European Journal of Operational Research 228 (2013) 169–180

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Robust multi-criteria ranking with additive value models and holistic pair-wise preference statements

Miłosz Kadziński [a,*], Tommi Tervonen [b]

[a] Institute of Computing Science, Poznań University of Technology, 60-965 Poznań, Poland
[b] Econometric Institute, Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 20 January 2012
Accepted 15 January 2013
Available online 1 February 2013

Keywords:
Decision analysis
Multiple criteria
Robust Ordinal Regression (ROR)
Stochastic Multicriteria Acceptability Analysis (SMAA)
Extreme Ranking Analysis (ERA)
Additive Value Function

ABSTRACT

We consider a problem of ranking alternatives based on their deterministic performance evaluations on multiple criteria. We apply additive value theory and assume the Decision Maker's (DM) preferences to be representable with general additive monotone value functions. The DM provides indirect preference information in form of pair-wise comparisons of reference alternatives, and we use this to derive the set of compatible value functions. Then, this set is analyzed to describe (1) the possible and necessary preference relations, (2) probabilities of the possible relations, (3) ranges of ranks the alternatives may obtain, and (4) the distributions of these ranks. Our work combines previous results from Robust Ordinal Regression, Extreme Ranking Analysis and Stochastic Multicriteria Acceptability Analysis under a unified decision support framework. We show how the four different results complement each other, discuss extensions of the main proposal, and demonstrate practical use of the approach by considering a problem of ranking 20 European countries in terms of 4 criteria reflecting the quality of their universities.

1. Introduction

Inadvertent biases and uncertainties constitute an indispensable part of many decision support processes. They are related to the specification of a decision problem, the environment in which the decision has to be made, and the character of the value system and preferences of a Decision Maker (DM) [3]. The complexity of this issue has led to the development of a framework for robustness analysis, i.e. a theoretical basis and a diversity of dedicated multiple criteria decision support methods that take into account internal and external uncertainties observed in the actual decision situations.

As noted by Vincke [29], robustness is often used to formulate requirements with respect to decision processes, methods, solutions, or conclusions. In this paper, we are interested in investigating the robustness of the provided conclusions, i.e. whether they are valid for all or for the most plausible sets of model parameters. We focus on multiple criteria ranking problems with deterministic performance evaluations, and model the DM's preferences with additive multi-attribute value models [13] defined through holistic pair-wise preference statements (i.e. alternative *a* is (weakly) preferred over *b*).

The holistic judgments may require smaller cognitive effort from the DM in answering questions concerning her preferences than direct elicitation of the value function, e.g., through the bisection method [13]. However, there is typically more than a single value function compatible with the holistic statements. Obviously, the ranking of the alternatives can vary depending on the compatible value function used, and often the set of compatible functions must be reduced in size by introducing additional preference information to obtain a complete preorder or to determine the most attractive alternative.

Robust Ordinal Regression (ROR) [6,4] allows taking into account all instances of a preference model (in our case, the monotone additive value functions) compatible with the provided indirect preference information. These instances do not involve any arbitrary parametrization, so the whole space of compatible value functions can be explored. ROR methods provide the DM with two results, the necessary and possible preference relations for the set of considered alternatives. As far as methods designed for dealing with multiple criteria ranking problems are concerned, ROR has been implemented for the first time in UTA^{GMS} [6] that is a generalization of the UTA method [8]. Apart from considering all compatible value functions rather than just a single one, the UTA^{GMS} does not require the DM's ranking of reference alternatives to be complete and it assumes the use of marginal value functions that are general monotone, and not piece-wise linear. [9] extended the framework to consider all complete preorders compatible with

* Corresponding author. Tel.: +48 61 6653022.
E-mail addresses: milosz.kadzinski@cs.put.poznan.pl (M. Kadziński), tervonen@ese.eur.nl (T. Tervonen).

0377-2217/$ - see front matter
http://dx.doi.org/10.1016/j.ejor.2013.01.022

the preference information and to determine the best and the worst ranks taken by each alternative.

A different way of handling multi-criteria problems having uncertain or imprecise values for the model was proposed in Stochastic Multicriteria Acceptability Analysis (SMAA). These methods apply simulation in order to provide the DM with indices describing the decision problem [23]; in particular, the original SMAA method [14] computes acceptability indices measuring the variety of different preferences that give each alternative the best rank, and SMAA-2 [16] extends it by introducing rank acceptability indices. They indicate the share of weights, criteria measurements, and other model parameters that assign an alternative to any rank from the best to the worst one. Ref. [18] proposed to derive pair-wise winning indices that indicate, for two alternatives, the probabilities of either being on a higher rank.

Both ROR and SMAA fail to consider some important issues. In particular, ROR methods analyse the sets of all, some, or no compatible instances of the preference model and the most and the least advantageous compatible model instances. However, in practical decision making situations, the necessary relation often leaves many pairs of alternatives incomparable, and it is desirable to answer how probable is it for an alternative to be preferred over another. Indication that an alternative could be ranked at its best or worst possible position with very high or extremely low shares of compatible preference models as well as knowing the most likely ranks an alternative can attain may change the preferred alternative of the DM similarly as risk attitude partially defines preference over risky outcomes in multi-attribute utility theory. Consequently, knowing the most and the least probable ranks and the probability of being preferred to another alternative may be valuable for practical decision support. In particular, a low probability of attaining a given rank indicates it to be sensitive for small changes in DM preferences.

SMAA-2 is traditionally applied with linear marginal value functions [27,28,24,19,1]. Such a limitation is arbitrary and restrictive, and it would be desirable for SMAA-2 to be applicable also with general monotone value functions. Furthermore, although SMAA-2 allows DMs to provide holistic preference judgments, they are used solely to derive linear constraints for the weights of the linear marginal value functions, and apart from the scaling, not to derive the piecewise linear functions themselves. Finally, although the rank acceptability indices of SMAA-2 can be estimated to within acceptable error bounds [25], they are not accurate. Therefore, an estimated rank acceptability or pair-wise winning index of 0 cannot be regarded with certainty, because they do not exclude the possibility of the alternative attaining a given position or being preferred over another alternative, respectively. Although the conditions under which such a situation is possible may be very specific, they are still consistent with the preference information provided by the DM. Thus, it is desirable to analyze estimations of the SMAA indices in the context of the necessary, possible, and extreme results of ROR and Extreme Ranking Analysis (ERA) to provide information on which particular outcomes occur with all, some, or no compatible preference models.

In this paper we overcome these shortcomings by combining ROR and SMAA in a joint approach. On the one hand, ROR is enriched by computing how probable are the possible relations. On the other hand, SMAA is extended by considering general instances of the preference model, admitting partial holistic judgments provided in an iterative manner, and confronting the indices estimated through Monte Carlo simulation with the results indicating necessary and possible preference relations and the corresponding extreme ranks.

The organization of the paper is the following. Section 2 presents the new combined approach for multiple criteria ranking problems. Section 3 considers extensions, discussing relations between the provided preference information and the outcomes of the combines approach, and introduces a procedure for selecting a representative value function. Section 4 provides an example application and Section 5 concludes.

2. The combined approach

We use the following notation:

- $A = \{a_1, \ldots, a_i, \ldots, a_n\}$ – a finite set of n alternatives;
- $A^R = \{a^*, b^*, \ldots\}$ – a finite set of reference alternatives on which the DM accepts to express preferences; we assume that $A^R \subseteq A$;
- $G = \{g_1, \ldots, g_j, \ldots, g_m\}$ – a finite set of m evaluation criteria, $g_j : A \rightarrow \mathbb{R}$;
- $X_j = \{g_j(a_i), a_i \in A\}$ – the set of deterministic evaluations on g_j; we assume, without loss of generality, that the greater $g_j(a_i)$, the more desirable is alternative a_i on criterion g_j;
- $x_j^1, \ldots, x_j^{n_j(A)}$ – the ordered values of X_j, $x_j^k < x_j^{k+1}$, $k = 1, \ldots, n_j(A) - 1$, where $n_j(A) = |X_j|$ and $n_j(A) \leqslant n$; consequently, $X = \prod_{j=1}^m X_j$ is the evaluation space.

The DM provides a partial preorder on the set of reference alternatives A^R, denoted by $\succsim$. In particular, the DM can state that a^* is at least as good as b^* ($a^* \succsim b^*$), a^* is indifferent to b^* ($a^* \sim b^*$), or a^* is strictly preferred to b^* ($a^* \succ b^*$). As a preference model, we use the additive value function:

$$U(a) = \sum_{j=1}^{m} u_j(a) \tag{1}$$

where the marginal value functions $u_j\left(x_j^k\right)$, $k = 1, \ldots, n_j(A)$ are monotone, non-decreasing and normalized so that the overall value (1) is bounded within the interval $[0,1]$.

The pair-wise comparisons provided by the DM form the input data for the ordinal regression that finds the whole set of value functions being able to reconstruct these judgments. Such value functions are *compatible* with the preference information. Precisely, a set of general additive value functions $\mathcal{U}_{ROR}^{A^R}$ compatible with the provided pair-wise comparisons is defined with the following set of constraints:

$$\left.\begin{array}{l} U(a^*) \geqslant U(b^*) + \varepsilon, \quad \text{if } a^* \succ b^* \text{ for } a^*, b^* \in A^R, \\ U(a^*) = U(b^*), \quad \text{if } a^* \sim b^* \text{ for } a^*, b^* \in A^R, \\ U(a^*) \geqslant U(b^*), \quad \text{if } a^* \succsim b^* \text{ for } a^*, b^* \in A^R, \\ \left.\begin{array}{l} u_j\left(x_j^k\right) - u_j\left(x_j^{(k-1)}\right) \geqslant 0, \ k = 2, \ldots, n_j(A), \\ u_j\left(x_j^1\right) = 0, \ \sum_{j=1}^{m} u_j\left(x_j^{n_j(A)}\right) = 1, \end{array}\right\} E_{base}^{A^R} \end{array}\right\} E_{ROR}^{A^R}, \tag{2}$$

where ε is an arbitrarily small positive value. If $\varepsilon^* = max\,\varepsilon$, s.t. $E_{ROR}^{A^R}$ is greater than 0 and $E_{ROR}^{A^R}$ is feasible, the set of compatible value functions is non-empty. Otherwise, the provided preference information is inconsistent with the assumed preference model.

2.1. Necessary and possible preference relations

Robust Ordinal Regression applies all compatible value functions $\mathcal{U}_{ROR}^{A^R}$, and defines two binary relations in the set of all alternatives A [6]:

- Necessary weak preference relation, $\succsim^N$, that holds for a pair of alternatives $(a,b) \in A \times A$, in case $U(a) \geqslant U(b)$ for all compatible value functions;
- Possible weak preference relation, $\succsim^P$, that holds for a pair of alternatives $(a,b) \in A \times A$, in case $U(a) \geqslant U(b)$ for at least one compatible value function.

The following linear programs (LPs) need to be solved to assess whether the relations hold:

$$Maximize:\ \varepsilon \tag{3}$$

s.t.

$$\left.\begin{array}{l} U(b) - U(a) \geqslant \varepsilon, \\ E^{A^R}_{ROR}, \end{array}\right\} E^N(a,b)$$

and

$$Maximize:\ \varepsilon \tag{4}$$

s.t.

$$\left.\begin{array}{l} U(a) - U(b) \geqslant 0, \\ E^{A^R}_{ROR}. \end{array}\right\} E^P(a,b)$$

Then, $a \succsim^N b$ if $E^N(a,b)$ is infeasible or $\varepsilon_* = max\,\varepsilon$, s.t. $E^N(a,b)$, is not greater than 0. $a \succsim^P b$ if $E^P(a,b)$ is feasible and $\varepsilon^* = max\,\varepsilon$, s.t. $E^P(a,b)$, is greater than 0.

2.2. Pair-wise outranking indices

Pair-wise outranking index $POI(a,b)$ is, for a pair of alternatives $(a,b) \in A \times A$, the share of compatible value functions for which a is not worse than b. Consequently, for any $(a,b) \in A \times A$:

$$POI(a,b) \in [0,1] \text{ and } POI(a,b) + POI(b,a) \geqslant 1,$$

and for any $a \in A$, $POI(a,a) = 1$. Note that the pair-wise winning index $PWI(a,b)$ [18] is $PWI(a,b) = 1 - POI(b,a)$. The stochastic index POI can be computed exactly only in very small problems, and in what follows we consider its Monte Carlo estimation POI'.

The necessary and possible weak preference relations and pair-wise outranking indices relate to each other as follows:

Remark 2.1. For any pair of alternatives $a,b \in A$:

1. $a \succsim^N b \Rightarrow POI'(a,b) = 1$,
2. $\neg(a \succsim^P b) \Rightarrow POI'(a,b) = 0$,
3. $POI'(a,b) > 0 \Rightarrow a \succsim^P b$; in particular, $POI'(a,b) = 1 \Rightarrow a \succsim^P b$,
4. $POI'(a,b) = 0 \Rightarrow \neg(a \succsim^N b)$.

The truth of the necessary relation and the falsity of the possible relation are the most certain recommendations provided by ROR. They indicate that one alternative is at least as good as the other for all or no compatible value functions, respectively. Thus, any value function sampled in a single Monte Carlo iteration needs to confirm such indication. However, the inverse is not necessarily true, i.e. if $POI'(a,b) = 1$ ($POI'(a,b) = 0$) does not imply that $a \succsim^N b$ ($\neg(a \succsim^P b)$) since the set of value functions $\mathcal{U}^{A^R}_{SMAA} \subset \mathcal{U}^{A^R}_{ROR}$ taken into account in the sample in the current iteration might not contain $U \in \mathcal{U}^{A^R}_{ROR}$ such that $U(b) > U(a)$($U(a) \geqslant U(b)$). Thus, outcomes of ROR may enrich an indication provided by SMAA with statements that are absolutely sure, certainly not true, or possible.

On the other hand, if a and b are related by the necessary incomparability (i.e., $\neg(a \succsim^N b)$ or $\neg(b \succsim^N a)$), there exist:

$$U_1, U_2 \in \mathcal{U}^{A^R}_{ROR}, \text{ such that } U_1(a) > U_1(b) \text{ and } U_2(a) < U_2(b). \tag{5}$$

If such ambiguity in the comparison for some pairs of alternatives occurs, it is useful to investigate the share of compatible value functions for which a is at least as good as b, and vice versa (i.e., $POI'(a,b)$ and $POI'(b,a)$). In this way, we would be able to state whether a definite possible relation is "almost sure", "sure on average", "barely sure", or "almost not possible at all". In particular, analyzing the difference between $POI'(a,b)$ and $POI'(b,a)$ may be helpful in designating a better alternative.

2.3. Extreme ranks

Definition 1 (*Ranking function*). The rank of an alternative a relative to all alternatives in A is defined with the ranking function

$$rank(U,a) = 1 + \sum_{b \in A \setminus \{a\}} h(U,a,b), \quad \text{where} \tag{6}$$

$$h(U,a,b) = \begin{cases} 1, & \text{if } U(b) > U(a) \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

To identify the best $P^*(a) = \max_{U \in \mathcal{U}^{A^R}_{ROR}}(rank(U,a))$ and the worst $P_*(a) = \min_{U \in \mathcal{U}^{A^R}_{ROR}}(rank(U,a))$ ranks that a particular alternative $a \in A$ can attain, [9] proposed ERA consisting of the following Mixed-Integer Linear Programming (MILP) models:

$$Minimize:\ f^{rank}_{max} = \sum_{b \in A \setminus \{a\}} v_b \tag{8}$$

s.t.

$$\left.\begin{array}{l} U(a) - U(b) + M v_b \geqslant 0, \text{ for all } b \in A \setminus \{a\}, \\ E^{A^R}_{ROR}, \end{array}\right\} E^{A^R}_{max},$$

and

$$Minimize:\ f^{rank}_{min} = \sum_{b \in A \setminus \{a\}} v_b \tag{9}$$

s.t.

$$\left.\begin{array}{l} U(b) - U(a) + M v_b \geqslant \varepsilon, \text{ for all } b \in A \setminus \{a\}, \\ E^{A^R}_{ROR}, \end{array}\right\} E^{A^R}_{min},$$

where M and ε are auxiliary variables equal to, respectively, big and small positive values, and v_b is a binary variable associated with comparison of a to alternative $b \in A \setminus \{a\}$. The best rank $P^*(a)$ of alternative a is $f^{rank}_{max} + 1$ and the worst rank $P_*(a)$ is $n - f^{rank}_{min}$.

2.4. Rank acceptability indices

The rank acceptability index $RAI(a,k) \in [0,1]$, for alternative $a \in A$ and rank $k = 1,\ldots,n$, is defined as the expected share of compatible value functions that grant alternative a rank k. Similarly to POI, RAI can only be computed exactly in very small problems and therefore in what follows we consider its Monte Carlo estimation RAI'. The extreme ranks and rank acceptability indices relate to each other as follows:

Remark 2.2. For any alternative $a \in A$:

1. $k \in [1,P^*(a)) \cup (P_*(a),n] \Rightarrow RAI'(a,k) = 0$,
2. $\sum_{k=P^*(a)}^{P_*(a)} RAI'(a,k) = 1$,
3. $RAI'(a,k) > 0 \Rightarrow P^*(a) \leqslant k$ and $k \leqslant P_*(a)$.

Extreme Ranking Analysis indicates the range of ranks $[P^*(a), P_*(a)]$ that can be attained by each alternative $a \in A$ when taking into account all compatible value functions $\mathcal{U}^{A^R}_{ROR}$. For ranks k outside this range, $RAI'(a,k)$ is 0. For ranks within this range, $RAI'(a,k)$ may be greater than 0, but is not necessarily greater than 0. Nevertheless, for each $a \in A$, the sum of rank acceptability indices for ranks between $P^*(a)$ and $P_*(a)$ is equal to 1. One should also bear in mind that the range of possible ranks indicated by ERA may be wider than the range of ranks in the SMAA indices, so only the recommendations of ERA should be treated with certainty.

Furthermore, if we limit $\mathcal{U}^{A^R}_{ROR}$ to exclude value functions for which two alternatives obtain an equal rank, the following Theorem holds:

Theorem 2.1 (No rank jumps). *Assume there are no shared ranks, i.e.* $\forall U \in \mathcal{U}^{A^R}_{ROR}, \forall (a_i, a_j) \in A \times A, i \neq j : U(a_i) \neq U(a_j)$. *Now, if* $\exists U',\ U'' \in \mathcal{U}^{A^R}_{ROR} : rank(U', a_i) = \rho, rank(U'', a_i) = \rho + \delta^\rho$ *with* $\delta^\rho \geqslant 2 \Rightarrow \forall \delta \in \{0, \ldots, \delta^\rho\} \exists U^* \in \mathcal{U}^{A^R}_{ROR} : rank(U^*, a_i) = \rho + \delta$

Proof. By induction on δ^ρ; the base case is $\delta^\rho = 2$. Let $a_i = a_1$ and a_2, a_3 be the alternatives that "overtake" a_1 in the ranking, i.e. $U'(a_1) > U'(a_2)$, $U'(a_1) > U'(a_3)$, $U''(a_1) < U''(a_2)$, $U''(a_1) < U''(a_3)$. Now, as $\mathcal{U}^{A^R}_{ROR}$ is convex (see the proof of Proposition 4.5 in [7]), $\mathcal{U}^{A^R}_{ROR} \ni U^*(a) = (1-\alpha)U'(a) + \alpha U''(a) \forall \alpha \in [0,1]$; particularly $\alpha = 0 \Rightarrow U^* = U'$, $\alpha = 1 \Rightarrow U^* = U''$. As there are no shared ranks, $\exists \alpha \in (0,1)$ such that either (i) $U^*(a_1) > U^*(a_2)$ and $U^*(a_1) < U^*(a_3)$, or (ii) $U^*(a_1) < U^*(a_2)$ and $U^*(a_1) > U^*(a_3)$ (i.e. exactly one of the alternatives a_2 and a_3 overtakes a_1 at once) $\Rightarrow \exists U^* \in \mathcal{U}^{A^R}_{ROR} : rank(U^*, a_1) = \rho + 1$. Induction assumption is that the theorem holds for $\delta^\rho = \Delta \geqslant 2$. Now, to prove that the Theorem holds for $\delta^\rho = \Delta + 1$, proceed analogously to $\delta^\rho = 2$. □

The assumption concerning no shared ranks is natural as otherwise the extreme ranks P^* and P_* would themselves be hard to interpret. In addition, the multiple simultaneous rank jumps occur rarely in practice and require an extremely constrained $\mathcal{U}^{A^R}_{ROR}$, e.g. with an indifference statement $a_2 \sim a_3$ that reduces $\mathcal{U}^{A^R}_{ROR}$ to an isopreference hyperplane, and then if $\exists a_1 \in A \exists u',\ u'' \in \mathcal{U}^{A^R}_{ROR} : u'(a_1) > u'(a_2) = u'(a_3)$ and $u''(a_1) < u''(a_2) = u''(a_3) \Rightarrow$ the rank of a_1 jumps by 2 when it simultaneously passes a_2 and a_3. Therefore, Theorem 2.1 implies that

(i) in practice we do not need to assess whether an alternative can obtain ranks between the extreme ones; assuming no shared ranks, an alternative can obtain them all, and
(ii) if the SMAA analysis indicates that the probability of an alternative a attaining rank k, such that $P^*(a) \leqslant k \leqslant P_*(a)$, is equal to zero (i.e. $RAI'(a,k) = 0$), this is most probably due to inaccuracy of the simulated index RAI' and actually $\exists U \in \mathcal{U}^{A^R}_{ROR} : rank(U, a) = k$.

Although ERA is focused on indicating a range of possible ranks, in practical decision making situations such a range may be very wide. Moreover, a binary answer with respect to the possibility of ranking a in the kth position may be insufficient to arrive at a recommendation. Thus, SMAA may enrich ERA with answering questions about the distribution of ranks between the extreme ones, the most probable range of positions for a given alternative, the expected position, or the probability of being ranked the best. In particular, knowing the distribution of ranks for a and b makes designating the better alternative easier. In particular, it may even change the result of the comparison of $[P^*(a), P_*(a)]$ and $[P^*(b), P_*(b)]$ in terms of interval orders (see e.g. [21]) because for one of these alternatives, the probability of being ranked at its best or worst position may be considered by the DM as negligibly low.

SMAA models are usually estimated through Monte Carlo simulation by sampling from the criteria-and weight distributions. In this paper, we consider only deterministic criteria measurements for compatibility with the ROR Linear Programming (LP) models. Computing the SMAA indices exactly would require us to (1) compute the volume of the full space exactly, (2) for each alternative, to partition the space in regions that grant an alternative a certain rank, and (3) to compute the volumes of these regions. Exact computation of the volume of a polytope is P#-hard [17]. Although Markov Chain Monte Carlo (MCMC) methods for volume estimation exist [12], they are complex, and a better use for the MCMC techniques is to apply them in sampling the actual value functions; [26] proposed a method for sampling weights with linear constraints, which can be formulated in terms of the reference alternatives in case linear marginal value functions are used. There do not exist efficient algorithms for sampling the complete value functions with non-linear marginal value functions, but a simple rejection sampling is feasible in low dimensionality ($n < 8$) problems. The SMAA-O technique [15] normally used for generating ordered weights can be applied [25] for generating the candidate draw marginal value functions. When pure rejection sampling is applied, the rejection rate grows exponentially with respect to the number of dimensions but only polynomially regarding the amount of preference statements.

Obviously, acceptable error limits for the stochastic indices can always be achieved, given sufficient computation time, with a certain amount of Monte Carlo iterations. Nevertheless, although the simulation results may be precise with high confidence, they can still be unstable with respect to changes in the set of pair-wise preference statements provided by the DM. Although the outcomes of both ROR and SMAA should be interpreted in the context of the provided preference information, analyzing the influence of different sets of preference statements may be useful for decision support in practice.

3. Extensions

3.1. Relation between the holistic preference statements and the analysis outcomes

The joint application of SMAA, ROR and ERA is designed for incremental specification of pair-wise comparisons. In this way, the DM can easily assess the impact of each of her/his statements on the final outcome. The provided results are designed to enhance the interaction between DM and the analyst. The suggested decision support process is to first analyze the necessary and possible preference relations, and in the following iterations enrich the preference relation for pairs (a,b) for which the relation was possible, but not necessary (i.e., $a \succsim^P b$, and $\neg(a \succsim^N b)$). Changing the possible preference relation into a necessary one can be supported through an analysis of $POI'(a,b)$. The DM may wish to state $a \succsim b$ if $POI'(a,b)$ is close to one, and indicate $\neg(a \succsim b)$ if $POI'(a,b)$ is close to zero. In this way, the proposed approach addresses the difficulty of supplying a large set of preference statements at once.

Note that the preference information provided by the DM is directly translated into the outcomes of ROR, ERA and SMAA in the following way:

Remark 3.1. For any reference alternatives $a^*, b^* \in A^R$:

$$a^* \succsim b^* \Rightarrow \begin{cases} a^* \succsim^N b^* \quad \text{and } POI'(a^*, b^*) = 1, \\ P^*(a^*) \leqslant P^*(b^*) \quad \text{and } P_*(a^*) \leqslant P_*(b^*), \\ \forall i \in [1, \ldots, n], \quad \sum_{k=1}^{i} RAI'(a^*, k) \geqslant \sum_{k=1}^{i} RAI'(b^*, k), \\ \forall j \in [1, \ldots, n], \quad \sum_{k=j}^{n} RAI'(a^*, k) \leqslant \sum_{k=j}^{n} RAI'(b^*, k), \end{cases} \tag{10}$$

and

$$a^* \succ b^* \Rightarrow \begin{cases} \neg(b^* \succsim^P a^*) \quad \text{and } POI'(b^*, a^*) = 0, \\ P^*(a^*) < P^*(b^*) \quad \text{and } P_*(a^*) < P_*(b^*), \\ \exists i \in [1, \ldots, n], \quad \sum_{k=1}^{i} RAI'(a^*, k) > \sum_{k=1}^{i} RAI'(b^*, k), \\ \exists j \in [1, \ldots, n], \quad \sum_{k=j}^{n} RAI'(a^*, k) < \sum_{k=j}^{n} RAI'(b^*, k). \end{cases} \tag{11}$$

Thus, by providing holistic pair-wise preference statements the DM can directly enrich the necessary relation or impoverish the possible relation. At the same, the number of pairs of alternatives $(a,b) \in A \times A$, for which the indication of $POI'(a,b)$ is unambiguous (i.e., $POI'(a,b) = 1$ or $POI'(a,b) = 0$) increases. Obviously, the provided preference information influences the relations between the extreme ranks for the pairs of compared reference alternatives. Finally, the relations $a^* \{\succsim, \succ\} b^*$ are reflected in the sum of RAIs corresponding to some i best ranks (i.e., $RAI'(x,1), \ldots, RAI'(x,i)$) and the sum of RAIs for some $n-j+1$ last ranks (i.e., $RAI'(x,j), \ldots, RAI'(x,n)$), with $i,j \in [1,\ldots,n]$ being any (in case $a^* \succsim b^*$) or some particular (in case $a^* \succ b^*$) ranks, such that $i \leqslant j$.

Let us consider incremental specification of preference information and denote with $B_1^R \subseteq B_2^R \subseteq \ldots \subseteq B_s^R$ nested sets of DM's pair-wise preference statements on the reference alternatives. Each of these sets B_t^R, $t = 1, \ldots, s$, is modeled with a set of constraints $E_{t,ROR}^{A^R}$ generating the set of compatible value functions $\mathcal{U}_{t,ROR}^{A^R}$. These are nested in an order inverse with respect to the related sets of pair-wise statements B_t^R, $t = 1, \ldots, s$, i.e. $\mathcal{U}_{1,ROR}^{A^R} \supseteq \mathcal{U}_{2,ROR}^{A^R} \supseteq \ldots \supseteq \mathcal{U}_{s,ROR}^{A^R}$. We will consider nested sets of pair-wise comparisons only until the generated sets of compatible value functions are not empty. For each iteration t, we can compute the corresponding possible and necessary weak preference relations $\succsim_t^P$ and $\succsim_t^N$ as well as the extreme ranks denoted as $P_{*,t}(a)$, $P_t^*(a)$. This requires employment of the procedures presented in Section 2 applied to the set $E_{t,ROR}^{A^R}$. An important property of these outcomes is stated in Proposition 3.1. Its proof is omitted for being obvious. Let us only note that it relies on the observation that each compatible value function considered in iteration t has been already considered in iteration $t-1$.

Proposition 3.1. *$\succsim_t^N$ and $\succsim_t^P, t = 1, \ldots, s$, are nested relations: $\succsim_{t-1}^N \subseteq \succsim_t^N$ and $\succsim_{t-1}^P \supseteq \succsim_t^P$, $t = 2, \ldots, s$. $[P_t^*(a), P_{*,t}(a)]$ are nested intervals $[P_t^*(a), P_{*,t}(a)] \subseteq [P_{t-1}^*(a), P_{*,t-1}(a)]$, $t = 2, \ldots, s$.*

With SMAA we cannot indicate general relations between indices computed in the subsequent iterations, because each iteration requires new Monte Carlo simulations and there is no guarantee that the sets of compatible value functions sampled in these simulations are related by inclusion.

Note that our approach is the constructivist one for Multiple Criteria Decision Aiding (MCDA) [20,22]. That is, we do not consider the DM's preference structure to be a pre-existing entity that needs to be discovered, but instead, we assume that a preference model has to be built in course of an interaction between the DM and the analyst. Confronting the DM's value system with the results obtained from applying the inferred model on a set of alternatives enables her to gain insight on her own preferences, and also allows her to understand better the employed method. For example, when pairs of alternatives are incomparable in terms of the necessary relation, we can exhibit the pair-wise outranking indices, and by providing this additional information, encourage the DM to supply statements for pairs of alternatives not present in the relation. Whether she chooses to provide the statements or not and which alternative she prefers remains her own decision.

3.2. Selection of a representative value function

Greco et al. [5,10] introduced the notion of a single representative value function that can be used to help the DM understand the results and to provide a complete ranking of the alternatives. The representative function is computed by maximizing differences between comprehensive values of pairs of alternatives $(a,b) \in A \times A$ related by the necessary preference relation (i.e., $a \succ^N b \iff a \succsim^N b$ and $\neg(b \succsim^N a)$), and minimizing differences between comprehensive values of pairs of alternatives $c,d \in A$ related by the necessary incomparability relation (i.e., $c ?^N d \iff \neg(c \succsim^N d)$ and $\neg(d \succsim^N c)$).

Combining SMAA and ROR allows indication of the desired difference between comprehensive values of specific alternatives. For a pair of alternatives $a,b \in A$, this intensity may be connected with the comparison between pair-wise winning indices $PWI'(a,b)$ and $PWI'(b,a)$, i.e. the shares of value functions confirming the advantage of a with respect to b, and vice versa. Precisely, one may select a value function promoting alternatives designated as better ones in the comparisons over all pairs of alternatives, and we would require that $U^R(a) > U^R(b)$, if $PWI'(a,b) > PWI'(b,a)$. The following procedure (let us call it REPROC) for selecting a representative value function corresponds to the above interpretation of representativeness:

1. For all $a,b \in A$, such that $PWI'(a,b) > PWI'(b,a)$, add the following constraints to the set of constraints $E_{ROR}^{A^R}$:

$$\begin{cases} U(a) - U(b) \geqslant \varepsilon(a,b), \\ \varepsilon(a,b) \geqslant \varepsilon. \end{cases} \tag{12}$$

2. Maximize ε s.t. the modified $E_{ROR}^{A^R}$.
3. Add the constraint $\varepsilon = \varepsilon^*$, with $\varepsilon^* = \max \varepsilon$ from the previous point, to the modified $E_{ROR}^{A^R}$.
4. Maximize $\sum_{a,b:PWI'(a,b)>PWI'(b,a)} \varepsilon(a,b)$ s.t. the modified $E_{ROR}^{A^R}$.
5. Read the representative comprehensive values $U^R(a)$ and corresponding marginal values from the solution of the LP problem considered in point 4.

Note that REPROC can be suitably adapted with respect to the preferences of the DM, for example, to emphasize the advantage of a over b only if the difference $PWI'(a,b) - PWI'(b,a)$ exceeds a pre-defined threshold $t_{PWI} > 0$. Such a threshold would be then used in REPROC by replacing the condition $PWI'(a,b) > PWI'(b,a)$ with $PWI'(a,b) - PWI'(b,a) \geqslant t_{PWI}$. Furthermore, inaccuracy of the PWI' computation can be accounted for by choosing a suitable t_{PWI}, e.g. 0.02.

4. Application

We analyze a real-world problem, originally considered in Webometrics [2], of ranking 20 European countries based on the quality of their universities. The quality is measured with respect to the presence of the universities in the web. The following four criteria are considered: g_1 (system), the number of universities in the Top 500 in the given country, divided by the mean position of these institutions; g_2 (access), a score built according to ranks (five points for a university in the Top 100, four points for 101–200, etc., divided by the population size of the country); g_3 (flagship), a normalized score based on the leading university rank for countries; and g_4 (economy), same score as for g_3 but divided by GDP per capita. The performance matrix is provided in Table 1.

Simulation of the RAI' and POI' indices was implemented via rejection sampling, and computation of the ROR indices by constructing LPs that were subsequently solved with gplk. The rejection sampling test code is freely available online[1] and the sampling procedure is provided as pseudo-code in Appendix A. In addition, the UTA$^{\text{GMS}}$, POI' and RAI' computation are implemented in the open source 'ror' R package available from CRAN.[2] For illustrative purpose, we also analyzed the problem with, in addition to the general value functions, with linear and two-piece linear ones. For

[1] http://github.com/tommite/pubs-code/tree/master/rorsmaa-ejor/.

[2] http://cran.r-project.org/web/packages/ror/index.html.

Table 1
Performance matrix for the problem of ranking European countries according to the presence of their universities in the web.

Country	Short name	g_1	g_2	g_3	g_4
Germany	GER	82	94	80	91
United Kingdom	UK	74	91	96	82
Spain	SPA	59	73	72	67
Sweden	SWE	47	77	90	46
Netherlands	NET	50	73	88	47
Italy	ITA	51	50	84	55
Finland	FIN	42	59	88	39
Belgium	BEL	44	57	84	41
Austria	AUS	42	53	88	38
Denmark	DEN	42	61	68	39
France	FRA	45	37	80	44
Czech Rep.	CZE	41	43	80	40
Portugal	POR	41	41	60	40
Slovenia	SLO	38	37	72	34
Ireland	IRE	40	40	60	34
Hungary	HUN	39	34	48	38
Estonia	EST	38	36	44	34
Greece	GRE	39	28	40	34
Poland	POL	39	26	36	37
Slovakia	SVK	37	21	8	37

the stochastic model, the candidate draws with the marginal linear value functions can be computed with standard SMAA-2 simulation technique by just varying the weights (see [25]). For the models with 2-piece linear marginal value functions, the candidate draws are sampled by generating the marginal value function middle point values (i.e. v (0.5)) uniformly from]0,1[. Then the marginal functions start from 0, end at 1, have the sampled middle point value, and are scaled with uniformly generated weights (similarly to sampling the candidate draws with linear marginal value functions).

4.1. Preference information

We apply three pair-wise comparisons of reference alternatives based on the actual ranks of these alternatives as determined by Webometrics:

$$\text{Denmark} \succ \text{Austria}, \quad \text{Spain} \succ \text{Sweden}, \quad \text{France} \succ \text{Czech Rep.}$$

These are consistent with respect to the considered additive model when assuming linear, two-piece linear, and general monotone non-linear value functions. Let us denote the sets of compatible additive value functions composed of linear marginal value functions by $\mathcal{U}_1^{A^R}$, of two-piece linear ones by $\mathcal{U}_2^{A^R}$, and of general monotone ones by $\mathcal{U}_G^{A^R}$. Obviously, the following inclusion relations are satisfied:

$$\mathcal{U}_1^{A^R} \subseteq \mathcal{U}_2^{A^R} \subseteq \mathcal{U}_G^{A^R}. \tag{13}$$

In the following subsections, we will discuss in detail results obtained with the general value functions, and subsequently compare them with those obtained with the linear and two-piece linear ones.

4.2. Necessary/Possible Preference Relations and Pair-wise Outranking Indices

For the provided preference information, the necessary and possible weak preference relation are computed using UTA^{GMS}. The Hasse diagram of the necessary relation $\succ_G^N$ assuming the use of general monotone marginal value functions is marked with solid arcs in Fig. 1. There are 149 pairs of alternatives $(a,b) \in A \times A$ related by the necessary preference $\succ_G^N$ (e.g., (Germany, Spain), (Italy, Portugal)). These relations are robust with respect to the given preference information. When analyzing the necessary relation, Germany and United Kingdom should be perceived as the best alternatives, Spain, Netherlands, Italy and Sweden should be viewed as relatively good countries, whereas Poland, Greece, Estonia, and Slovakia need to be considered as the worst ones.

When using general value functions, the degree of freedom in assessing compatible instances of the preference model is the greatest. Using the most flexible model, one obtains also the most general results. Hence, when taking into account two-piece linear marginal functions, it is easier to satisfy the necessary relation as the set of compatible value functions is smaller, and $\mathcal{U}_2^{A^R} \subseteq \mathcal{U}_G^{A^R} \Rightarrow \succ_2^N \supseteq \succ_G^N$. In Fig. 1, the necessary relations that become true when limiting the set of compatible value function from general to two-piece linear are marked with dashed arcs. Precisely, there are three such relations: (SPA, ITA), (SPA, NET), (ITA, BEL). Further, when considering linear marginal value functions the necessary relation $\succ_1^N$ is even richer: there are another five pairs of alternatives (i.e., (GER, UK), (ITA, FIN), (ITA, DEN), (BEL, AUS), (POR, SLO)) for which $\succ_1^N$ is true whereas $\succ_2^N$ is false. These relations are marked with dotted arcs in Fig. 1. In general, the necessary relation corresponding to different types of marginal value functions is nested as follows:

$$\succ_1^N \supset \succ_2^N \supset \succ_G^N. \tag{14}$$

If the necessary relation holds for a given pair of alternatives, then the possible relation holds as well. However, for the case of general value functions there are 41 pairs of alternatives $(c,d) \in A \times A$ related by the necessary incomparability $?^N$, i.e., $\neg\left(c \succsim_G^N d\right)$ and $\neg\left(d \succsim_G^N c\right)$ (e.g., (Germany, UK), (Finland, Denmark), (Slovenia, Slovakia)). The nodes corresponding to these countries are not related by an arc (neither directly nor when considering transitivity of the necessary relation) in Fig. 1. The original ROR methods leave these pairs of alternatives equally desirable, stating that there is at least one compatible value function for which c is ranked better than d, and at least one compatible value function for which this order is reversed. With our combined approach we are able to estimate the share of compatible value functions that confirm the possible preference relation. In Fig. 2, we present pair-wise outranking indices $POI'(c,d)$ and $POI'(d,c)$ for pairs of alternatives $(c,d) \in A \times A$ related by $?_G^N$. When considering the nodes corresponding to a particular pair (c,d), we indicate with a smaller (greater) head of the arc alternative c (d) for which the result of the SMAA-based comparison is positive (negative) (i.e., $POI'(c,d) > POI'(d,c)$). The values of the indices are provided near the corresponding heads of the arc. In particular, for (Germany, UK), $POI'(\text{Germany},\text{UK}) = 78.11\%$ and $POI'(\text{UK},\text{Germany}) = 21.89\%$.

When it comes to indicating the best country, analysis of pair-wise outranking indices supports Germany rather than UK (i.e., $POI'(\text{Germany},\text{UK}) > POI'(\text{UK},\text{Germany})$). When comparing the subset of the worst alternatives, for a majority of compatible value functions Slovakia is ranked worse than all other alternatives that are not necessarily preferred to any other alternative. Furthermore, Ireland and Hungary compare positively to Slovenia, despite they are placed at a lower level in the necessary relation-based ranking. Note that for some pairs of alternatives, designating the better alternative on the basis of *POI*s is straightforward (see, e.g., (Spain, Italy), (Portugal, Slovenia)). For some other pairs, such indication is ambiguous, since *POI*s do not differ significantly (see, e.g., (Italy, Sweden), (Poland, Greece)). Finally, although the pair-wise outranking index for the pair (Germany, Italy) is equal to 1, one needs to bear in mind that according to ROR there are some compatible value functions for which Italy is ranked better than Germany, and therefore the necessary weak preference relation does not hold for this pair of alternatives.

As far as pair-wise outranking indices for different types of marginal value functions are concerned, for clarity of presentation, we will focus on six alternatives that could be considered as potential

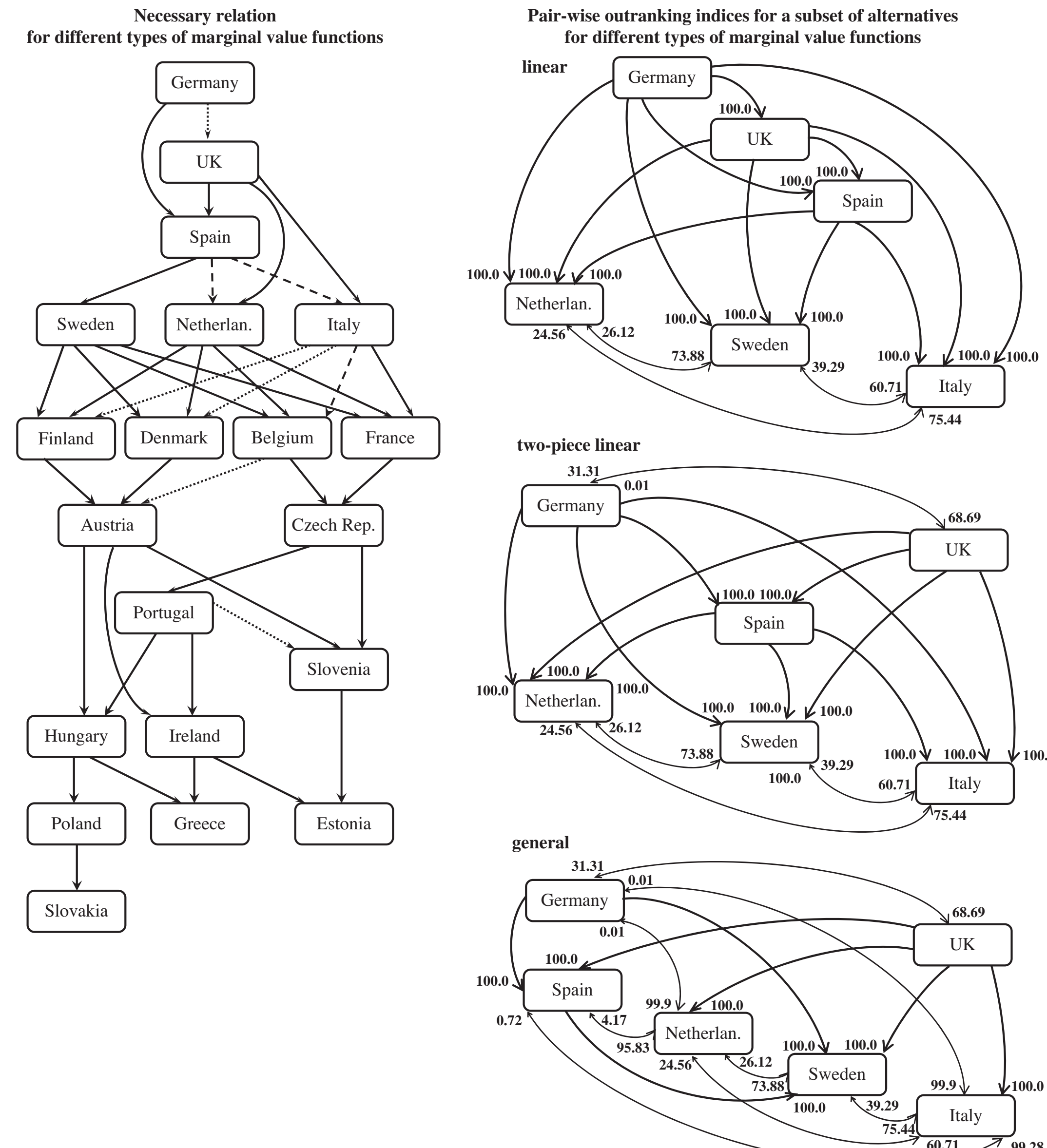

Fig. 1. Hasse diagram of the necessary relation when considering marginal value functions which are linear ($\succ_1^N$-solid, dashed, and dotted), two-piece linear ($\succ_2^N$-solid and dashed), or general ($\succ_G^N$-solid)-to the left. Note that the relation is transitive, and the arcs obtained through transitive closure are omitted. Pair-wise outranking indices (in %) for a subset of alternatives for different types of marginal value functions-to the right.

best options. In Fig. 1 (to the right), we present pair-wise outranking indices corresponding to the different types of marginal value functions. For all linear value functions Germany is preferred to UK, for two-piece linear functions there is a low probability that UK may be ranked better, and with general value functions Germany is ranked lower than UK already for over 20% of the compatible value functions. Moreover, for some pairs of alternatives, even designation of the better country depends on the underlying preference model. In case of Sweden and Italy, the majority of compatible linear value functions support Sweden, for two-piece linear functions the difference between pair-wise outranking indices for these countries is negligible, whereas analysis conducted with general value functions favors Italy.

4.3. The representative value function

The representative value function composed of marginal general value functions, which has been selected with the REPROC procedure, is presented in Fig. 3. The corresponding comprehensive values are provided in Table 2, column U_R. For all pairs of alternatives this function assigns a larger value to the alternative with a higher pair-wise winning index (i.e., $U_R(a) > U_R(b)$, for all $a,b \in A$, such that $PWI'(a,b) > PWI'(b,a)$). In particular, when considering the complete ranking determined by U_R, Germany is ranked first with score 1.0, because $PWI'(\text{Germany},b) > PWI'(b,\text{Germany})$, for all $b \in A \setminus \{\text{Germany}\}$. On the contrary, Slovakia is ranked at the very bottom, since it looses pair-wise comparisons with all the remaining alternatives, i.e., $PWI'(\text{Slovakia},b) < PWI'(b,\text{Slovakia})$, for all $b \in A \setminus \{\text{Slovakia}\}$.

Note that the selected value function "flattens" the consequences of applying the set of compatible value functions sampled by SMAA. In this way, it extends the capacity of the proposed approach in explaining the outputs in terms of a value function that could be displayed to the DM. Then, the DM can see a score for each alternative, and easily assess relative importance of the criteria understood as a share of a given criterion in the comprehensive value.

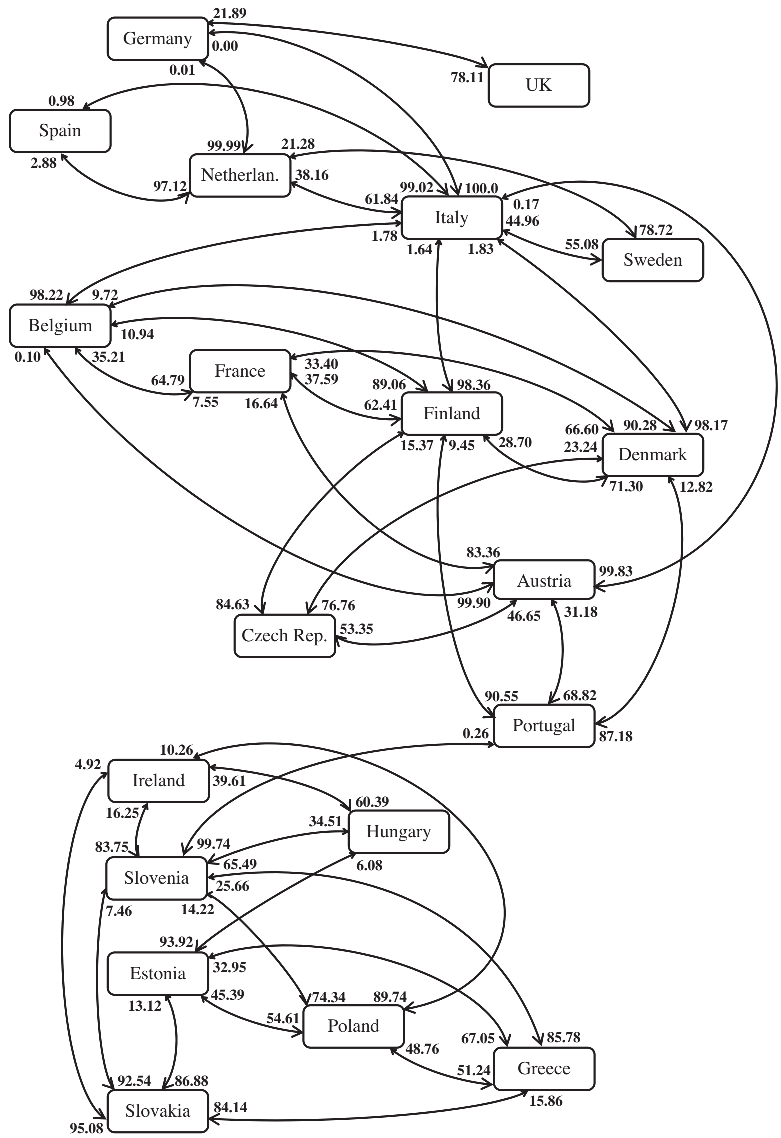

Fig. 2. Possible preference relations (arcs) and the corresponding pair-wise outranking indices (numbers, in %) for the problem of ranking European countries when considering general marginal value functions.

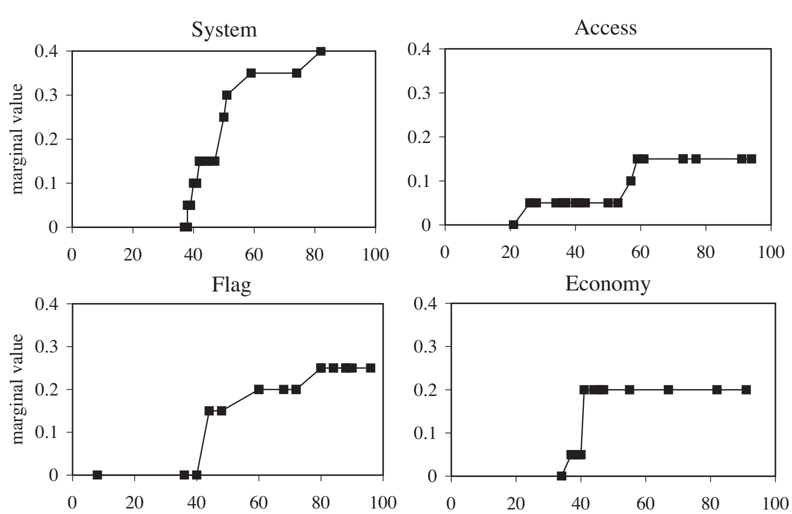

Fig. 3. Representative value function for the problem of ranking European countries according to the presence of their universities in the web.

Table 2
Representative values assuming the use of general value functions. Extreme Ranking Analysis with respect to marginal value functions that are linear (1), two-piece linear (2), or general (G).

Country	U_R	P_1^*	$P_{*,1}$	P_2^*	$P_{*,2}$	P_G^*	$P_{*,G}$
GER	1.00 (1)	1	1	1	2	1	4
UK	0.95 (2)	2	2	1	2	1	2
SPA	0.90 (3)	3	3	3	3	3	5
SWE	0.75 (6)	4	6	4	6	4	6
NET	0.85 (4)	4	6	4	6	2	6
ITA	0.80 (5)	4	6	4	9	2	10
FIN	0.60 (9)	7	12	6	12	6	12
BEL	0.70 (7)	7	8	7	10	6	11
AUS	0.50 (11)	10	13	8	13	8	13
DEN	0.55 (10)	7	12	6	12	6	12
FRA	0.65 (8)	7	11	7	11	7	11
CZE	0.45 (12)	9	12	9	12	9	12
POR	0.40 (13)	10	13	10	14	10	14
SLO	0.25 (16)	14	19	13	19	13	19
IRE	0.35 (14)	14	18	14	18	14	18
HUN	0.30 (15)	14	17	14	17	14	17
EST	0.20 (17)	16	20	16	20	16	20
GRE	0.10 (19)	16	20	16	20	16	20
POL	0.15 (18)	15	19	15	19	15	19
SVK	0.05 (20)	16	20	16	20	16	20

4.4. Extreme ranks and rank acceptability indices

Table 2 shows the best and the worst ranks of each alternative $a \in A$ for all compatible value functions when considering different types of marginal value functions. For the general ones, Germany and UK are potential top alternatives, but Germany is more sensitive to the choice of a compatible value function because its rank may drop to 4. Spain, Sweden, Netherlands, and Italy possibly take place in the top 5. However, among these four countries only Spain never falls out top 5, and in the worst case Italy may be ranked only 10th. Another 5 countries (Belgium, Austria, Denmark, France, and Czech Rep.) are always ranked in the middle of the ranking (i.e., $P_G^*(a) \geqslant 6$ and $P_{*,G}(a) \leqslant 15$). Estonia, Greece, Poland, and Slovakia are the least ranked alternatives though Poland is never ranked at the very bottom. The distributions of extreme ranks are provided in Table 3.

As suggested in [9], when analyzing the extreme ranks for all alternatives, one can easily work out a recommendation in terms of a multiple criteria choice problem. In this case, it is sufficient to define the limits on the best and/or the worst ranks that potentially best options need to attain. Furthermore, on the basis of the extreme ranks, we could compare alternatives in terms of ranking intervals. In particular, we could indicate the preference of a over b, if $P_G^*(a) < P_G^*(b)$ and $P_{*,G}(a) < P_{*,G}(b)$ (e.g., (Germany, Italy), (Belgium, Austria)), or state indifference between a and b, if $\left[P_G^*(a), P_{*,G}(a)\right] \subset \left[P_G^*(b), P_{*,G}(b)\right]$ (e.g., (Spain, Netherlands), (Slovenia, Hungary)).

The ranges of possible ranks for different types of marginal value functions are nested as follows (see Table 2), for each $a \in A$:

$$\left[P_1^*(a), P_{*,1}(a)\right] \subseteq \left[P_2^*(a), P_{*,2}(a)\right] \subseteq \left[P_G^*(a), P_{*,G}(a)\right]. \quad (15)$$

In particular, when using linear marginal value functions, Germany is always ranked first. When admitting two-piece linear or general functions, in the worst case Germany could be ranked second or fourth, respectively. Furthermore, Italy could be ranked in three, six, or nine different positions depending on the underlying preference model. Nevertheless, for some countries (see, e.g., France,

Table 3
Distribution of extreme ranks.

Range	$1 \leqslant P_G^*(a) \leqslant 5$	$6 < P_G^*(a) \leqslant 10$	$11 < P_G^*(a) \leqslant 15$	$16 \leqslant P_G^*(a) \leqslant 20$
Number of countries	6	7	4	3
Range	$1 \leqslant P_{*,G}(a) \leqslant 5$	$6 < P_{*,G}(a) \leqslant 10$	$11 < P_{*,G}(a) \leqslant 15$	$16 \leqslant P_{*,G}(a) \leqslant 20$
Number of countries	3	3	7	7

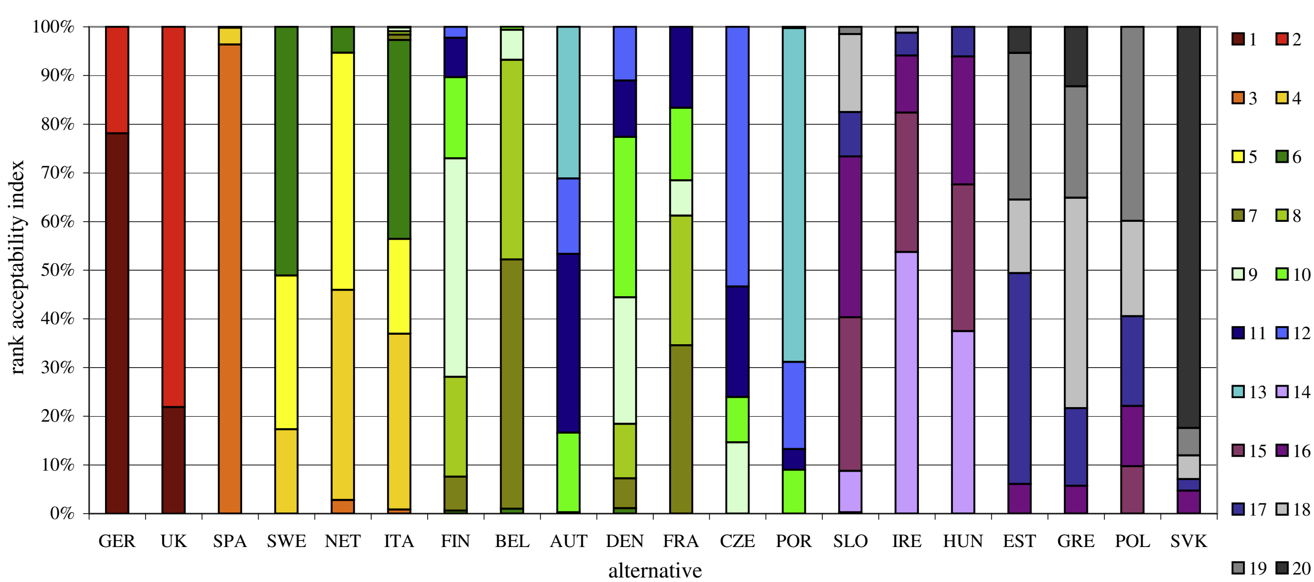

Fig. 4. The rank acceptability indices (see Table B.4 in Appendix B for the numerical values).

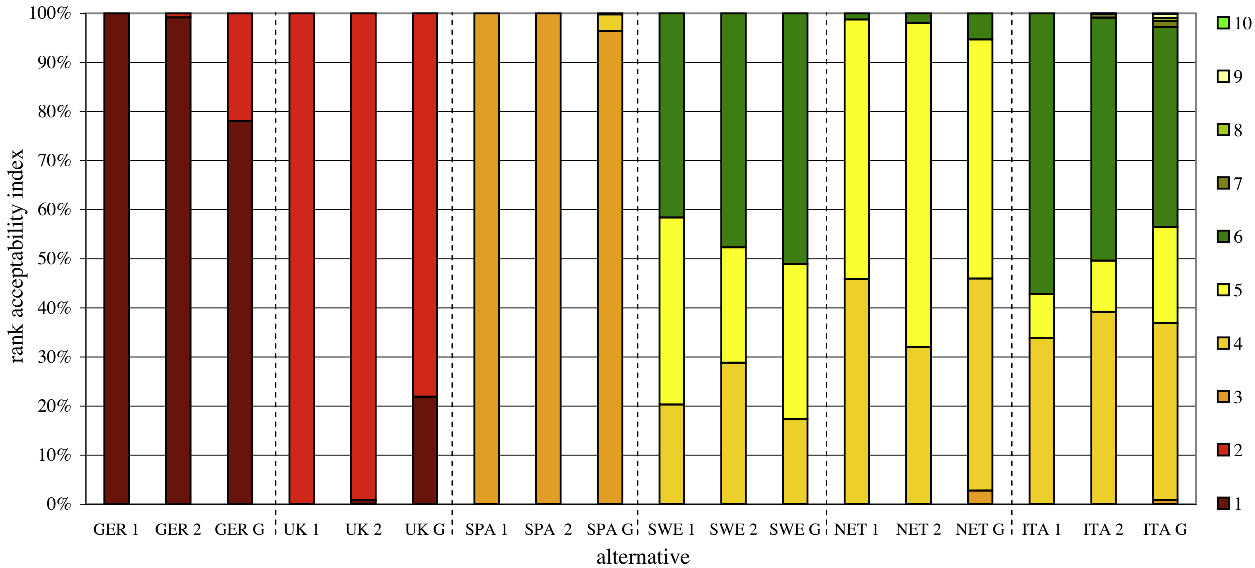

Fig. 5. Rank acceptability indices for six chosen alternatives and different types of marginal value functions (1 – linear, 2 – two-piece linear, G – general) (see Table B.5 in Appendix B for the numerical values).

Hungary, Slovakia) the extreme ranks are the same irrespective of the type of marginal value function used. In general, the average width of the ranges of possible ranks for each alternative is 2.9 for linear functions, 3.55 if we admit two linear pieces for each marginal function, and 4.1 for general functions. For alternatives whose evaluation vectors are typical within some significant subset of alternatives, this range is rather narrow (e.g., UK, Spain, Sweden, Czech Rep., Hungary). For alternatives good on some criteria while being relatively bad on the others the rank may differ significantly (e.g., Italy, Finland, Slovenia). Nevertheless, the observed ranges confirm that when increasing the number of characteristic points, the underlying preference model becomes more flexible, and additional additive value functions compatible with the preference information can be found.

The rank acceptability indices obtained for the general value functions are presented in Fig. 4. Although Germany may be ranked between first and fourth, for the majority of compatible value functions it is ranked at the top and the probability that it is placed outside top 2 is very low. In the same spirit, for many alternatives we could indicate a single rank attained by them for a significant share of compatible value functions (e.g., $RAI'(\text{Spain}, 3) = 96.35\%$, $RAI'(\text{Portugal}, 13) = 68.56\%$, $RAI'(\text{Slovakia}, 20) = 82.39\%$). For some other countries, analysis of the outcomes of SMAA enables to narrow down the range of the most probable ranks. For example, for over 95% of compatible value functions, Italy is ranked in positions between 4 and 6, whereas, in general, it could be ranked in positions between 2 and 10. Note that for a few countries some ranks have acceptabilities of 0.0 (e.g., (Germany, 4), (Italy, 2), (Austria, 7)), although ERA reveals they do attain these positions with at least one compatible value function.

For different types of marginal value functions, one could also note the differences in the provided recommendations with

Table B.4
The rank acceptability indices (in %) corresponding to Fig. 4.

RANK	1	2	3	4	5	6	7	8	9	10
GER	78.11	21.88	0.01							
UK	21.89	78.11								
SPA			96.35	3.44	0.21					
SWE				17.31	31.58	51.11				
NET		0.01	2.78	43.18	48.71	5.32				
ITA			0.86	36.07	19.50	40.85	1.12	0.67	0.76	0.17
	5	6	7	8	9	10	11	12	13	14
FIN		0.60	6.98	20.50	44.90	16.68	8.09	2.25		
BEL		1.01	51.20	41.00	6.16	0.62				
AUS					0.27	16.37	36.71	15.50	31.18	
DEN		1.11	6.14	11.20	26.00	32.95	11.58	11.00		
FRA			34.60	26.60	7.25	14.91	16.64			
CZE					14.60	9.31	22.69	53.40		
POR						8.99	4.29	17.90	68.56	0.26
	11	12	13	14	15	16	17	18	19	20
SLO			0.26	8.50	31.56	33.03	9.15	15.99	1.51	
IRE				53.74	28.61	11.76	4.65	1.24		
HUN				37.50	30.10	26.32	6.08			
EST						6.08	43.30	15.08	30.12	5.38
GRE						5.73	15.90	43.23	22.87	12.23
POL					9.73	12.38	18.50	19.59	39.85	
SVK						4.70	2.39	4.87	5.65	82.39

Table B.5
Rank acceptability indices (in %) for six chosen alternatives and different types of marginal value functions corresponding to Fig. 5.

	1	2	3	4	5	6	7	8	9	10
GER 1	100.0									
GER 2	99.16	0.84								
GER G	78.11	21.88	0.01							
UK 1		100.0								
UK 2	0.84	99.16								
UK G	21.89	78.11								
SPA 1			100.0							
SPA 2			100.0							
SPA G			96.35	3.44	0.21					
SWE 1				20.32	38.09	41.59				
SWE 2				28.83	23.50	47.67				
SWE G				17.31	31.58	51.11				
NET 1				45.86	52.89	1.25				
NET 2				31.96	66.09	1.95				
NET G		0.01	2.78	43.18	48.71	5.32				
ITA 1				33.82	9.02	57.16				
ITA 2				39.21	10.41	49.49	0.80	0.09		
ITA G			0.86	36.07	19.50	40.85	1.12	0.67	0.76	0.17

respect to the rank acceptability indices (see Fig. 5 for RAIs of the exemplary six countries). Although it is not possible to formulate general observations related to the comparison of shares of each rank for particular alternatives, we can see a greater diversity of the attained ranks with the growth of flexibility of the underlying preference model, and draw conclusions for each alternative individually. As for UK, its *RAI*′ with respect to the first positions grows from 0.0 (for linear functions) through 0.84–21.89% (for general functions). For the Netherlands, the probability that it is ranked fourth or fifth amounts to over 98% when considering linear or two-piece linear functions, whereas for general functions the *RAI*s corresponding to the third or sixth position become non-negligible.

5. Conclusions

We presented a new approach for multiple criteria ranking problems. The approach considers a set of additive value functions compatible with the holistic pair-wise preference statements provided by the DM. Then, the necessary and the possible preference relations are computed through Robust Ordinal Regression LPs, and these crisp relations are enriched with estimations of their probabilities. Furthermore, we considered the best and the worst ranks for each alternative and examined the distributions of ranks between the extreme positions. In this way, purely ordinal results can be confronted with the pair-wise outranking indices and rank acceptability indices of SMAA. We emphasized how the outcomes of the ordinal and stochastic analyses complement each other and how they can support incremental specification of preference information. We also presented an extension of the method for selecting a representative value function.

Although not discussed in detail, the introduced approach remains valid when considering preference information in form of rank-related requirements [11]. In this case the DM may provide a range of ranks a particular reference alternatives should attain, and the SMAA simulation process can additionally be used to support specification of such ranges.

Acknowledgments

The authors thank the three anonymous reviewers whose comments helped to improve previous version of the paper. In particular, the need of including Theorem 2.1 was suggested by one of the reviewers. The first author acknowledges financial support from the National Science Center.

Appendix A. Pseudo-code for rejection-sampling a set of general value functions

Algorithm 1.

Input: n_{vf}, the number of general value functions to sample
Input: P, the set of weak preference statements
Input: A, the set of alternatives
Output: S, the set of sampled value functions

```
S ← ∅
for l ∈ {1,...,n_vf} do
  ok ← FALSE
  while ok = FALSE do
    for j ∈ {1,...,m} do
      U_j(x_j^1) ← 0
      for k ∈ {2,...,n_j(A)} do
        U_j(x_j^k) ← Unif(0,1)
      end for
      SORT(U_j)
    end for
    ok ← TRUE
    for (a ≿ b) ∈ Γ do
      if U(a) < U(b) then
        ok ← FALSE
        break
      end if
    end for
  end while
  S ← S ∪ U
end for
```

Symbol explanations:

$U_j(x_j^k)$ the value of criterion evaluation x_j^k with the marginal value function of criterion j

$SORT(U_j)$ sort the function values of U_j in an ascending order (i.e. permute the values so, that $\forall i,k \in \{1,\ldots,n_j(A)\}$, $i < k : U_j\left(x_j^i\right) \leqslant U_j\left(x_j^k\right)$)

Γ a set of pair-wise weak preference statements provided by the DM (in case the statements contain indifference and/or strong preference ones, modify the if-check of lines 14–17 accordingly).

Appendix B. Rank acceptability indices for the illustrative example from Section 4

Tables B.4 and B.5.

References

[1] W. Aertsen, V. Kint, J. van Orshoven, B. Muys, Evaluation of modelling techniques for forest site productivity prediction in contrasting ecoregions using stochastic multicriteria acceptability analysis (SMAA), Environmental Modelling & Software 26 (7) (2011) 929–937.
[2] CSIC, 2010. Cybermetrics Lab at Consejo Superior de Investigaciones Cientficas (CSIC), Webometrics Ranking of World Universities, Madrid, Spain.
[3] S. French, Uncertainty and imprecision: modelling and analysis, Journal of the Operational Research Society 46 (1) (1995) 70–79.
[4] S. Greco, M. Kadziński, V. Mousseau, R. Słowiński, ELECTRE^{GKMS}: robust ordinal regression for outranking methods, European Journal of Operational Research 214 (1) (2011) 118–135.
[5] S. Greco, M. Kadziński, R. Słowiński, Selection of a representative value function in robust multiple criteria sorting, Computers & Operations Research 38 (11) (2011) 1620–1637.
[6] S. Greco, V. Mousseau, R. Słowiński, Ordinal regression revisited: multiple criteria ranking using a set of additive value functions, European Journal of Operational Research 191 (2) (2008) 415–435.
[7] S. Greco, V. Mousseau, R. Słowiński, Multiple criteria sorting with a set of additive value functions, European Journal of Operational Research 207 (4) (2010) 1455–1470.
[8] E. Jacquet-Lagrèze, Y. Siskos, Assessing a set of additive utility functions for multicriteria decision making: the UTA method, European Journal of Operational Research 10 (1982) 151–164.
[9] M. Kadziński, S. Greco, R. Słowiński, Extreme ranking analysis in robust ordinal regression, Omega 40 (4) (2012) 488–501.
[10] M. Kadziński, S. Greco, R. Słowiński, Selection of a representative value function in robust multiple criteria ranking and choice, European Journal of Operational Research 217 (3) (2012) 541–553.
[11] M. Kadziński, S. Greco, R. Słowiński, RUTA: a framework for assessing and selecting additive value functions on the basis of rank related requirements, Omega 41 (4) (2013) 735–751.
[12] R. Kannan, L. Lovász, M. Simonovits, Random walks and an $O*(n(5))$ volume algorithm for convex bodies, Random Structures & Algorithms 11 (1) (1997) 1–50.
[13] R. Keeney, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Trade-offs, Cambridge University Press, Cambridge, 1976.
[14] R. Lahdelma, J. Hokkanen, P. Salminen, SMAA-stochastic multiobjective acceptability analysis, European Journal of Operational Research 106 (1) (1998) 137–143.
[15] R. Lahdelma, K. Miettinen, P. Salminen, Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA), European Journal of Operational Research 147 (1) (2003) 117–127.
[16] R. Lahdelma, P. Salminen, SMAA-2: stochastic multicriteria acceptability analysis for group decision making, Operations Research 49 (3) (2001) 444–454.
[17] J. Lawrence, Polytope volume computation, Mathematics of Computation 57 (1991) 259–271.
[18] P. Leskinen, J. Viitanen, A. Kangas, J. Kangas, Alternatives to incorporate uncertainty and risk attitude in multicriteria evaluation of forest plans, Forest Science 52 (3) (2006) 304–314.
[19] A. Menou, A. Benallou, R. Lahdelma, P. Salminen, Decision support for centralizing cargo at a moroccan airport hub using stochastic multicriteria acceptability analysis, European Journal of Operational Research 204 (3) (2010) 621–629.
[20] Mousseau, V., 2005. A General Framework for Constructive Learning Preference Elicitation in Multiple Criteria Decision Aid. Cahier du Lamsade 223, Universite Paris Dauphine.
[21] M. Roubens, P. Vincke, On families of semiorders and interval orders imbedded in a valued structure of preference: a survey, Information Sciences 34 (1984) 187–198.
[22] B. Roy, Two conceptions of decision aiding, International Journal of Multicriteria Decision Making (IJMCDM) 1 (1) (2010) 74–79.
[23] T. Tervonen, J. Figueira, A survey on stochastic multicriteria acceptability analysis methods, Journal of Multi-Criteria Decision Analysis 15 (1–2) (2008) 1–14.
[24] T. Tervonen, H. Hakonen, R. Lahdelma, Elevator planning with stochastic multicriteria acceptability analysis, Omega 36 (3) (2008) 352–362.
[25] T. Tervonen, R. Lahdelma, Implementing stochastic multicriteria acceptability analysis, European Journal of Operational Research 178 (2) (2007) 500–513.
[26] T. Tervonen, G. van Valkenhoef, N. Baştürk, D. Postmus, Hit-and-run enables efficient weight generation for simulation-based multiple criteria decision analysis, European Journal of Operational Research 224 (3) (2013) 552–559.
[27] T. Tervonen, G. van Valkenhoef, E. Buskens, H.L. Hillege, D. Postmus, A stochastic multicriteria model for evidence-based decision making in drug benefit-risk analysis, Statistics in Medicine 30 (12) (2011) 1419–1428.
[28] G. van Valkenhoef, T. Tervonen, J. Zhao, B. de Brock, H.L. Hillege, D. Postmus, Multi-criteria benefit-risk assessment using network meta-analysis, Journal of Clinical Epidemiology 65 (4) (2012) 394–403.
[29] P. Vincke, Robust solutions and methods in decision-aid, Journal of Multi-Criteria Decision Analysis 8 (3) (1999) 181–187.