorem 2. *Let \tilde{S}, S be information sets such that $\tilde{S} \subset S$ and $\text{int}(S) \cap \tilde{S} \neq \emptyset$. Then, $P_N(\tilde{S}) \subseteq P_N(S)$.*

The requirement $\text{int}(S) \cap \tilde{S} \neq \emptyset$ is a necessary condition for the conclusion of Theorem 2. This can be shown by considering projects x^1, x^2 with complete score information $v_1^1 = 0.5, v_2^1 = 0.5, v_1^2 = 1, v_2^2 = 0$ and weight information $S_w = \{w \in S_w^0 \mid w_2 \geq w_1\}$. Clearly, $x^1 \succ_S x^2$ because $V(x^1, w, v) = 0.5 \geq w_1 = V(x^2, w, v)$ whenever $0 \leq w_1 \leq 0.5$. If additional preference information were to result in the complete weight information $\tilde{S}_w = \{(0.5, 0.5)^T\}$ – in violation to the requirement of Theorem 2 – the dominance relation $x^1 \succ_{\tilde{S}} x^2$ would not necessarily hold: for example, if only one project can be started with the available resources, we would have $\{x^2\} \in P_N(\tilde{S})$ and $\{x^2\} \notin P_N(S)$.

From the computational point of view, a major implication of Theorem 2 is that the set of non-dominated portfolios needs to be computed by the dynamic programming algorithm in Section 3 with regard to the initial information set S only. Thereafter, the set $P_N(\tilde{S})$ can be obtained from $P_N(S)$ by pairwise dominance checks

within $P_N(S)$, i.e., $P_N(\tilde{S}) = \{p \in P_N(S) | p' \not\prec_{\tilde{S}} p \ \forall p' \in P_N(S)\}$, because the dominance relation is asymmetric, irreflexive and transitive.

4.2. Robust projects

Even if the number of non-dominated portfolios is high, it may be possible to provide incontestable recommendations about individual projects. Indeed, the characterization of projects that should be surely selected or rejected is one of the defining features of RPM. The proposed approach can also be extended to other discrete portfolio problems where the DM is presented with a set of non-dominated solutions instead of a unique solution.

Definition 3. The core index of project $x^j \in X$ with regard to the information set S , denoted by $CI(x^j, S)$, is

$$CI(x^j, S) = \frac{|\{p \in P_N(S) | x^j \in p\}|}{|P_N(S)|},$$

where $|\{\cdot\}|$ denotes the number of portfolios in the respective set.

If the core index of a project is 1, the project is included in *all* non-dominated portfolios; it is consequently called a *core project*. At the other extreme, if its core index is 0, the project is not included in *any* non-dominated portfolio; it is therefore referred to as an *exterior project*. Finally, projects whose core index is strictly greater than zero but less than one are called *borderline projects*.

Definition 4. With regard to the information set S , we define the sets of

- Core projects : $X_C(S) = \{x^j \in X | CI(x^j, S) = 1\}$,
- Borderline projects : $X_B(S) = \{x^j \in X | 0 < CI(x^j, S) < 1\}$,
- Exterior projects : $X_E(S) = \{x^j \in X | CI(x^j, S) = 0\}$.

The two following corollaries can be derived from [Theorem 2](#) to characterize how the core indexes of projects respond to additional information. Specifically, core (exterior) projects remain core (exterior) even if additional information is given. Moreover, additional score information can reduce the set of non-dominated portfolios only if it relates to borderline projects, because providing narrower score intervals for core or exterior projects has no impact on the set of non-dominated portfolios.

Corollary 1. Let $\tilde{S} \subseteq S$ such that $\text{int}(S) \cap \tilde{S} \neq \emptyset$. Then, $X_C(S) \subseteq X_C(\tilde{S})$ and $X_E(S) \subseteq X_E(\tilde{S})$.

Corollary 2. Let $\tilde{S} \subseteq S$ such that $\text{int}(S) \cap \tilde{S} \neq \emptyset$. If $\tilde{S}_w = S_w$ and $\tilde{v}_i^j = v_i^j$, $\tilde{v}_i^j = \bar{v}_i^j \ \forall i = 1, \dots, n, x^j \in X_B(S)$, then $P_N(S) = P_N(\tilde{S})$.

Core projects are *robust* choices in the sense that if point estimates for weight and score parameters $(w, v) \in S$ were to be acquired through the elicitation of additional information, these projects would be included in the portfolio that maximizes the overall portfolio value. Conversely, any portfolio which does not contain all core projects, or which contains an exterior project, is either dominated or infeasible. By [Corollary 1](#), additional information cannot alter the core (exterior) status of projects. The decision to select or to reject a project can be taken as soon as such a status is first established.

Non-dominated portfolios differ from each other only in terms of the borderline projects that they contain. Thus, from the viewpoint of decision support, a key objective of eliciting additional information is to reduce the set of non-dominated portfolios (and hence the number of borderline projects). In view of [Corollary 2](#), these elicitation efforts can be focused on obtaining narrower score intervals for borderline projects and/or more restrictive weight information. In this sense, core indexes help identify further information needs, which as such is one of the key purposes of sensitivity and robustness analysis (e.g., [Pannell, 1997](#)).

Apart from guiding the elicitation of additional information, core indexes justify project-specific yes/no decisions in a manner which accounts for incomplete information, resource constraints and alternative project

opportunities. These indexes also lead to a transparent project selection process, because for each project they transform information about non-dominated portfolios into a single performance measure.

4.3. Robustness measures for portfolios

Apart from considering robustness at the level of individual projects, it is instructive to analyze which portfolios are robust subject to the given information set. Such an analysis is called for especially when no additional information can be acquired, but decision recommendations at portfolio level must nevertheless be provided. In this situation, non-dominated portfolios can be regarded as discrete decision alternatives whose performance can be analyzed through suitable robustness measures.

Kouvelis and Yu (1997) present two robustness measures in the context of discrete optimization problems. *Absolute robustness* is defined as the worst-case performance of a solution while *robust deviation* is defined as the worst-case performance difference between the given solution and the best solution. These two robustness measures parallel the *maximin* and the *minimax-regret* decision rules in Preference Programming (see, e.g., Salo and Hämäläinen, 2001).

Maximin rule recommends the portfolio for which the minimum of its overall portfolio value over the information set is highest. Thus, the robustness measure of a portfolio is based on its ‘worst-case’ value (i.e., lower bound of its overall value interval (4)). The recommended portfolio is therefore in the set

$$P_{\min} = \arg \max_{p \in P_N} \min_{w \in S_w} \underline{V}(p, w).$$

Minimax-regret rule recommends the portfolio for which the maximum regret – defined as the greatest possible loss of value relative to some other portfolio over the information set – is smallest. The maximum regret of portfolio p can be computed by maximizing the value difference between other non-dominated portfolios and p over the information set S . Thus, the recommended portfolio belongs to the set

$$P_{mmr} = \arg \min_{p \in P_N} \max_{\substack{p' \in P_N, \\ w \in S_w}} [\overline{V}(p' \setminus p, w) - \underline{V}(p \setminus p', w)].$$

For any non-dominated portfolio p , it may be instructive to present the DM with the corresponding maximum loss of value ($\max_{p' \in P_N, w \in S_w} [\overline{V}(p' \setminus p, w) - \underline{V}(p \setminus p', w)]$) on the basis of which the minimax-regret recommendations are given.

The maximin and minimax-regret decision rules are based on different robustness measures and may therefore recommend different portfolios. The focus on non-dominated portfolios is well-motivated: for if a dominated portfolio were recommended to the DM, there would exist a non-dominated portfolio with a higher or equal portfolio value over the entire information set. Such a portfolio would outperform the dominated one in view of any robustness measure (see, e.g., Kouvelis and Yu, 1997).

4.4. Interactive decision support

Fig. 1 gives a schematic outline of the interactive RPM decision support process. At the outset, the DM is advised to supply wide enough score intervals and loose enough weight constraints so that the ‘true’ parameter values are contained in the initial information set. The corresponding non-dominated portfolios are computed by the dynamic programming algorithm with regard to the initial information set. When additional information is given, the set of non-dominated portfolios can be updated by pairwise comparisons (Theorem 2).

The DM can analyze non-dominated portfolios in terms of their criterion-specific scores and overall value intervals (4). She is also guided by core indexes at the project-level (Corollaries 1 and 2) and robustness measures at the portfolio-level (maximin, minimax-regret). If the robustness measures for non-dominated portfolios are not acceptable to the DM, she may seek to reduce the number of non-dominated portfolios by providing additional weight statements and/or narrower score intervals for borderline projects.

Core indexes indicate which projects should be surely selected or rejected in the view of the information set. These indexes are useful also in that they allow the DM to focus on borderline projects, which reduces the complexity of the portfolio selection problem when the number of projects is high.

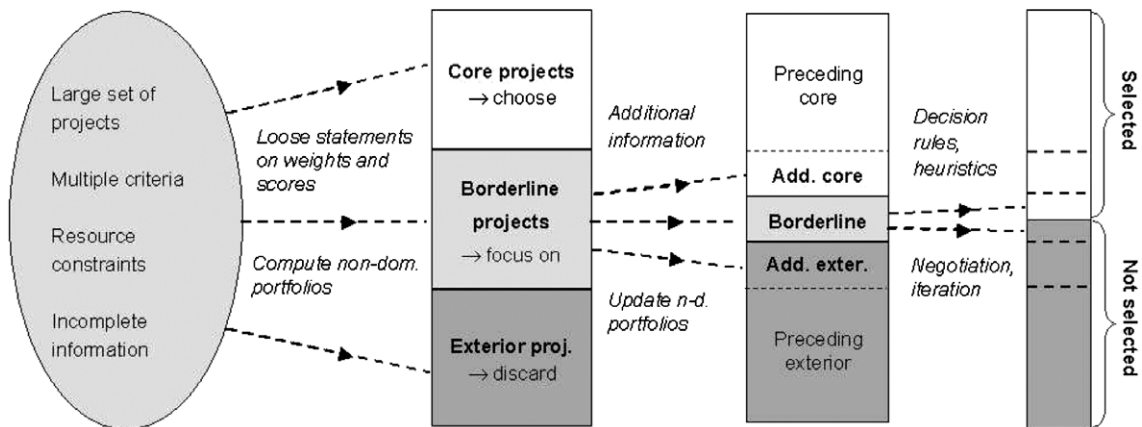


Fig. 1. RPM – decision support process.

For the purpose of promoting learning, it may be advisable to proceed from a relatively incomplete information set towards a more complete one. In the course of such an iterative process, the DM can learn, for instance, when a particular project is identified as one of the core or exterior projects. Moreover, the selection of the final project portfolio can be defended by showing which projects were among the core and borderline projects, respectively. It is even possible to backtrack at what stage core projects acquired their core status.

By construction, the portfolio recommendations based on two robustness measures are contingent on the information set. It is therefore advisable *not* to follow them early on when the number of non-dominated portfolios is still high and, specifically, when these measures (e.g., minimum overall portfolio value, maximum loss of value) are not yet acceptable to the DM. One reason for this is that one cannot exclude the possibility that these measures provide support for portfolios that would be outperformed by other portfolios, if additional information were to be acquired (Salo and Hämäläinen, 2001).

If several non-dominated portfolios remain after the iterative elicitation of additional information, the project portfolio can be selected by relying on one of the robustness measures, or on judgemental comparisons or other less formal approaches. But even in this case, the DM is strongly advised to select core projects, to reject exterior projects, and to focus on the analysis of borderline projects.

The RPM methodology seems particularly helpful in settings where the decision is to be taken by several DMs who act in public interest, on behalf of their constituencies, or as representatives of specific fields of expertise. In this case, the information set can be defined so that it subsumes the DMs' complementary viewpoints about the performance of proposed projects and the relative importance of evaluation criteria. The use of this information in RPM analysis yields core index values which help identify those projects that would either belong to or be absent from the DMs' optimal portfolios, if all the DMs were to solve the project portfolio selection problem independently. Thus, the RPM methodology does identify 'consensus projects' so that further attention can be devoted to the reduction of the information set, or to informal negotiations which may even address considerations that are not necessarily captured by the formal model.

Our belief in the practical significance of the RPM methodology is based on several case studies. In a telecommunication company, for instance, RPM was deployed for product strategy development (Lindstedt et al., 2006). In this case study, the members of the management group provided scores for 41 products, while the CEO of the company outlined information about the relative importance of the three criteria. Based on this information, core index values were assigned to all products and adopted as a useful point of departure in an ensuing management group meeting. In another case study, RPM was employed for screening promising innovation ideas in an extensive consultation process that was run on behalf of the Finnish Ministry of Trade and Industry (Könnölä et al., 2006). Here, well over 100 participants produced some 166 innovation ideas which were first evaluated with regard to three criteria (novelty, feasibility, societal relevance), and then extensively discussed in foresight workshops by using the resulting core index values as a guiding principle. The RPM

methodology can also be used as an ex post evaluation framework: that is, because RPM helps identify which projects are incontestably ‘best’ and ‘worst’ (in terms of their ex post impacts), corresponding longitudinal data can be examined to find out what ex ante characteristics may have contributed to such exceptional over- or underperformance (Salo et al., 2006). For further case studies and examples, see <http://www.rpm.tkk.fi/>.

5. Application of RPM to road pavement projects

We next illustrate the RPM process with actual data from a recent case study for the Finnish Road Administration (Finnra). At Finnra, the local experts develop an annual road pavement programme by deciding which ones out of the competing road pavement projects are undertaken in their road district (there are 9 districts in Finland with some 10 000 km of roads in each). Although multi-criteria data is systematically collected on potential pavement projects, the prevailing praxis of programme development has been essentially based on the use of a single criterion and holistic judgemental iteration. In this situation, Finnra managers were interested in the possibilities of obtaining multi-criteria decision support.

The salient features of this case study are presented through an illustrative ex post analysis of $m = 50$ project proposals which are evaluated with regard to $n = 4$ criteria, subject to a single budgetary constraint $q = 1$. The projects $X = \{x^1, \dots, x^{50}\}$ are generated by a road information management system such that each project proposal corresponds to a continuous road segment with scattered damages. This system provides data measurements on the criteria which correspond to performance indicators on (1) damage density, (2) attainable improvement in a composite driving cost index, (3) durability life of the repair, and (4) urgency index derived from the deterioration rate, respectively. On each indicator, proposals with higher data measurements have a higher priority than those with lower measurements. The cost of each project, $C(x^j) = c_1^j$, is given in euros. The budget $B = B_1$ is set at 60% of the sum of all proposals’ costs. The project proposals perform differently with regard to the criteria, wherefore the proposed programme depends on how the criteria are accounted for in the optimization problem.

The implications of possible inaccuracies in the measurement data are demonstrated by applying a relative variation of $\pm 2.5\%$ to the measurements on of each project. Thus, we assume that the ‘true’ performance belongs to the interval $[0.975x_i^j, 1.025x_i^j]$ where x_i^j denotes the point estimate about project’s performance, as recorded in the information system. On each criterion, indicator measurements are transformed into project-specific scores through a linear value function

$$v_i(x) = \frac{x - 0.975x_i^0}{1.025x_i^* - 0.975x_i^0}, \tag{10}$$

where x_i^0 and x_i^* are the worst and best recorded performance measurements on the i th criterion (e.g., Keeney and Raiffa, 1976). Lower and upper bounds for score intervals are given by $\underline{v}_i^j = v_i(0.975x_i^j)$ and $\bar{v}_i^j = v_i(1.025x_i^j)$, respectively, while scores for the initial point estimates are denoted by $\hat{v}_i^j = v_i(x_i^j)$. Because the criterion-specific value function (10) is linear, the weight ratio w_i/w_j represents the constant trade-off rate between one unit of x_i and x_j in the overall value function (2).

The example is presented in four phases which correspond to consecutive information sets S^1, S^2, S^3 and S^4 , each of which is a proper subset of its predecessor (see Table 1).

Table 1
Information sets, $S^1 \supset S^2 \supset S^3 \supset S^4$, at the different phases

Set	Weight information S_w	Score information S_v
S^1	$\{w \in S_w^0 w_1 \geq w_i, w_i \geq \frac{1}{20} \forall i\}$	$v_i^j \in [\underline{v}_i^j, \bar{v}_i^j] \forall i, j$
S^2	$\{w \in S_w^0 w_1 \geq w_2 \geq w_3 \geq w_4, w_i \geq \frac{1}{20} \forall i\}$	As above
S^3	$\{w \in S_w^0 0.9w_i^{roc} \leq w_i \leq 1.1w_i^{roc} \forall i\}$	As above
S^4	As above	$v_i^j = \hat{v}_i^j \forall i, x^j \in X_B(S^3)$

- In the first phase, the weight set S_w^1 is defined by stating the most important criterion only (damage density), in keeping with the earlier selection praxis at Finnra. The relevance of all criteria is ensured by imposing a lower bound on their criterion weights such that this bound is set at one fifth of the average weight $0.25/5 = 0.05$ (see Salo and Punkka, 2005). Score information is given by the intervals as described in the preceding paragraph.
- In the second phase, a full rank-ordering of criterion weights is specified ($w_1 \geq w_2 \geq w_3 \geq w_4$).
- In the third phase, the weight set is constructed around the rank order centroid (ROC) weight vector which is at the center of the feasible region defined by the rank-ordering (Edwards and Barron, 1994). For the given rank ordering, the ROC vector is $w^{roc} = (0.5208, 0.2708, 0.1458, 0.0625)^T$, around which each component of the weight vector let to vary $\pm 10\%$ in relative terms subject to the constraint $w \in S_w^0$.
- In the fourth phase, the score intervals of the remaining borderline projects are replaced by their initial point estimates while using weight information from the third phase.

Fig. 2 and Table 2 illustrate how the set of non-dominated portfolios and the projects' core indexes change when moving from one information set to the next. For the purpose of facilitating comparisons, the projects are numbered consecutively in the ascending order of their core indexes from the fourth phase ($CI(x^j, S^4)$). In most analyses, however, this 'last' information set would not be known at the outset, wherefore it is more meaningful to list projects in the ascending order of their core indexes for each information set.

In the first phase, the specification of the most important criterion leads to the identification of 15 core projects and 5 exterior projects, leaving 30 borderline projects. In the second phase, the additional information reduces the number of non-dominated portfolios from 234 to 77 and makes it possible to identify 5 new core

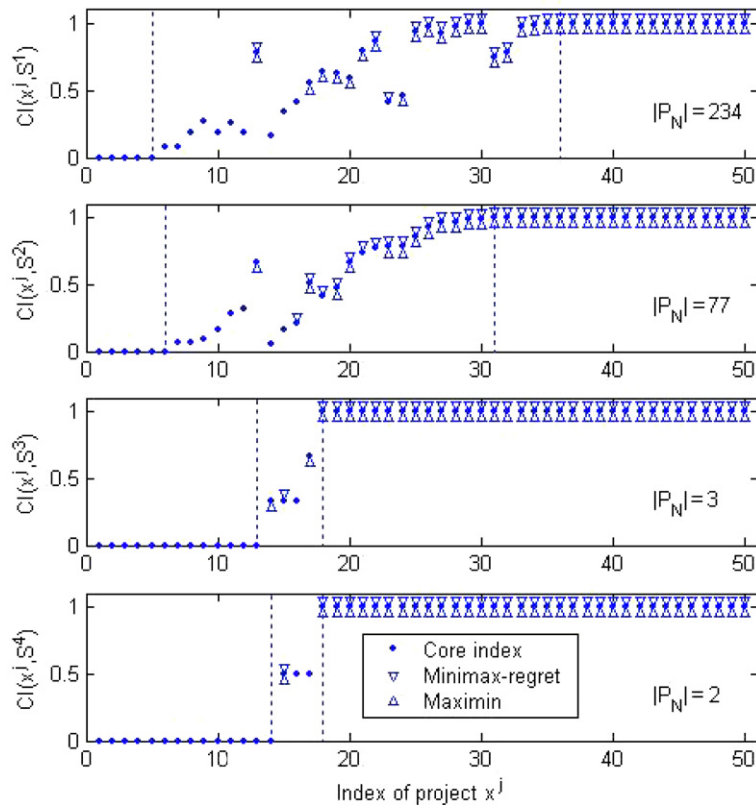


Fig. 2. The projects' core indexes with regard to the different information sets. Projects recommended by the minimax regret and maximin decision rules are marked with triangles. Projects are numbered consecutively based on their core indexes with regard to information set S^4 .

Table 2
Set sizes at the four phases of the example

Set	$ P_N $	$ X_C $	$ X_B $	$ X_E $	$\max_{p \in P_N} p $	$\min_{p \in P_N} p $
S^1	234	15	30	5	37	29
S^2	77	20	24	6	36	30
S^3	3	33	4	13	35	34
S^4	2	33	3	14	35	34

projects (x^{31}, \dots, x^{35}) and a new exterior project (x^6), too. In the third phase, the feasible weight set is far smaller: there are now only 3 non-dominated portfolios and 4 borderline projects, out of which one or two can be selected. The narrower score intervals in the fourth phase exclude yet another borderline project so that the DM is left with the problem of choosing either x^{15} or both x^{16} and x^{17} in addition to the 33 core projects.

Fig. 2 also shows the projects that are recommended by the maximin and minimax-regret decision rules (upward and downward triangles, respectively). In general, these recommended portfolios include projects with high core indexes, although individual projects may be excluded from the recommended portfolios when additional information is acquired. For instance, in the first phase (S^1) both decision rules recommend portfolios in which project x^{13} is contained. Yet, this project is among the exterior projects subject to information sets S^3 and S^4 . It is worth noting that in the fourth phase (S^4) both decision rules recommend the same portfolio.

This example with 50 projects cannot be readily solved by explicit enumeration. The set of non-dominated portfolios $P_N(S^1)$ was consequently computed by the algorithm in Section 3, whereafter the sets $P_N(S^2)$, $P_N(S^3)$ and $P_N(S^4)$ were obtained from their predecessors through pairwise comparisons between portfolios. The total computation time for determining $P_N(S^1)$ was about 27 minutes on a personal computer (Intel Pentium 1.3 GHz). The largest of the auxiliary sets (8) was P_N^{43} which contained 20649 portfolios. At the final round, P_N^{50} contained 1959 portfolios of which 234 were non-dominated. The two most time consuming steps 2(b) and 2(c) took about 26 minutes together. However, step 2(b) discarded 167469 portfolios in about 25 minutes while step 2(c) discarded 146663 portfolios in less than one minute when a reference set P_D of 10 portfolios was used. Because the generation of P_D by ILP took less than one second, it seems that the benefits of using reference portfolios for reducing the size of auxiliary sets P_N^k outweigh the additional effort involved in the generation and utilization of P_D .

The original case study with real DMs was carried out with a larger data set of more than 200 projects. The non-dominated portfolios were determined approximately by solving the ILP in Eq. (3) in a systematic grid of feasible weight vectors and point estimate scores. The results – which were essentially similar to those illustrated here – were presented to the Finnra managers for analysis and discussion. They found the RPM methodology intuitive and transparent; they could also readily understand its key concepts (e.g., maximization of multiple criteria, incomplete information, dominance concepts, core index values), even though they did not have a strong background in mathematics.

The ability to deal with incomplete information, in particular, was deemed useful. The managers noted that they were comfortable with ordinal statements about the importance of criteria (e.g., incomplete (S^1) or complete (S^2) rank ordering), but not with the specification of precise point estimates for criterion weights. In this sense, robust decision recommendations based on incomplete information were regarded more acceptable than a single optimal recommendation based on the use of complete information. Furthermore, the inconclusiveness of the results (cf. information set S^2 , for example) was not a problem: to the contrary, it was regarded beneficial in highlighting which projects should be analyzed further through judgmental iteration and negotiation.

The core index was found instructive, because the road pavement program is developed mainly on the basis of project-level considerations rather than through explicit portfolio optimization. By construction, this index serves to inform such considerations by distinguishing between core, borderline and exterior projects. The core index also illustrated the impacts of incomplete information through the ‘breadth’ of the band of borderline projects. Thus, it can support decision making in several ways, even if the actual project selection decisions were to differ from the recommendations based on robustness measures.

6. Discussion and conclusions

The RPM methodology developed in this paper offers robust decision recommendations for project portfolio selection in the presence of incomplete information. In essence, this methodology extends several Preference Programming concepts (e.g., decision rules, modeling of incomplete information about criterion weights and/or projects' scores) to the portfolio context and provides efficient algorithms towards their implementation.

The proposed methodology features a novel project-specific measure – core index – which is based on the share of those non-dominated portfolios in which a particular project is contained. By construction, this index helps identify projects that should be surely selected or rejected; it also guides further efforts towards the acquisition of additional information. Together with the proposed robustness measures and computational algorithms, the core index enables interactive decision support processes which guide the DM through recommendations for choosing among individual projects or entire portfolios. Feedback from real decision makers suggests that RPM is an intuitively appealing, transparent and readily applicable methodology for approaching multi-criteria portfolio selection problems.

The RPM methodology can be extended in several ways. First, because the strong additivity assumptions do not apply in the presence of strong project interactions (i.e., if the overall value of a portfolio differs from the value sum of its constituent projects), there is a need to accommodate such interactions through additional dummy projects and corresponding constraints, for instance (see, e.g., [Stummer and Heidenberger, 2003](#)). Second, problems with a very large number of projects or criteria are likely to remain intractable for exact algorithms, wherefore approximative (see, e.g., [Erlebach et al., 2002](#)) or heuristic algorithms (see, e.g., [Zitzler and Thiele, 1999](#)) are called for. Third, it may be of interest to determine how the recommendations would change when the levels of resource constraints assume values within some specified intervals. Towards this end, the algorithm in Section 3 can be extended to determine the non-dominated portfolios that correspond to varying levels of resource availability.

Acknowledgements

This research has been supported by the National Technology Agency of Finland (Tekes). We are grateful to Antti Punkka and three anonymous referees for their constructive comments.

Appendix A

Proof of Theorem 1. We show that the two conditions of [Definition 1](#) hold if and only if the minimization and maximization conditions of [Theorem 1](#) hold.

1°

$$\begin{aligned}
 &V(p, w, v) \geq V(p', w, v) \quad \forall (w, v) \in S \\
 &\iff \min_{w \in S_w, v \in S_v} [V(p, w, v) - V(p', w, v)] \geq 0 \\
 &\iff \min_{w \in S_w, v \in S_v} [V(p \setminus p', w, v) + V(p \cap p', w, v) - V(p \cap p', w, v) - V(p' \setminus p, w, v)] \geq 0 \\
 &\iff \min_{w \in S_w} \left[\min_{v \in S_v} [V(p \setminus p', w, v) - V(p' \setminus p, w, v)] \right] \geq 0.
 \end{aligned}$$

Since $\nexists x^j \in X$ s.t. $x^j \in p \setminus p'$ and $x^j \in p' \setminus p$, the minimization of the value difference with regard to v can be carried out by separation. Thus,

$$\begin{aligned}
 &\iff \min_{w \in S_w} \left[\min_{\substack{v^j \in [L^j, \bar{P}^j], \\ x^j \in (p \setminus p')}} [V(p \setminus p', w, v)] - \max_{\substack{v^j \in [L^j, \bar{P}^j], \\ x^j \in (p' \setminus p)}} [V(p' \setminus p, w, v)] \right] \geq 0 \\
 &\iff \min_{w \in S_w} [\underline{V}(p \setminus p', w) - \bar{V}(p' \setminus p, w)] \geq 0,
 \end{aligned}$$

2°

$$\begin{aligned} & \exists (w, v) \in S \text{ s.t. } V(p, w, v) > V(p', w, v) \\ & \iff \max_{w \in S_w, v \in S_p} [V(p, w, v) - V(p', w, v)] > 0 \\ & \iff \max_{w \in S_w} \left[\max_{\substack{v^j \in [\underline{v}^j, \bar{v}^j], \\ x^j \in (p \setminus p')}} [V(p \setminus p', w, v)] - \min_{\substack{v^j \in [\underline{v}^j, \bar{v}^j], \\ x^j \in (p' \setminus p)}} [V(p' \setminus p, w, v)] \right] > 0 \\ & \iff \max_{w \in S_w} [\bar{V}(p \setminus p', w) - \underline{V}(p' \setminus p, w)] > 0. \end{aligned}$$

The claim follows from cases 1° and 2°. \square

Proof of Lemma 1

(i) Assume contrary to the claim that $\exists p \in P_N^k$ s.t. $(p \setminus \{x^k\}) \notin P_N^{k-1}$. Then by (8), $\exists p' \in P_F^{k-1}$ s.t. $C(p') \leq C(p \setminus \{x^k\})$ and $p' \succ (p \setminus \{x^k\})$. If $x^k \notin p$ (i.e. $p = p \setminus \{x^k\}$), then $C(p') \leq C(p)$ and $p' \succ p$ and thus $p \notin P_N^k$, which is a contradiction. On the other hand, if $x^k \in p$, then

$$\begin{aligned} C(p') & \leq C(p \setminus \{x^k\}) \wedge p' \succ (p \setminus \{x^k\}) & (11) \\ \iff C(p' \cup \{x^k\}) & \leq C(p) \wedge (p' \cup \{x^k\}) \succ p. & (12) \end{aligned}$$

Since $C(p' \cup \{x^k\}) \leq C(p) \leq B$, $(p' \cup \{x^k\}) \in P_F^k$ and thus by (8) $p \notin P_N^k$, which is a contradiction.

(ii) Since $p \in P_F^k \setminus P_N$, $\exists p^1 \in P_F^k$ such that $p^1 \succ p$ and $C(p^1) \leq C(p)$. If $p^1 \in P_N^k$ then the lemma holds. On the other hand, if $p^1 \in P_F^k \setminus P_N^k$ then $\exists p^2 \in P_F^k$ such that $p^2 \succ p^1 \succ p$ and $C(p^2) \leq C(p^1) \leq C(p)$. If $p^2 \in P_N^k$ then the lemma holds and otherwise the deduction is continued until $p^h \in P_N^k$ such that $p^h \succ p^{h-1} \succ \dots \succ p^1 \succ p$ and $C(p^h) \leq C(p^{h-1}) \leq \dots \leq C(p^1) \leq C(p)$ is found. Since the set P_F^k is finite, such a portfolio always exists and the lemma holds.

(iii) Assume contrary to the claim that $\exists p \in P_N$ s.t. $p \notin P_N^m$. Then, by (8), $\exists p' \in P_F^m = P_F$ s.t. $p' \succ p$. Therefore, $p \notin P_N$, which is a contradiction. \square

Proof of Lemma 2. Assume that

$$\min_{w \in \text{ext}(S_w)} [\underline{V}(p \setminus (p' \setminus a), w) - \bar{V}((p' \setminus a) \setminus p, w) - U_{k+1}(B - C(p' \setminus a), w)] > 0. \tag{13}$$

The first term in (13) can be divided into two terms

$$\underline{V}(p \setminus (p' \setminus a), w) = \underline{V}(p \setminus p', w) + \underline{V}(p \cap p' \cap a, w)$$

and the second term in (13) can be divided into three terms

$$\begin{aligned} \bar{V}((p' \setminus a) \setminus p, w) & = \bar{V}((p' \setminus p) \setminus a, w) = \bar{V}(p' \setminus p, w) - \bar{V}((p' \setminus p) \cap a, w) \\ & = \bar{V}(p' \setminus p, w) - \bar{V}(p' \cap a, w) + \bar{V}(p' \cap a \cap p, w). \end{aligned}$$

Since $\underline{V}(p \cap p' \cap a, w) - \bar{V}(p' \cap a \cap p, w) \leq 0$, (13) implies that

$$\min_{w \in \text{ext}(S_w)} [\underline{V}(p \setminus p', w) - \bar{V}(p' \setminus p, w) + \bar{V}(p' \cap a, w) - U_{k+1}(B - C(p' \setminus a), w)] > 0. \tag{14}$$

Since portfolio $p' \in P_F$, i.e., it satisfies all resource constraints, we have $B - C(p') = B - C(p' \setminus a) - C(p' \cap a) \geq 0$, which is equivalent to $B - C(p' \setminus a) \geq C(p' \cap a)$. Since $j \geq k + 1$ for all $x^j \in p' \cap a$, (9) implies $\bar{V}(p' \cap a, w) \leq U_{k+1}(B - C(p' \setminus a), w)$. Thus, $\bar{V}(p' \cap a, w) - U_{k+1}(B - C(p' \setminus a), w) \leq 0$ and (14) implies that

$$\begin{aligned} \min_{w \in \tilde{S}_w} [\underline{V}(p \setminus p', w) - \overline{V}(p' \setminus p, w)] &> 0 \\ \Rightarrow \max_{w \in \tilde{S}_w} [\overline{V}(p \setminus p', w) - \underline{V}(p' \setminus p, w)] &> 0. \end{aligned}$$

Thus, $p \succ p'$ by **Theorem 1**. \square

Proof of Theorem 2. Assume contrary to the claim that $\exists p' \in P_N(\tilde{S}), p' \notin P_N(S)$. Then, $\exists p \in P_N(S)$ s.t. $p \succ_S p'$. That is,

$$V(p, w, v) \geq V(p', w, v) \quad \forall (w, v) \in S \wedge \tag{15}$$

$$\exists (w^*, v^*) \in S \text{ s.t. } V(p, w^*, v^*) > V(p', w^*, v^*). \tag{16}$$

Since $\tilde{S} \subset S$, it holds that

$$V(p, w, v) \geq V(p', w, v) \quad \forall (w, v) \in \tilde{S}. \tag{17}$$

By assumption, $\exists s' = (w', v') \in \text{int}(S) \cap \tilde{S}$. Let $s^0 = (w^0, v^0)$ such that $w^0 = w' + \varepsilon(w' - w^*)$ and $v^0 = v' + \varepsilon(v' - v^*)$, where $\varepsilon > 0$. Since $s' \in \text{int}(S)$, $\exists \varepsilon > 0$ s.t. $s^0 \in S$. By rearranging the terms we have

$$w' = \frac{1}{1 + \varepsilon} w^0 + \frac{\varepsilon}{1 + \varepsilon} w^* \equiv \alpha w^0 + \beta w^*,$$

$$v' = \frac{1}{1 + \varepsilon} v^0 + \frac{\varepsilon}{1 + \varepsilon} v^* \equiv \alpha v^0 + \beta v^*.$$

Note that $\alpha, \beta > 0$. In what follows we use a vector presentation $V(p, w, v) = z(p)v w$, where $z(p) \in R^{1 \times m}$, such that $z_j = 1$ if $x^j \in p$, and $z_j = 0$ otherwise

$$\begin{aligned} V(p, w', v') - V(p', w', v') &= (z(p) - z(p'))v'w' = (z(p) - z(p'))(\alpha w^0 + \beta w^*)(\alpha v^0 + \beta v^*) \\ &= (z(p) - z(p'))(\alpha^2 v^0 w^0 + \alpha \beta v^* w^0 + \alpha \beta v^0 w^* + \beta^2 v^* w^*) \\ &= \alpha^2 (V(p, w^0, v^0) - V(p', w^0, v^0)) + \alpha \beta (V(p, w^0, v^*) - V(p', w^0, v^*)) \\ &\quad + \alpha \beta (V(p, w^*, v^0) - V(p', w^*, v^0)) + \beta^2 (V(p, w^*, v^*) - V(p', w^*, v^*)) \\ &> 0, \end{aligned}$$

since all terms are non-negative by inequality (15) and the last one is strictly positive by inequality (16). Thus,

$$V(p, w', v') > V(p', w', v'). \tag{18}$$

Since inequalities (17) and (18) hold, $p \succ_{\tilde{S}} p'$. Thus, $p' \notin P_N(\tilde{S})$, which is a contradiction. \square

Proof of Corollary 1. Follows from **Theorem 2**. \square

Proof of Corollary 2. 1° ‘ \subseteq ’. **Theorem 2** implies that $P_N(S) \supseteq P_N(\tilde{S})$.

2° ‘ \subseteq ’. Let $\tilde{\underline{V}}(p) = \sum_{x^j \in p} \sum_{i=1}^n w_i \tilde{v}_i^j$ and $\tilde{\overline{V}}(p) = \sum_{x^j \in p} \sum_{i=1}^n w_i \tilde{v}_i^j$. Assume contrary to the claim that $\exists p' \in P_N(S)$ s.t. $p' \notin P_N(\tilde{S})$. Then, $\exists p \in P_N(\tilde{S})$ s.t. $p \succ_{\tilde{S}} p'$, which, by **Theorem 1**, is equivalent to

$$\min_{w \in \tilde{S}_w} [\tilde{\overline{V}}(p \setminus p', w) - \tilde{\underline{V}}(p' \setminus p, w)] \geq 0$$

$$\max_{w \in \tilde{S}_w} [\tilde{\overline{V}}(p \setminus p', w) - \tilde{\underline{V}}(p' \setminus p, w)] > 0.$$

Corollary 1 implies that $x^j \in X_C(\tilde{S}) \forall x^j \in X_C(S)$ and $x^j \in X_E(\tilde{S}) \forall x^j \in X_E(S)$. Therefore, $x^j \in X_B(S) \forall x^j \in (p \setminus p')$, $\forall x^j \in (p' \setminus p)$. Since we have assumed that $\tilde{v}_i^j = \underline{v}_i^j, \tilde{v}_i^j = \overline{v}_i^j \forall i = 1, \dots, n, \forall x^j \in X_B(S)$ and that $\tilde{S}_w = S_w$, we have

$$\min_{w \in S_w} [\bar{V}(p \setminus p', w) - \underline{V}(p' \setminus p, w)] = \min_{w \in S_w} [\tilde{\bar{V}}(p \setminus p', w) - \tilde{\underline{V}}(p' \setminus p, w)] \geq 0,$$

$$\max_{w \in S_w} [\bar{V}(p \setminus p', w) - \underline{V}(p' \setminus p, w)] = \max_{w \in S_w} [\tilde{\bar{V}}(p \setminus p', w) - \tilde{\underline{V}}(p' \setminus p, w)] > 0.$$

Thus, $p \succ_S p'$ and therefore $p' \notin P_N(S)$, which is a contradiction. \square

References

- Arbel, A., 1989. Approximate articulation of preference and priority derivation. *European Journal of Operational Research* 43, 317–326.
- Archer, N.P., Ghasemzadeh, F., 1999. An integrated framework for project portfolio selection. *International Journal of Project Management* 17, 207–216.
- Cooper, R.G., Edgett, S.J., Kleinschmidt, E.J., 1999. New product portfolio management: Practices and performance. *Journal of Product Innovation Management* 16, 333–351.
- Dias, L.C., Clímaco, J.N., 2000. Additive aggregation with variable interdependent parameters: The VIP analysis software. *Journal of Operational Research Society* 51, 1070–1082.
- Edwards, W., Barron, F.H., 1994. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organizational Behavior and Human Decision Processes* 60, 306–325.
- Ehrgott, M., Gandibleux, X., 2000. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum* 22, 425–460.
- Erlebach, T., Kellerer, H., Pferschy, U., 2002. Approximating multiobjective knapsack problems. *Management Science* 48, 1603–1612.
- Golabi, K., 1987. Selecting a group of dissimilar projects for funding. *IEEE Transactions on Engineering Management* 34, 138–145.
- Golabi, K., Kirkwood, C.W., Sichernman, A., 1981. Selecting a portfolio of solar energy projects using multiattribute preference theory. *Management Science* 27, 174–189.
- Hämäläinen, R.P., 2004. Reversing the perspective on the applications of decision analysis. *Decision Analysis* 1, 26–31.
- Hazen, G.B., 1986. Partial information, dominance, and potential optimality in multiattribute utility theory. *Operations Research* 34, 296–310.
- Henriksen, A., Traynor, A.J., 1999. A practical R&D project-selection scoring tool. *IEEE Transactions on Engineering Management* 46, 158–170.
- Keefer, D.L., Kirkwood, C.W., Corner, J.L., 2004. Perspective on decision analysis applications, 1990–2001. *Decision Analysis* 1, 4–22.
- Keeney, R.L., Raiffa, H., 1976. *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. John Wiley & Sons, New York.
- Kirkwood, C.W., Sarin, R.K., 1985. Ranking with partial information: A method and an application. *Operations Research* 33, 38–48.
- Kleinmuntz, C.E., Kleinmuntz, D.N., 1999. Strategic approaches for allocating capital in healthcare organizations. *Healthcare Financial Management* 53, 52–58.
- Könnölä, T., Brummer, V., Salo, A., 2006. Diversity in foresight: Insights from the fostering of innovation ideas. Helsinki University of Technology, Systems Analysis Laboratory, Research Report E19.
- Kouvelis, P., Yu, G., 1997. *Robust Discrete Optimization and Its Applications*. Kluwer Academic, Dordrecht.
- Liberatore, M.J., 1987. An extension of analytic hierarchy process for industrial R&D project selection and resource allocation. *IEEE Transactions on Engineering Management* 34, 12–18.
- Lindstedt, M., Liesiö, J., Salo, A., 2006. Participatory development of strategic product portfolio in a telecommunication company. *International Journal of Technology Management*, forthcoming.
- Mármol, A.M., Puerto, J., Fernández, F.R., 1998. The use of partial information on weights in multicriteria decision problems. *Journal of Multi-Criteria Decision Analysis* 7, 322–329.
- Marthello, S., Toth, P., 1990. *Knapsack Problems*. John Wiley & Sons, Chichester.
- Memtsas, D.P., 2003. Multipleobjective programming methods in reserve selection problem. *European Journal of Operational Research* 150, 640–652.
- Pannell, D., 1997. Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics* 16, 139–152.
- Park, K.S., Kim, S.H., 1997. Tools for interactive decision making with incompletely identified information. *European Journal of Operational Research* 98, 111–123.
- Rios Insua, D., French, S., 1991. A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research* 54, 176–190.
- Salo, A., Hämäläinen, R.P., 1992. Preference assessment by imprecise ratio statements. *Operations Research* 40, 1053–1061.
- Salo, A., Hämäläinen, R.P., 1995. Preference programming through approximate ratio comparisons. *European Journal of Operational Research* 82, 458–475.
- Salo, A., Hämäläinen, R.P., 2001. Preference ratios in multiattribute evaluation (PRIME) – elicitation and decision procedures under incomplete information. *IEEE Transactions on Systems, Man, and Cybernetics* 31, 533–545.
- Salo, A., Punkka, A., 2005. Rank inclusion in criteria hierarchies. *European Journal of Operational Research* 163, 338–356.
- Salo, A., Gustafsson, T., Mild, P., 2004. Prospective evaluation of a cluster program for Finnish forestry and forest industries. *International Transactions in Operational Research* 11, 139–154.

- Salo, A., Mild, P., Pentikäinen, T., 2006. Exploring causal relationships in an innovation program with robust portfolio modeling. *Technological Forecasting and Social Change* (Special Issue on ‘Tech Mining’), forthcoming.
- Steuer, R.E., 1986. *Multiple Criteria Optimization: Theory, Computation and Application*. John Wiley & Sons, New York.
- Stummer, C., Heidenberger, K., 2003. Interactive R&D portfolio analysis with project interdependencies and time profiles of multiple objectives. *IEEE Transactions on Engineering Management* 50, 175–183.
- Villarreal, B., Karwan, M.H., 1981. Multicriteria integer programming: A (hybrid) dynamic programming recursive algorithm. *Mathematical Programming* 21, 204–223.
- Weber, M., 1987. Decision making with incomplete information. *European Journal of Operational Research* 28, 44–57.
- White III, C.C., Sage, A.P., Dozono, S., 1984. A model of multiattribute decision making and trade-off weight determination under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics* 14, 223–229.
- Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computing* 3, 257–271.