Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 174 (2006) 278–292

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Decision Support

Using intervals for global sensitivity and worst-case analyses in multiattribute value trees

Jyri Mustajoki, Raimo P. Hämäläinen *, Mats R.K. Lindstedt

Systems Analysis Laboratory, Helsinki University of Technology, P.O. Box 1100, FIN-02015 HUT, Finland

Received 10 January 2003; accepted 17 February 2005
Available online 13 June 2005

Abstract

Sensitivity analyses have long been used to assess the impacts of uncertainties on outcomes of decision models. Several approaches have been suggested, but it has been problematic to get a quick overview of the total impact of all the uncertainties. Here we show how interval modeling can be used for global sensitivity analyses in multiattribute value trees, and a nuclear emergency case is used to illustrate the method. The approach is conceptually simple and computationally feasible. With intervals, the decision maker can include all the possible uncertainties and quickly estimate their combined impact. This is especially useful in high-risk decisions where a worst-case type of sensitivity analysis is essential. By varying the intervals one can also examine which uncertainties have the greatest impact and thus need the most consideration. Global sensitivity analysis reveals how the outcome is affected by many simultaneous variations in the model.

Keywords: Multiple criteria analysis; Sensitivity analysis; Preference programming; Interval modeling; Worst-case analysis; Nuclear emergency management

1. Introduction

Dealing with uncertainties related to data and preferential judgments is an essential part of a practical decision analysis project. How this is done is likely to strongly affect the confidence that the decision makers have in the results. If the uncertainties are not well accounted for, the credibility of the decision analysis method used can suffer. Thus an easy-to-use and transparent method for examining the effects of the uncertainties is needed to ensure the decision makers' commitment to the decision.

Sensitivity analyses are commonly used to analytically assess the impacts of uncertainties on outcomes of decision models. In this paper, our focus

* Corresponding author. Tel.: +358 9 451 3054; fax: +358 9 451 3096.
E-mail addresses: jyri.mustajoki@hut.fi (J. Mustajoki), raimo@hut.fi (R.P. Hämäläinen), mats.lindstedt@sitcom.fi (M.R.K. Lindstedt).

0377-2217/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2005.02.070

is on multiattribute value tree analysis (MAVT) as described in Keeney and Raiffa (1976). MAVT is based on structuring the decision problem into a value hierarchy, or a value tree. The topmost objective is the overall goal that the decision maker wishes to achieve. This objective is divided into sub-objectives, and on the lowest level are the measurable attributes that are important for the decision problem at hand. The attributes are weighted according to their importance, and the overall performance of each alternative is measured by a value function, which aggregates the performance measures of this alternative on each attribute into a single overall value measure.

Previous research has suggested that intervals and incomplete judgments can be used to incorporate uncertainties directly in the modeling phase (see e.g. Arbel, 1989; Salo and Hämäläinen, 1992, 1995, 2001; Weber, 1985; White et al., 1984). In an interval approach, preference judgments and the outcomes of the alternatives are presented as ranges including all the possible value estimates for these. Consequently, the overall values of the alternatives will be intervals. The approach is also called Preference Programming (Arbel, 1989; Salo and Hämäläinen, 1995, 2003) which reflects the fact that one can see how the preferences over the alternatives evolve as the information increases and the preference statements become less incomplete. Interval methods have been successfully used in applications including environmental decision making and policy analysis (see e.g. Hämäläinen and Leikola, 1996). Intervals can also be used in group decision making to incorporate the preferences of all participants in a single model (see e.g. Hämäläinen et al., 1992; Hämäläinen and Pöyhönen, 1996; Kim et al., 1999).

In this paper, we present a way to carry out global sensitivity analysis in MAVT by Preference Programming. That is, we allow the model parameters to vary within given intervals representing the uncertainty ranges and study the consequential effects in the results. The aim is to quickly assess the total impact of all uncertainties. The main benefit of our approach is that it is conceptually straightforward. It captures all the uncertainties in a single analysis, and intervals used in the model are easy to understand for non-mathematicians who might not be familiar or comfortable with probability distributions associated with the preferences and the ratings of the alternatives. In this type of analysis, it is the extreme possibilities (i.e. the worst-case results) that are important and not the probability distributions. Yet, the proposed approach explicitly studies the sensitivity with respect to the input elicited from the decision maker (i.e. the weight ratios between the attributes). The previous approaches including the traditional one-way sensitivity analysis have typically only considered the sensitivity of the final weights.

This work fits into the framework of sensitivity analysis proposed by Rios Insua and French (1991) and Proll et al. (2001), in which one employs constraints on the model parameters to describe uncertain or imprecise information. The Preference Programming approach applied here gives a convenient and computationally efficient way to include the constraints in the weight ratios of the attributes and in the outcomes of the alternatives, also in hierarchical value trees. However, some related concepts such as potential optimality (see Hazen, 1985; Rios Insua and French, 1991) may not be applicable when considering the analysis from the worst-case perspective, as it may eliminate alternatives that are never optimal but perform reasonably well in all situations and are thus less risky.

The interval model leads to a set of linear extremum problems. In hierarchical MAVT models, the overall problem is computationally quick to solve, as the extremum problems on each branch of the value tree can be separately solved (Salo and Hämäläinen, 1992, 1995), and in a typical model there are seldom more than ten sub-objectives under any objective in the hierarchy. This makes it possible to carry out interactive analyses. For example, with our WINPRE software (Workbench for INteractive PREference Programming; Hämäläinen and Helenius, 1997), the decision maker can immediately see the changes in the results when adjusting the intervals. This requires that the calculations can be done very fast, preferably within a few tenths of a second, which is not always possible with other approaches.

Our focus is on MAVT, and thus other models such as decision trees (see Clemen, 1996) are beyond the scope of this work. The suggested approach may also be well suited for other types of models, although the number of parameters may set limitations for its use (e.g. the model presented in Francos et al. (2003) involves 82 input parameters). For an overview of suitable global sensitivity analysis approaches for these cases (such as variance based importance measures, Bayesian networks, etc.), the reader is referred to the special issue of Reliability Engineering and System Safety on sensitivity analysis (Tarantola and Saltelli, 2003).

Global worst-case analysis is especially relevant in high-consequence decisions with a high level of uncertainty, such as in the case of nuclear emergencies. We have previously successfully applied multiattribute decision analysis and decision conferences to support nuclear emergency management (Hämäläinen et al., 2000), and the example case used in this paper is taken from one of these studies (Ammann et al., 2001).

This paper is organized as follows. Section 2 describes different approaches to carry out sensitivity analysis. Section 3 describes the interval sensitivity analysis method, and how to use it in practice. An example of using interval sensitivity analysis in a nuclear accident case is given in Section 4, and Section 5 concludes the paper.

2. Sensitivity analyses—Why and how?

Uncertainties can be grouped in several different ways. French (1995) suggested a classification into three groups depending on which step of the analysis they belong; modeling, interpreting the results or exploring the model (Table 1). One should note that in this paper the term uncertainty is used in a similar way as in French (1995) to also cover such types of uncertainties that may arise from imprecision or ambiguity.

Another perspective is to focus on the origins of uncertainties. For example, in the field of risk analysis Salo (2001) suggests three dimensions of technological risks; physical causation (e.g. uncertainties in causal relationships), value concerns (e.g. changes in the stakeholders' preferences) and policy response (e.g. the effectiveness of the actions taken).

In the different cases, methods are needed to support the decision maker in dealing with the uncertainties in a constructive way. Sensitivity analyses are commonly used, and the reader is referred to Saltelli et al. (2000a,b), Tarantola and Saltelli (2003) or French (2003) for a perspective on different approaches in general, or to Belton and Stewart (2001), French and Rios Insua (1999) or Rios Insua and French (1991) on different approaches specific to MAVT.

Sensitivity analyses can be used for a wide range of purposes. Pannell (1997) grouped these into four categories:

1. Decision making (identifying critical values/parameters, testing robustness, overall riskiness of decision).
2. Communication (increasing commitment/confidence/credibility, explicitly showing critical assumptions).
3. Increased understanding (understanding relationship between input/output variables).
4. Model development (identifying needs for more accurate measurements/more information).

The single parameter test is a common sensitivity analysis method to examine how sensitive a

Table 1
Classifying uncertainties following French (1995)

When modeling the decision problem	When interpreting the results	When exploring the model
• Uncertainty about what might happen or what can be done • Uncertainty about meaning or ambiguity in terminology • Uncertainty about related decisions	• Uncertainty about the appropriateness of a descriptive/normative model • Uncertainty about the depth to which to conduct the analysis	• Uncertainty resulting from physical randomness or lack of knowledge • Uncertainty about the evolution of future beliefs and preferences • Uncertainty about judgments • Uncertainty about accuracy of calculations

model is to small changes in one parameter. That is, all other parameters are held fixed except a single one that is allowed to vary. The analysis is usually visualized by graphs showing the consequential variations in the overall results. The tornado diagram is another common method providing useful diagrams. It is usually drawn for single parameter tests and employed to compare the base alternative to another option (Felli and Hazen, 1999). For a detailed discussion of tornado diagrams, the reader is referred to Clemen (1996). These methods do not, however, account for parameter interactions nor do they cover the worst-case settings.

Often there is more than one parameter that the decision maker might be uncertain about, and with the traditional approaches it can be difficult to estimate the combined overall impact. For example, Felli and Hazen (1999) showed that single parameter tests tend to overestimate the overall sensitivity. This leads to the need for a multiparameter test.

Global sensitivity analysis incorporates the influences of the whole ranges of variation in model parameters and these variations are allowed in multiple parameters simultaneously (Saltelli et al., 2000a). Especially in complex simulation models, for example, when modeling environmental phenomena, the model can have numerous input parameters. Different types of global sensitivity analysis approaches have been developed for dealing with uncertainties in these models, and these approaches often rely on statistical or probabilistic calculations. Examples of such approaches include Monte Carlo analysis, ANOVA, FAST and Bayesian models (see e.g. Saltelli et al., 2000a,b). Rank based methods can also be used to study the sensitivity of the model with respect to uncertainties in the rankings of parameters (see e.g. Barron, 1992; Salo and Punkka, 2005).

In MAVT, the Monte Carlo simulation technique can be used to analyze model uncertainties and to statistically rank the alternatives (see e.g. Arbel and Vargas, 1993; Butler et al., 1997; Stam and Silva, 1997). The main advantage of this method is that one gets a lot of information, such as mean values, variances and fractiles, about the characteristics of the decision model subject to uncertainties. However, the normalization of the weights and the reciprocity of the weight ratios make the application of distributions on the weights very difficult. In addition, the effort needed for the calculations can become substantial and one cannot always carry out what-if analyses without computational delays. To avoid this, Kirkwood (1992) has suggested a method to estimate the impact of uncertainty on the results of a multiattribute model prior to a complete probabilistic analysis. However, this would still demand the approximation of expected values, variances and covariances.

Kirkwood (1997) provides a summary of literature on scenario analysis in decision making. In scenario analysis, parameters are given (sometimes extreme) values so as to make the analysis favor a certain view of the actual situation. Often the set of scenarios contains expected, optimistic and pessimistic alternatives. The process of scenario generation is not, however, a straightforward one. For example, there is no clear rule how pessimistic the pessimistic scenario should be.

The precautionary principle (see e.g. Goldstein and Carruth, 2004; Graham, 2000) has recently received growing attention. The precautionary approach addresses the problems of multidimensionality, humility about knowledge and openness to alternatives. It places the burden of proof on the advocates to prove the soundness of the suggested decision. Thus, also in that approach, a global worst-case analysis would be useful.

3. Worst-case sensitivity analysis with intervals

3.1. Interval methods

In MAVT the overall values of the alternatives are composed of the ratings of the alternatives with respect to each attribute and of the weights of the attributes. If the attributes are mutually preferentially independent (see e.g. Keeney and Raiffa, 1976), an additive value function can be used to derive the overall values for the alternatives to represent the decision maker's overall preference over the alternatives. The overall value of alternative x is

$$v(x) = \sum_{i=1}^{n} w_i v_i(x_i), \tag{1}$$

where n is the number of attributes, w_i is the weight of attribute i, x_i is the consequence of alternative x with respect to attribute i and $v_i(x_i)$ is its rating. The weights denote the relative importances of the attributes changing from their worst level to their best level compared to the changes in the other attributes. The sum of the weights is normalized to one, and the ratings $v_i(x_i)$ are scaled onto the range $[0, 1]$.

The value tree can also be constructed hierarchically. Then the upper level objectives are divided into sub-objectives, and the weighting is carried out locally on each set of these. In this case, w_i in (1) denotes the overall weight of the lowest level attribute i, which is calculated as a product of the local weight of this attribute and the local weights of all the preceding upper level objectives.

Preference Programming can be used to model uncertainties in the decision maker's preference statements with intervals. PAIRS (Salo and Hämäläinen, 1992) is a Preference Programming method in which intervals are directly given to constrain both the weight ratios of any attribute pairs and the ratings of the alternatives. For example, instead of giving an exact weight ratio $w_1/w_2 = 2$, the decision maker can define that $w_1/w_2 \in [1, 4]$ (i.e. the ratio is at least 1 but no more than 4). The given intervals constrain the feasible region of the weights. Fig. 1 illustrates the feasible region S constrained by intervals $w_1/w_2 \in [1, 4]$, $w_1/w_3 \in [1, 4]$ and $w_2/w_3 \in [1/2, 2]$. Similarly, uncertainties in decision outcomes can be modeled with ranges of possible values (e.g. $v_i(x_i) \in [0.3, 0.5]$). As a result, the overall values of the alternatives will also be intervals, which can be calculated as extremes of (1) with linear programming. The lower bound for the value of alternative x is

$$\underline{v}(x) = \min_{w \in S} \sum_{i=1}^{n} w_i \underline{v}_i(x_i), \tag{2}$$

where S is the feasible region of the weights, $\underline{v}_i(x_i)$ is the least allowed value for $v_i(x_i)$ and $w = (w_1, \ldots, w_n) \in S$. The upper bound is calculated analogously. For details refer to Salo and Hämäläinen (1992).

Interval models require that we specify the dominance concepts to analyze the results (see e.g. Weber, 1987; Salo and Hämäläinen, 1992). Alternative x dominates alternative y absolutely, if the lower bound of the overall value interval of x is higher than the upper bound of the interval of y. Alternative x dominates y in a pairwise sense, if the overall value of x is higher than or equal to that of y for every feasible weight and value combination, i.e. if

$$\min_{w \in S} \sum_{i=1}^{n} w_i [\underline{v}_i(x_i) - \overline{v}_i(y_i)] \geqslant 0, \tag{3}$$

and the inequality is strict at least for some $w \in S$. Thus, pairwise dominance can also exist on overlapping value intervals. Absolute dominance im-

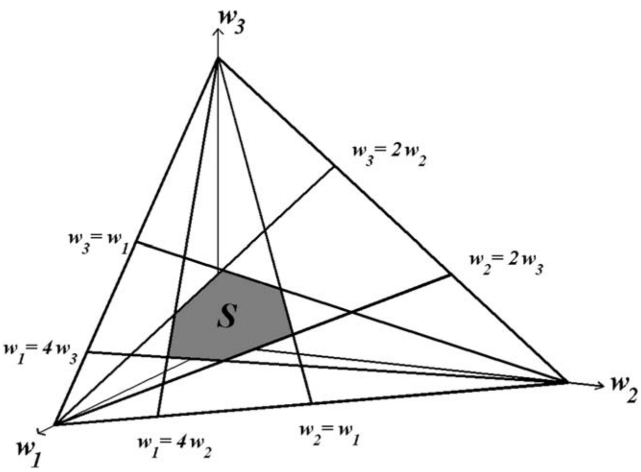

Fig. 1. Example of the feasible region of the weights S on the weight plane $w_1 + w_2 + w_3 = 1$.

plies pairwise dominance, and in general, the term dominance is considered to refer to pairwise dominance.

3.2. Interval sensitivity analysis

In this paper, we apply the PAIRS method to carry out interval sensitivity analysis in MAVT. In this approach, the decision maker extends the initial point estimates of the weight ratios and attribute ratings to intervals indicating the ranges within which these estimates are allowed to vary. These intervals describe the possible uncertainties in the weight ratios and attribute ratings. By studying the consequential changes in the overall value intervals and in dominance relations, the decision maker can elicit how sensitive these are to all the possible variation in the model simultaneously. For example, if the dominance relations between the alternatives remain unchanged when assigning the uncertainty intervals, the rank order of the alternatives is not sensitive to any allowed variation in the model. However, if more than one alternative become non-dominated, any of these can be considered as a suitable candidate for the most preferred alternative.

Intervals can be simultaneously assigned both to the weight ratios between the attributes on any level of the value tree, and to the ratings of the alternatives. This makes it possible to study the joint effects of different types of uncertainties, for example, due to the decision maker's inability to accurately estimate the relative importances of the attributes and due to the imprecision in the input data on the alternatives.

Interval sensitivity analysis can also be applied to the Analytic Hierarchy Process (AHP; Saaty, 1980). In the AHP, the attribute weights are computed from a matrix of all the pairwise weight ratio judgments. Thus, Preference Programming can be used to carry out interval sensitivity analyses by setting intervals on all the pairwise judgments. However, in AHP one also has to take into account feasibility constraints due to the redundancy of comparisons.

3.2.1. Interpreting the intervals

In practice there are different ways to interpret the intervals. The strictest one is to set the intervals so that they cover all the possible variation in the weight ratios and in the ratings. Then the sensitivity analysis can be seen as a way to find a true worst-case solution. That is, if a dominating alternative is found, it is the best alternative for every feasible combination of model parameters, including the worst ones. However, with this interpretation, the intervals may easily become so wide that no dominance relations between the alternatives can be established.

One can also use tighter intervals that may not cover all the possible variation in parameter values. The analysis should then be interpreted as a "what-if" type sensitivity analysis. That is, we study *what* would the overall value intervals be, *if* we allowed the weight ratios and ratings to be any values within the given intervals.

Yet another approach is to consider intervals as confidence intervals (e.g. 95% confidence interval). However, if the parameter distributions on these intervals are not given, the overall value intervals would not be true confidence intervals, but they would merely describe the bounds for the possible variation in the overall values when allowing each parameter to be any value within its confidence interval. To get true confidence intervals for the overall values, one has to assign distributions to the parameters, but this leads to the use of a Monte Carlo simulation approach.

Attributes can have different levels of uncertainty, which can be taken into account by using intervals of different size. If the decision maker is not willing or able to give the intervals explicitly, he/she can, for example, assign error ratios to the initial point estimates (see e.g. Bryson et al., 1995; Salo and Hämäläinen, 2001). They provide a quick way to model the proportional uncertainties. An error ratio is a coefficient by which we multiply each weight ratio to get an upper bound for this ratio. For example, with error ratio 2, the weight ratio $w_1/w_2 = 3$ extends to interval $w_1/w_2 \in [3/2, 3 \times 2] = [1.5, 6]$. Similarly, the ratings of the alternatives can be extended to intervals by setting a general value for the maximum allowed variation in the initial ratings. For example, with the maximum allowed variation of ± 0.1, the rating estimate $v_i(x_i) = 0.3$ extends to interval $v_i(x_i) \in [0.3 - 0.1, 0.3 + 0.1] = [0.2, 0.4]$.

The choice of intervals could, however, be difficult if there is a high level of uncertainty involved. This could lead to a set of wide intervals, when it may be too risky to draw any conclusions from the results. In this case, the value of the analysis would be in finding out the key factors that resulted in the non-acceptable uncertainty intervals. That is, we can assign the uncertainty intervals to different model parameters or parameter groups one-by-one, and study what uncertainties have the greatest impact and thus need the most consideration. A detailed analysis might even reveal how much the uncertainties need to be reduced to arrive at an acceptable outcome.

3.2.2. Potential optimality and the decision rules

In many cases it may be useful to first try to reduce the number of alternatives before the final decision on which alternative to choose. The dominance test eliminates clearly inferior alternatives but the remaining set can still be extensive. Rios Insua and French (1991) discuss the use of potential optimality (Hazen, 1985) for reducing the number of alternatives. This method aims at the optimal solution and only considers alternatives that could potentially be optimal. In many cases this can be a useful approach but not necessarily in the worst-case analysis, as it might reject non-optimal alternatives that would have been acceptable to the decision maker. For example, in the nuclear accident case presented here, a non-optimal alternative might be the best decision if it performs reasonably well in all the possible scenarios.

As a result of a worst-case analysis there may be several non-dominated alternatives. The decision maker can then carry out what-if analyses to study with which uncertainty intervals there would only be one non-dominated alternative. If a set of such intervals are found, the decision maker should consider whether these intervals could be accepted to represent the related uncertainties. If he/she is not willing or able to modify the intervals, decision rules (see e.g. Salo and Hämäläinen, 2001) can be applied to rank non-dominated alternatives, for example, according to the minimum values of the overall value intervals.

The possible loss-of-value estimates (Salo and Hämäläinen, 2001) can also be used to further compare non-dominated alternatives. The possible loss-of-value for alternative x is calculated as

$$\max_{w \in S, y \neq x} \sum_{i=1}^{n} w_i [\overline{v}_i(y) - \underline{v}_i(x)]. \tag{4}$$

It indicates how much the decision maker can at the most lose in the overall value by choosing alternative x instead of any other alternative. Thus, it can also be seen to describe the decision maker's maximum regret if he/she chooses this alternative. Possible loss-of-value can be used, for example, to reduce the set of non-dominated alternatives by specifying an acceptable limit for how far from optimality any alternative can be at most. If the possible loss-of-value for the originally chosen alternative is within acceptable limits, this alternative might still be acceptable, even if the uncertainties might cause this to be overtaken by another alternative in some circumstances.

3.2.3. Origins of uncertainty

Depending on the origins of uncertainty (see e.g. French, 1995) one can assign intervals to the ratings of the alternatives or to the weight ratios between the attributes. An interval assigned to a rating of an alternative describes uncertainty in this alternative only, and any variation within the given interval is assumed to be independent of the allowed variations on the other alternatives. On the other hand, by definition the weight of an attribute describes the importance of the range of this attribute compared to the other attribute ranges (Keeney and Raiffa, 1976). In addition, any variation in the weight of an attribute affects linearly to the attribute-specific value (i.e. the share of the overall value contributed through this attribute) of each alternative. Thus, linear correlations in the uncertainties of the alternatives' measurement values on some attribute can be modeled by keeping the corresponding ratings constant and assigning an uncertainty interval to the weight of this attribute. Then, the relative ratings of the alternatives on this attribute remain the same, but the relative importance of this attribute is allowed to vary with respect to the other attributes.

As a simple example, consider attribute *Costs* on two *alternatives A* and *B*. If there is uncertainty in, for example, some general costs, it is likely to affect both alternatives in the same way so that the relative *Costs* between *A* and *B* remain the same (i.e. if the *Costs* of *A* double, so do the *Costs* of *B*). This kind of uncertainty is actually related to the range of the measurement values of the *Costs*, and thus should be assigned to the weight of the *Costs*. In contrast, if uncertainty in the *Costs* of *A* is independent of any uncertainty in the *Costs* of *B*, these should be modeled by assigning uncertainty intervals directly to the ratings of the *Costs*.

3.2.4. Computer support

In practice, computer support is needed as the solutions are elicited through linear programming. In hierarchical MAVT models, the overall linear extremum problem can be decomposed into smaller problems so that there is one extremum problem for each branch of the hierarchy (Salo and Hämäläinen, 1992, 1995). Thus, when varying a model parameter, the linear programming problems need to be re-calculated only on those branches of the value tree in which the change is made, and upwards thereof. This makes the model computationally quick to solve and update, which is especially important in *what-if* type of analyses. Computer support can also be used, for example, to visualize the results.

As noted above, there already exists software for the interval techniques considered here, and sensitivity analyses can easily be run on this software. The WINPRE software provides an interactive approach by presenting to the user the overall values of the alternatives and the dominance relations immediately when making changes to the model parameters. The example case in Section 4 is analyzed using WINPRE and Figs. 2–6 are screen captures from the software. WINPRE is freely available for academic purposes on the Decisionarium Web site (www.decisionarium.hut.fi; Hämäläinen, 2000, 2003).

3.3. Comparison with other sensitivity analysis approaches

Next we shall compare interval sensitivity analysis to one-way sensitivity analysis and related techniques such as tornado diagrams, as well as to the Monte Carlo simulation technique, which are commonly used sensitivity analysis approaches in MAVT. We shall also discuss for which types of situations the different approaches are best suited.

3.3.1. One-way sensitivity analysis

The resulting graph in one-way sensitivity analysis shows the overall values of the alternatives with respect to each possible weight of an attribute under consideration (see e.g. Clemen, 1996). Two-way sensitivity analysis extends this analysis into two attributes, when the combined effects of variations in these are presented with a three-dimensional graph. However, with more than two attributes, visualization of the analysis becomes impossible. With tornado diagrams one can study the effects of several different parameter variations in the same graph, but these effects are also calculated by varying a single weight at a time. Similar analyses can be carried out on the ratings of the alternatives.

In MAVT the attribute weights are typically estimated from a process in which the decision maker gives ratio statements about the relative importances of the attribute ranges. For example, in the SWING method (Von Winterfeldt and Edwards, 1986), the decision maker is first asked to give a hundred points to the most important attribute (i.e. to an attribute whose consequence he/she most preferably would change from its worst level to its best level). Then, he/she is asked to assign fewer points to each other attribute to denote the relative importance of the corresponding consequence change in this attribute compared to the change in the most important attribute. Finally, the actual attribute weights are elicited by normalizing the sum of these points to one. Thus, uncertainty in the personal judgments is related to the weight ratio statements between the attributes. However, one-way sensitivity analysis only considers changes in an individual weight at a time, and it does not directly reflect the uncertainty originating from the weight elicitation process. In contrast, interval sensitivity analysis explicitly considers the weight ratios between the attributes, and consequently uncertainty in the weight assessment. One should note that one-way sensitivity analysis could also be carried out, for example, by varying the points given to the attributes in the SWING method. As far as the authors know, this approach has not yet been applied.

Especially in hierarchical problems, interval sensitivity analysis as a multiparameter analysis may provide useful additional information compared to one-way sensitivity analysis. In these problems, the weighting is carried out locally on each branch of the value tree, and the overall weight of the lowest-level attribute is obtained as a product of the local weights of this and all the preceding upper level attributes. With one-way analysis we can only consider uncertainty in one branch of a value tree at a time. However, if one wants to take the joint effects of uncertainties in different attribute levels as well as in alternatives into account, multiparameter analysis is needed.

One-way sensitivity analysis shows the resulting overall values for all the possible values of the parameter under consideration in one graph. In contrast, interval sensitivity analysis only shows the extreme values of the overall values with respect to allowed variations in model parameters. Thus, if one wants to study how the changes in the given parameter intervals affect the overall value intervals, he/she has to produce a separate graph for each set of intervals. Another way is to use the method interactively by varying the parameter intervals and studying the immediate response in the overall value intervals.

3.3.2. Monte Carlo simulation

In Monte Carlo simulation, one assigns full probability distributions to the ratings of the alternatives and to the weights of the attributes. The overall values of the alternatives are calculated from each sample over these distributions and, as a result, one gets overall value distributions reflecting the variation in the model parameters. One can also estimate some other statistical measures from this data, such as the probability of one alternative being better than some other alternative.

It is conceptually straightforward to assign distributions to the ratings of the alternatives. However, the situation is different when we consider the weights of the attributes, which are dependent of each other through the normalization. Earlier researchers have found a specific distribution for the weights (i.e. the Dirichlet distribution) with which the simulation can be done (Butler et al., 1997; Moskowitz et al., 2000). However, the use of general distributions on the weights becomes intractable due to the normalization of the sum of the weights to one. In addition, the approach only considers the final weights, and thus the interpretation of the given distributions may become problematic, as was noted earlier. Another simulation technique is to assign distributions onto the feasible regions of the weights (Haines, 1998). However, in this case the marginal distributions of the individual weight ratios become ambiguous. To simulate the effects of imprecision in the statements of the decision maker, one should assign distributions to the given weight ratios, similarly to the way that intervals are applied in the interval SMART/SWING method (Mustajoki et al., 2005). However, due to the normalization of the weights and the reciprocity of the weight ratios this approach also leads to problems in those weight ratios that are not given explicitly. For example, a uniform distribution assigned to a weight ratio on interval $w_1/w_2 \in [1, 2]$ implies that the corresponding distribution on the reciprocal ratio interval $w_2/w_1 \in [1/2, 1]$ will not be uniform.

In many cases, the decision maker is interested in the extreme values, for example, to get a worst-case solution, or to study deterministic dominance relations between the alternatives. For these purposes, interval sensitivity analysis should be used, as it explicitly concentrates on the extremes of the intervals. Monte Carlo simulation can be used to find additional statistical information on the value distributions and relations between the alternatives. However, the assignment of the distributions and sampling under constraints could often become an overwhelming task.

4. Interval sensitivity analysis in a nuclear accident exercise

We demonstrate the use of interval sensitivity analysis by applying it to a model developed in a decision conference (i.e. a facilitated training workshop) exercise. The aim of the conference was to plan protective actions on the milk production chain in a case of a hypothetical nuclear accident (for details, see Ammann et al., 2001). The conference was one in a series of decision conferences on nuclear emergency management (see also Hämäläinen et al., 2000). Although no real decisions were made in the exercise, the gained experiences are to be utilized should a real accident take place. The participants of the conference and the preparatory meetings were representatives of the safety authorities and experts on radiation, farming and the dairy industry.

4.1. Multiattribute value tree

The MAVT approach was used to structure the problem. The value tree (Fig. 2) was developed on the basis of the discussion in the decision conference and in the preliminary meetings. It describes the objective hierarchy in which the *Overall* goal of finding the best action is composed of three main objectives: *Health* effects, *Socio-psychological* effects and *Costs*. These are further divided into sub-objectives (i.e. attributes) and the alternatives are measured with respect to these. For the full names of the attributes, see Table 2.

It was assumed that due to the accident, fodder in the area becomes contaminated, and if nothing is done, the radioactivity migrates into milk products. As a precautionary action, the cattle were sheltered and provided with uncontaminated fodder and water for the first week after the accident. Our focus was on the later phase actions (from 1 to 12 weeks after the accident). Three protective action policies were considered: (i) *supplying clean fodder* (‘Fod’), in which uncontaminated fodder is transported into the contaminated area, (ii) *production change* (‘Prod’), in which the milk production is replaced by other dairy products, as the production processes of these can enrich, dilute or secrete radio nuclides, and (iii) *banning milk* (‘Ban’), in which the use of contaminated milk is totally banned. In addition, an action where *nothing* is done (‘– – –’) was included in the analysis as a reference. The actions were divided into two phases. The first phase covers the actions during weeks 2–5 after the accident and the second phase covers weeks 6–12. The actual alternatives considered were combinations of these. For example, ‘Ban+Fod’ represents an alternative where the use of milk is banned during weeks 2–5 and clean fodder is supplied during weeks 6–12. In Fig. 2, the alternatives are shown as the rightmost elements of the value tree.

The initial point estimates for the attribute weights and the alternatives’ ratings in the model (Table 2) are the ones given by one of the participant groups in the conference. The estimates for

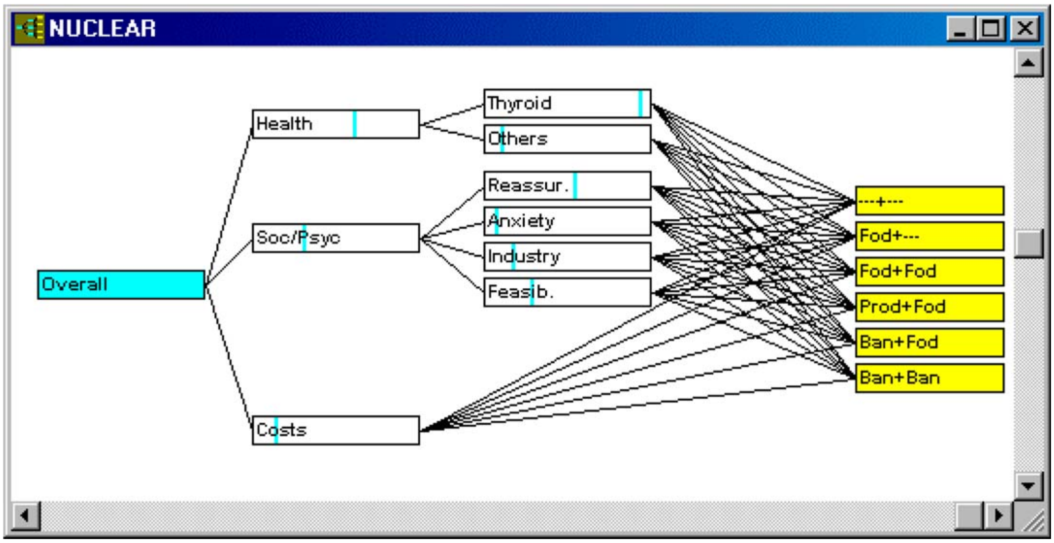

Fig. 2. Value tree for evaluating the protective actions in the case. A screen capture from the WINPRE software.

Table 2
Weights of the attributes and the ratings of the alternatives

	Weights		Ratings					
	Local weight	Overall weight	---+---	Fod+---	Fod+Fod	Prod+Fod	Ban+Fod	Ban+Ban
Health	*0.588*							
Thyroid cancers	0.909	0.534	0.01	0.98	1.00	0.98	1.00	1.00
Other cancers	0.091	0.053	0.04	0.90	0.98	0.95	0.98	1.00
Socio-psychological	*0.294*							
Reassurance	0.526	0.155	0.00	0.18	0.60	0.48	0.76	0.86
Anxiety	0.053	0.016	0.00	0.21	0.87	0.75	0.49	0.35
Industry	0.158	0.046	0.10	0.46	0.67	0.34	0.19	0.14
Feasibility	0.263	0.077	1.00	0.65	0.55	0.78	0.67	0.46
Costs	*0.118*	0.118	1.00	0.88	0.82	0.70	0.26	0.00

the consequences of *Health* effects and *Costs* were calculated with a Real-time on-line Decision Support system called RODOS (Ehrhardt and Weis, 2000). The values of *Socio-psychological* attributes were directly rated by the group. Attribute weighting was carried out with the SWING method.

The resulting overall values for the alternatives are shown in Fig. 3. The alternatives are listed at the bottom, and the upper and lower bounds for the overall value of each alternative are shown above its name. The bounds are also graphically shown on a [0–1] value scale between the numerical values. One should note that on each alternative the lower and upper bounds are now the same, as initially both the attribute weights and the alternatives' consequences are exact point estimates. The figure indicates that alternative 'Fod+Fod', where clean fodder is provided for both periods, is the best alternative with an overall value of 0.86.

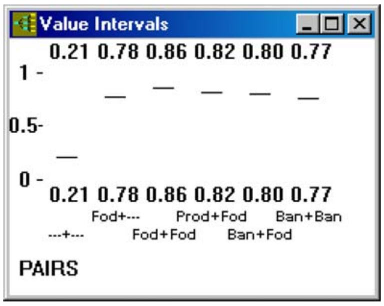

Fig. 3. Overall values for the alternatives.

4.2. Sensitivity analysis

This example is a typical case where interval sensitivity analysis is useful. In emergency planning, we often want to design precautionary actions following the worst-case approach. The alternative with the highest performance score might not be the right one to choose if a worst-case analysis reveals a risk of getting an unacceptably low performance under certain circumstances. An alternative with a slightly lower scoring on an average might be preferred, if the analysis shows that it performs in an acceptable way in all the possible circumstances. When applying the worst-case approach, we need to take into account all the possible uncertainties both in the ratings of the alternatives and in the weight ratios between the attributes in different levels of the value tree. Thus, multiparameter analysis is needed.

Next, we demonstrate the use of interval sensitivity analysis in this case. The objective is to study the changes in the relations between the alternatives when all the possible uncertainties in the problem are considered. The analysis consists of three phases. First, the sensitivity in the weight assessment is studied by extending the initial weight ratio point estimates to intervals. In the next phase, the effects of the possible variation in the alternatives' ratings are studied by giving these as intervals. Finally, uncertainties in both of these are simultaneously taken into account. The approximation of uncertainties is based on the results of the survey carried out among the decision makers in the conference.

To model uncertainty in the weight assessment, we extend each initial weight ratio estimate between the attributes to intervals. For each weight ratio, an error ratio 2 is used to reflect this uncertainty. For example, the weight ratio estimate between the attributes *Health* and *Socio-psychological effects* extends to an interval $[(0.588/0.294)/2, (0.588/0.294) * 2] = [1, 4]$. Fig. 4 shows the resulting value intervals and dominance relations (e.g. "$A \rightarrow B$" denotes that A dominates B). Alternative 'Fod+Fod' still dominates all the other alternatives. Thus, the choice of the most preferred alternative is not sensitive to considerable variation in the weights. One should note that alternatives 'Fod+- - -', 'Prod+Fod', 'Ban+Fod' and 'Ban+Ban' are all dominated, although the upper bound of these is higher than the lower bound of 'Fod+Fod'. Thus, dominance is pairwise, which means that there is no single feasible weight combination with which the overall value of any of these is higher than the overall value of 'Fod+Fod'.

In the second phase, uncertainty in the decision makers' subjective estimates in the *Socio-psychological* effects is studied. This uncertainty can be taken into account by extending the initial rating estimates to intervals. We assume an uncertainty level of ±10% of the value interval so that each initial rating estimate $v_i(x_i)$ is extended to an interval $[v_i(x_i) - 0.1,\ v_i(x_i) + 0.1]$. In the cases where the lower or the upper bound of the interval exceeds the $[0, 1]$ rating range, this bound is set to 0 or 1, respectively. One should note that any variation within these intervals is assumed to be independent of variations in the other alternatives, and the possible uncertainties in the ranges of the attributes are taken into account in the variation of the weights. Fig. 5 shows the resulting value intervals and dominance relations. In this case, 'Fod+Fod' dominates all the other alternatives except 'Prod+Fod'. Thus, the choice of the best alternative is not very sensitive to considerable variations in the participants' value estimates either.

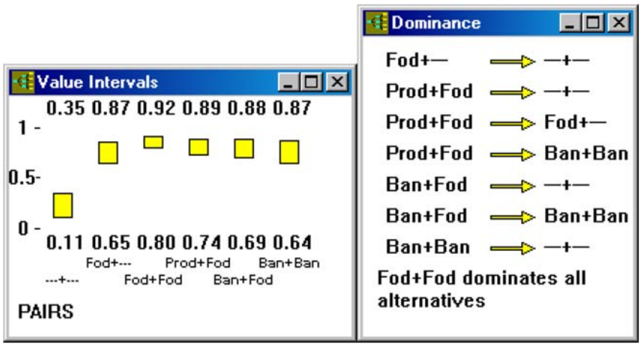

Fig. 4. Overall value intervals and dominance relations calculated with uncertainty in the weight assessment.

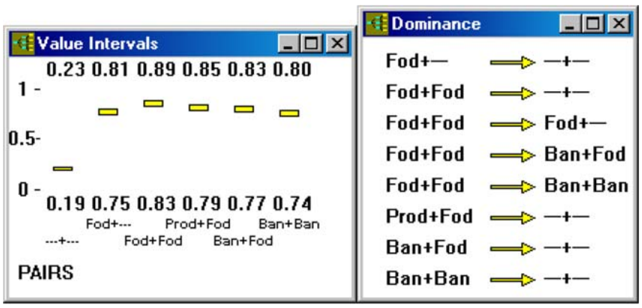

Fig. 5. Overall value intervals and dominance relations calculated with uncertainty in value estimation.

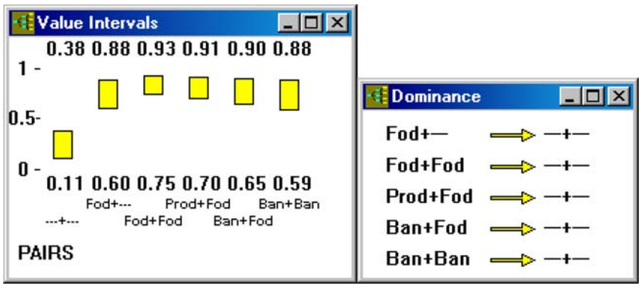

Fig. 6. Overall value intervals and dominance relations when all the possible uncertainty in the problem is taken into account.

Finally, Fig. 6 shows the overall value intervals when there is uncertainty both in the weight assessment and in the value estimation. Although the previous analyses did not show the problem to be very sensitive to these uncertainties alone, under these joint effects all the alternatives except '– – – + – – –' (i.e. *do nothing*) become non-dominated. Thus, the decision maker can safely only eliminate this alternative. However, of the non-dominated alternatives, 'Fod+Fod' still performs well. For example, it has the highest lower and upper bounds of the overall value interval (0.75 and 0.93, respectively). Thus, this analysis can also increase the confidence on the alternative 'Fod+Fod', as even in the worst-case it is not considerably worse than any of the other alternatives.

The above analysis showed how the decision maker can quickly assign uncertainties to the model. In practice the decision maker can continue the process by further adjusting the intervals on the weight ratios and on the ratings, and by studying how the overall value intervals are affected by these changes. He/she can, for example, tighten the intervals and evaluate what are the widest intervals for which 'Fod+Fod' is the only dominant alternative, and then consider whether to accept these tightened intervals to represent the allowed uncertainties in the problem.

5. Conclusion

In this paper we have described how to use interval modeling in global sensitivity analyses in multilevel value trees to analyze the effects of the total impacts of all combined uncertainties in the model. Our approach is concerned with the extreme values of the intervals which are needed, for example, if one is interested in a worst-case scenario. The proposed approach is computationally fast and the interpretation of the results is conceptually straightforward for non-mathematicians. With software, such as WINPRE, intervals can be easily given and graphical output visualizes the magnitude of the total uncertainty in an easy-to-understand way.

The intervals can be constructed in different ways. Strict maximum and minimum values will enable worst-case analyses. The decision maker can also assign error ratios to point estimates or treat the intervals as a kind of confidence interval, although in this case one needs to be careful when interpreting the results. One can assign different levels of uncertainties to all the model parameters simultaneously. By varying the intervals, the decision maker can also carry out what-if analyses. This is useful, for example, for finding the level of uncertainty allowed for the decision to still be the best in all circumstances. In a more detailed analysis the interval method can also show which factors affect the outcome the most. That is, by adding imprecision to different factors one by one, the analysis will reveal what information should be collected and how much that will reduce the uncertainty in the model.

As demonstrated in our example, the proposed approach is likely to be attractive in high-risk situations where a worst-case analysis is needed and where the decision makers might want to quickly try out different what-if analyses. Also when following the precautionary principle, a global sensitivity analysis using intervals would be useful.

Acknowledgements

Jyri Mustajoki acknowledges the financial support of the Academy of Finland (project 32641), the Finnish Cultural Foundation, and Jenny and Antti Wihuri Foundation. We also wish to acknowledge two anonymous referees for their constructive comments.

References

Ammann, M., Sinkko, K., Kostiainen, E., Salo, A., Liskola, K., Hämäläinen, R.P., Mustajoki, J., 2001. Decision analysis of countermeasures for the milk pathway after an accidental release of radionuclides. Radiation and Nuclear Safety Authority in Finland, STUK-A186, December 2001. Available from: <www.stuk.fi/julkaisut/stuk-a/stuk-a186.pdf>.

Arbel, A., 1989. Approximate articulation of preference and priority derivation. European Journal of Operational Research 43, 317–326.

Arbel, A., Vargas, L.G., 1993. Preference simulation and preference programming: Robustness issues in priority derivation. European Journal of Operational Research 69, 200–209.

Barron, F.H., 1992. Selecting a best multiattribute alternative with partial information about attribute weights. Acta Psychologica 80, 91–103.

Belton, V., Stewart, T.J., 2001. Multiple Criteria Decision Analysis: An Integrated Approach. Kluwer Academic Publishers, Boston, MA.

Bryson, N., Mobolurin, A., Ngwenyama, O., 1995. Modeling pairwise comparisons on ratio scales. European Journal of Operational Research 83, 639–654.

Butler, J., Jia, J., Dyer, J., 1997. Simulation techniques for the sensitivity analysis of multi-criteria decision models. European Journal of Operational Research 103, 531–546.

Clemen, R.T., 1996. Making Hard Decisions: An Introduction to Decision Analysis. Duxbury Press, California.

Ehrhardt, J., Weis, A. (Eds.), 2000. RODOS: Decision support system for off-site nuclear emergency management in Europe, European Commission, EUR 19144 EN. Available from: <www.rodos.fzk.de>.

Felli, J., Hazen, G., 1999. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. Risk Decision and Policy 4 (2), 79–98.

Francos, A., Elorza, F.J., Bouraoui, F., Bidoglio, G., Galbiati, L., 2003. Sensitivity analysis of distributed environmental simulation models: Understanding the model behaviour in hydrological studies at the catchment scale. Reliability Engineering and System Safety 79, 205–218.

French, S., 1995. Uncertainty and imprecision: Modelling and analysis. Journal of Operational Research Society 46, 70–79.

French, S., 2003. Modelling, making inferences and making decisions: The roles of sensitivity analysis. Top (The Journal of the Spanish Statistical and Operations Research Society) 11 (2), 229–252.

French, S., Rios Insua, D. (Eds.), 1999. Journal of Multi-Criteria Decision Analysis, Special Issue, Sensitivity Analysis in MCDA 8(3).

Goldstein, B., Carruth, R.S., 2004. The precautionary principle and/or risk assessment in World Trade Organization decisions: A possible role for risk perception. Risk Analysis 24 (2), 491–499.

Graham, J.D., 2000. Perspectives on the precautionary principle. Human and Ecological Risk Assessment 6 (3), 383–385.

Haines, L.M., 1998. A statistical approach to the analytic hierarchy process with interval judgments. (I). Distributions on feasible regions. European Journal of Operational Research 110, 112–125.

Hämäläinen, R.P., 2000. Decisionarium—Global Space for Decision Support. Systems Analysis Laboratory, Helsinki University of Technology. Available from: <www.decisionarium.hut.fi>.

Hämäläinen, R.P., 2003. Decisionarium—aiding decisions, negotiating and collecting opinions on the Web. Journal of Multi-Criteria Decision Making 12 (2–3), 101–110.

Hämäläinen, R.P., Helenius, J., 1997. WINPRE—Workbench for Interactive Preference Programming. Computer software. Systems Analysis Laboratory, Helsinki University of Technology. Available from: <www.decisionarium.hut.fi>.

Hämäläinen, R.P., Leikola, O., 1996. Spontaneous decision conferencing with top-level politicians. OR Insight 9 (1), 24–28.

Hämäläinen, R.P., Pöyhönen, M., 1996. On-line group decision support by preference programming in traffic planning. Group Decision and Negotiation 5, 185–200.

Hämäläinen, R.P., Salo, A., Pöysti, K., 1992. Observation about consensus seeking in a multiple criteria environment. In: Proceedings of the 25th Hawaii International Conference on Systems Sciences, Hawaii, vol. IV, pp. 190–198.

Hämäläinen, R.P., Lindstedt, M.R.K., Sinkko, K., 2000. Multiattribute risk analysis in nuclear emergency management. Risk Analysis 20 (4), 455–468.

Hazen, G.B., 1985. Partial information, dominance, and potential optimality in multiattribute utility theory. Operations Research 34 (2), 296–310.

Keeney, R.L., Raiffa, H., 1976. Decisions with Multiple Objectives. Preferences and Value Tradeoffs. John Wiley & Sons, Inc.

Kim, S.H., Choi, S.H., Kim, J.K., 1999. An interactive procedure for multiple attribute group decision making with incomplete information: Range-based approach. European Journal of Operational Research 118, 139–152.

Kirkwood, C.W., 1992. Estimating the impact of uncertainty on a deterministic multiattribute evaluation. Management Science 38 (6), 819–826.

Kirkwood, C.W., 1997. Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets. Duxbury Press, Belmont, CA.

Moskowitz, H., Tang, J., Lam, P., 2000. Distribution of aggregate utility using stochastic elements of additive multiattribute utility models. Decision Sciences 31 (2), 327–360.

Mustajoki, J., Hämäläinen, R.P., Salo, A., 2005. Decision support by interval SMART/SWING—Incorporating imprecision in the SMART and SWING methods. Decision Sciences 36 (2), 317–339.

Pannell, D.J., 1997. Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. Agricultural Economics 16, 139–152.

Proll, L.G., Salhi, A., Rios Insua, D., 2001. Improving an optimization-based framework for sensitivity analysis in multi-criteria decision-making. Journal of Multi-Criteria Decision Analysis 10, 1–9.

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.

Saaty, T.L., 1980. The Analytic Hierarchy Process. McGraw-Hill, Inc.

Salo, A., 2001. On the role of decision analytic modelling. In: Stirling, A. (Ed.), On Science and Precaution in the Management of Technological Risk, vol. II. Case Studies, European Commission, Joint Research Centre, EUR 19056/EN/2, pp. 123–141.

Salo, A., Hämäläinen, R.P., 1992. Preference assessment by imprecise ratio statements. Operations Research 40 (6), 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—Part A: Systems and Humans 31 (6), 533–545.

Salo, A., Hämäläinen, R.P., 2003. Preference Programming. Manuscript. Available from: <http://www.sal.hut.fi/Publications/pdf-files/msal03b.pdf>.

Salo, A., Punkka, A., 2005. Rank inclusion in criteria hierarchies. European Journal of Operational Research 163 (2), 338–356.

Saltelli, A., Chan, K., Scott, E.M. (Eds.), 2000a. Sensitivity Analysis. John Wiley & Sons, Ltd.

Saltelli, A., Tarantola, S., Campolongo, F., 2000b. Sensitivity analysis as an ingredient of modelling. Statistical Science 15 (4), 377–395.

Stam, A., Silva, A.P.D., 1997. Stochastic judgments in the AHP: The measurement of rank reversal probabilities. Decision Sciences 28 (3), 655–688.

Tarantola, S., Saltelli, A., (Eds.), 2003. SAMO 2001: Methodological advances and innovative applications of sensitivity analysis. Special Issue on Reliability Engineering and System Safety 79 (2).

Von Winterfeldt, D., Edwards, W., 1986. Decision Analysis and Behavioral Research. Cambridge University Press.

Weber, M., 1985. A method of multiattribute decision making with incomplete information. Management Science 31 (11), 1365–1371.

Weber, M., 1987. Decision making with incomplete information. European Journal of Operational Research 28, 44–57.

White, C.C., Sage, A.P., Dozono, S., 1984. A model of multiattribute decisionmaking and trade-off weight determination under uncertainty. IEEE Transactions on Systems, Man, and Cybernetics 14 (2), 223–229.