# Working Paper

## Context-Related Scaling of Human Judgement in the Multiplicative AHP, SMART, and ELECTRE

*Freerk A. Lootsma*

WP-93-39
July 1993

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria

Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

# Context-Related Scaling of Human Judgement in the Multiplicative AHP, SMART, and ELECTRE

*Freerk A. Lootsma*

WP-93-39
July 1993

*F.A. Lootsma is Professor at the Delft University of Technology, Faculty of Technical Mathematics and Informatics, 2628 CD Delft, The Netherlands*



IIASA International Institute for Applied Systems Analysis ◻ A-2361 Laxenburg ◻ Austria
Telephone: +43 2236 715210 ◻ Telex: 079 137 iiasa a ◻ Telefax: +43 2236 71313

# Foreword

This paper is part of an ongoing research in the Risk Analysis and Policy (RAP) project addressing issues of siting hazardous facilities, which is one of the most intractable public policy problems in most industrialized countries. An important part of this problem is creating credible and trusted processes for making these difficult choices. Decision analysis has traditionally served to clarify tradeoffs which underlie these kinds of decisions. This paper presents and compares different methods for eliciting preferences, which can certainly be a major contribution to systematic analyses of difficult policy choices of this sort.

This work was carried out, in part, during Freerk A. Lootsma's visit to IIASA from May to August, 1992.

# Contents

# CONTEXT-RELATED SCALING OF HUMAN JUDGEMENT IN THE MULTIPLICATIVE AHP, SMART, AND ELECTRE

F.A. Lootsma
Faculty of Technical Mathematics and Informatics
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

**Abstract**

Since decisions are invariably made within a given context, we model relative preferences as ratios of increments or decrements in an interval on the axis of desirability. Next, we sort the ratio magnitudes into a small number of categories, represented by numerical values on a geometric scale. We explain why the Analytic Hierarchy Process (AHP) and the French collection of ELECTRE systems, typically based on pairwise - comparison methods, are concerned with category judgement of ratio magnitudes, whereas the Simple Multi-Attribute Rating Technique (SMART) essentially uses the orders of magnitude of these ratios. This phenomenon, well-known in psycho-physics, provides a common basis for the analysis of the methods in question and for a cross-validation of their results. Throughout the paper, we illustrate the approach via a well-known case study, the choice of a location for a nuclear power plant.

## 1 Introduction

We are concerned with three well-known approaches for multi-criteria decision analysis (MCDA): the direct-rating method in the Simple Multi-Attribute Rating Technique (S-MART), as well as the pairwise-comparison methods in the Analytic Hierarchy Process (AHP) and in the French set of ELECTRE methods. We concentrate on the basic question which is differently answered in each of them: how to model the decision maker's preferential judgement when he considers the alternatives under a particular criterion?

The above methods are well-known in the field of MCDA, where support is given to the subjective weighing of a finite number of alternatives $A_1, \ldots, A_n$ under a finite number of conflicting performance criteria $C_1, \ldots, C_m$. SMART (Von Winterfeldt and Edwards (1985)) and the AHP (Saaty (1980), see also Zahedi (1986) and French (1988)) are firmly rooted in the American school of MCDA, where the preference intensities for the alternatives are modelled by a partial value function on the set of alternatives, under each of the criteria separately. ELECTRE (Roy (1985), see also Schärlig (1985) and Vincke (1989)) is at the origin of the French school, where the preference relations between the alternatives are modelled via a system of binary outranking relations, again under each of the criteria separately. In all these methods there is eventually an aggregation step to produce either a global value function or a global system of outranking relations, enabling the decision maker to rank the alternatives in a subjective order of preference.

In the basic pairwise-comparison experiment of the AHP and ELECTRE, two alternatives $A_j$ and $A_k$ are presented to the decision maker whereafter he is requested to judge them under a particular criterion $C_i$ and to express his indifference between them or his graded preference for one of the two. In SMART, however, the decision maker is asked to rank the alternatives $A_1, \ldots, A_n$ under the criterion $C_i$ and to refine his judgement by the assignment of grades to them. Thus, SMART enables the decision maker to keep a rather holistic view on the alternatives. In the AHP and in ELECTRE, his judgement is more fragmented.

In several papers (Lootsma (1987, 1990)) we studied the numerical model for preferential judgement in the AHP, and we proposed a multiplicative variant where the gradations of comparative judgement are put on a scale with geometric progression. We started from the assumption that the subjective weighing of the alternatives under a particular criterion is carried out in a given context, represented by an interval on the corresponding axis. This interval is partitioned into subintervals which are felt to be of the same order of magnitude; the echelons of the partitioning constitute a sequence with geometric progression. This property is well-known in psycho-physics.

In fact, the multiplicative AHP has been designed to estimate ratios of subjective values. In the basic experiment just mentioned, the decision maker is requested to estimate the relative value of $A_j$ and $A_k$ under criterion $C_i$. Instead of a real magnitude scale, however, we use a category scale requiring that his responses are restricted to the gradations indifference, weak, strict, or strong preference, or to thresholds between adjacent gradations. Finally, the above echelons enable us to assign numerical values to the gradations of comparative judgement.

In psycho-physical measurement, the ratios of audible sound or visible light intensities are usually recorded as differences on the deci-Bell scale. This means that not the ratio magnitudes themselves, but their orders of magnitude are considered. The observation suggested us to assume that a difference of grades in SMART represents the order of magnitude of a ratio of subjective values in the multiplicative AHP. In doing so, we obtain a simple, straightforward relationship between the two MCDA methods, enabling us to carry out a cross-validation of the results.

In ELECTRE, the outranking relation between the alternatives $A_j$ and $A_k$ under criterion $C_i$ depends on the difference between the physical or monetary values $g_i(A_j)$ and $g_i(A_k)$ expressing the performance of the two alternatives under $C_i$. The key question is to find certain discrimination thresholds around $g_i(A_j)$, that is, indifference, preference, and veto thresholds, so that the preference for $A_k$ with respect to $A_j$ is determined by the zone where $g_i(A_k)$ is situated. The context, however, a uniform framework when we compare the alternatives under criterion $C_i$, is not explicitly used in the model. We therefore propose to model the context as an interval on the axis corresponding with $C_i$, and to relate the choice of the thresholds to the endpoints of the range.

## 2 Categorization of price and reliability ranges

We start with the example which is frequently used to illustrate MCDA: the evaluation and the selection of a car. Usually, low costs are important for the decision maker so that he carefully considers the consumer price. and possibly the annual expenditures for maintenance and insurance. The consumer price as such, however, cannot tell us whether the car in question would be more or less acceptable to him. That depends on the context of the decision problem, that is, on the spending power of the decision maker and on the alternative cars which he seriously has in mind. In what follows, we shall be assuming that the acceptable prices are anchored between a minimum price $C_{\min}$ which he is prepared to pay, and a maximum price $C_{\max}$ which he can afford and which he does not really want to exceed. Intuitively, he will subdivide the price range $(C_{\min}, C_{\max})$ into a number of price categories (sub-intervals) which are felt to be of the same order of magnitude. We take the grid points $C_{\min}, C_{\min} + e_0, C_{\min} + e_1, \ldots$ to denote the price levels which demarcate these sub-intervals. The price increments $e_0, e_1, e_2, \ldots$ represent the echelons of the category scale under construction. In order to model the requirement that the sub-intervals must subjectively be equal, we recall Weber's law (1834) in psychophysics, stating that the just noticeable difference $\Delta s$ of stimulus intensities must be proportional to the actual stimulus level $s$. The just noticeable difference is a step of the smallest possible order of magnitude when we move from $C_{\min}$ to $C_{\max}$; we assume that it is practically the step carried out in the construction of our model. Thus, taking the price increment above $C_{\min}$ as the stimulus intensity, that is, assuming that the decision maker is not really sensitive to the price as such but to the excess above the minimum price $C_{\min}$ which he has to pay anyway, we set

$$e_\mu - e_{\mu-1} = \varepsilon e_{\mu-1} \ , \ \mu = 1, 2, \ldots, \tag{1}$$

which yields

$$e_\mu = (1 + \varepsilon) e_{\mu-1} = \ldots = (1 + \varepsilon)^\mu e_0.$$

Obviously, the echelons constitute a sequence with geometric progression. The initial step is $e_0$, and $(1 + \varepsilon)$ is the progression factor. The integer-valued parameter $\mu$ is chosen to designate the order of magnitude of the echelons.

The number of categories is rather small. because our linguistic capacity to describe the categories unambiguously in verbal terms is limited. We introduce the following gradations to identify the subsequent price categories:

cheap,
cheap/somewhat more expensive,
somewhat more expensive,
somewhat more/more expensive,
more expensive,
more/much more expensive,
much more expensive.

Thus, we have four major, linguistically distinct categories: cheap, somewhat more, more and much more expensive cars. Moreover, there are three so-called threshold categories between them, which can be used if the decision maker hesitates between the neighbouring gradations.

In earlier papers (the author (1990, 1991)) we employed a number of examples such as the progression of historical periods and planning horizons, the classification of nations according to size, and the perception of light and sound intensities, in order to show

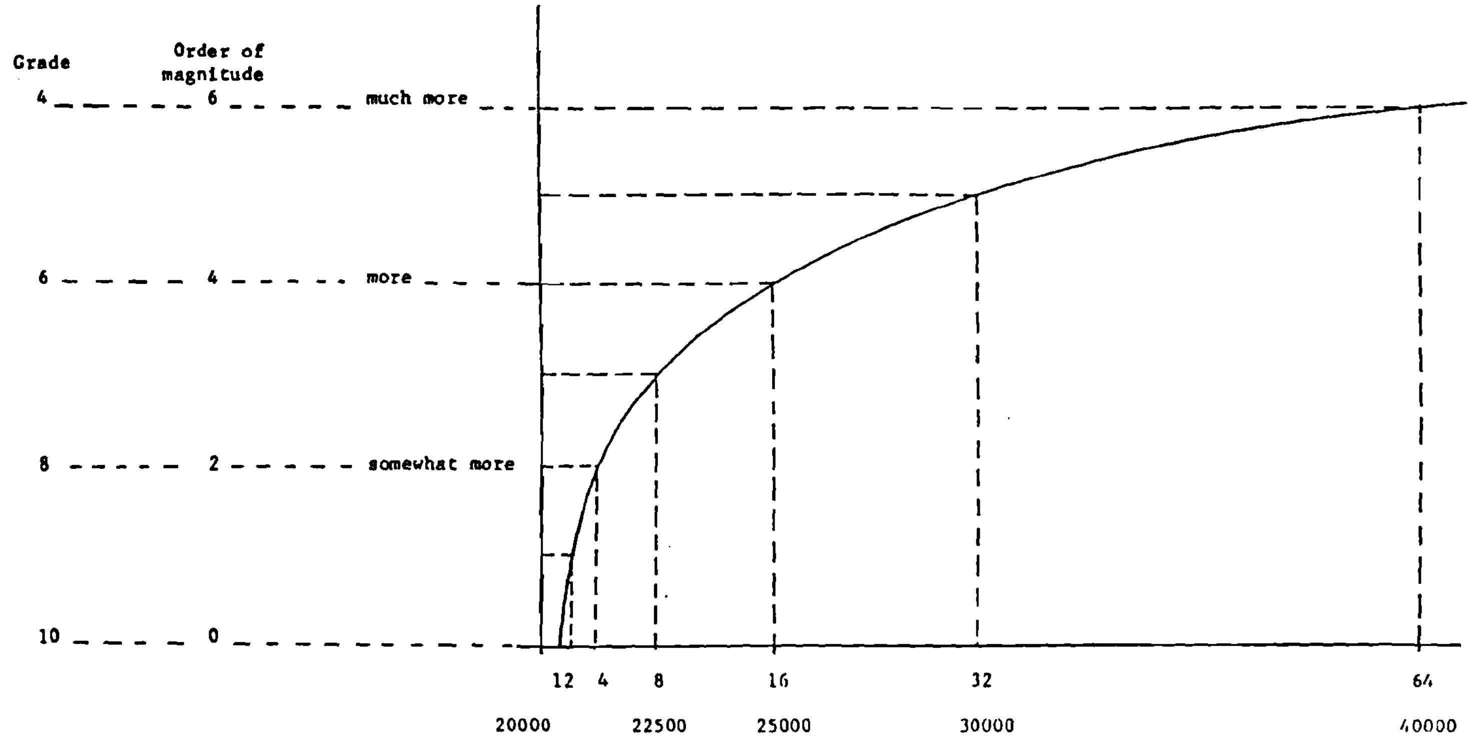


Figure 1. Categorization of a price range. Concave relationship between echelons on the dimension of price and orders of magnitude of subjective judgement.

that human beings follow the same pattern in many unrelated areas when they categorize an interval. They introduce three to five major categories, and the categorization of the range of audible sound or visible light intensities demonstrates that the progression factor $(1 + \varepsilon)^2$ is roughly 4. By the interpolation of threshold categories, they have a more refined subdivision of the given interval. Then there are six to nine categories, and the progression factor $(1 + \varepsilon)$ is roughly 2, as the progression of historical periods and the categorization of nations readily show. With these results we can easily complete the categorization of a price range. Let us, for instance, take the range between Dfl 20,000 (ECU 9,000) for a modest Renault 5 and Dfl 40,000 (ECU 18,000) for a well-equipped Renault 21 in the Netherlands. The length of the range is Dfl 20,000. Hence, setting the last price level $C_{\min} + e_6$ roughly at $C_{\max}$ we have

$$e_6 = C_{\max} - C_{\min},$$

$$(1 + \varepsilon)^6 e_0 = 20,000; \quad 1 + \varepsilon = 2,$$

$$e_0 = 20,000/64 \approx 300.$$

It is sometimes more convenient to associate the above-named qualifications, not with the sub-intervals, but with the price levels. Thus, cheap cars are roughly found at the price $C_{\min} + e_0$, somewhat more expensive cars at $C_{\min} + e_2$, etc. This will eventually lead to the following subdivision:

| | | |
|---|---|---|
| $C_{\min} + e_0$ | Dfl 20,300 | cheap cars, |
| $C_{\min} + e_1$ | Dfl 20,600 | cheap/somewhat more expensive cars, |
| $C_{\min} + e_2$ | Dfl 21,200 | somewhat more expensive cars, |
| $C_{\min} + e_3$ | Dfl 22,500 | somewhat more/more expensive cars, |
| $C_{\min} + e_4$ | Dfl 25,000 | more expensive cars, |
| $C_{\min} + e_5$ | Dfl 30,000 | more/much more expensive cars, |
| $C_{\min} + e_6$ | Dfl 40,000 | much more expensive cars. |

We can now give a more precise interpretation for the gradations of comparative judgement. A somewhat more expensive car has a price increment $e_2$, which is 4 times the price increment $e_0$ of a cheap car, etc. We use this observation to identify the so-called modifiers "somewhat more", "more", and "much more" with ratios 4:1, 16:1, and 64:1 respectively. Note that, by this convention, a car of Dfl 25,000 is somewhat more expensive than a car of Dfl 21,200 because the price increments also have the ratio 4:1. By the same token, a car of Dfl 21,200 is somewhat cheaper than a car of Dfl 25,000. We ignore the possibility of hysteresis when we invert the orientation of comparative judgement.

When the alternative cars are judged under the consumer-price criterion, the target is at the lower end $C_{\min}$ of the interval of possible prices. From this point the decision maker looks at less favourable alternatives. That is the reason why the above categorization, in principle an asymmetric subdivision of the interval under consideration, has an orientation from the lower end: the upward direction is typically the line of sight of the decision maker, at least under the given criterion. Figure 1 shows the concave form of the relationship

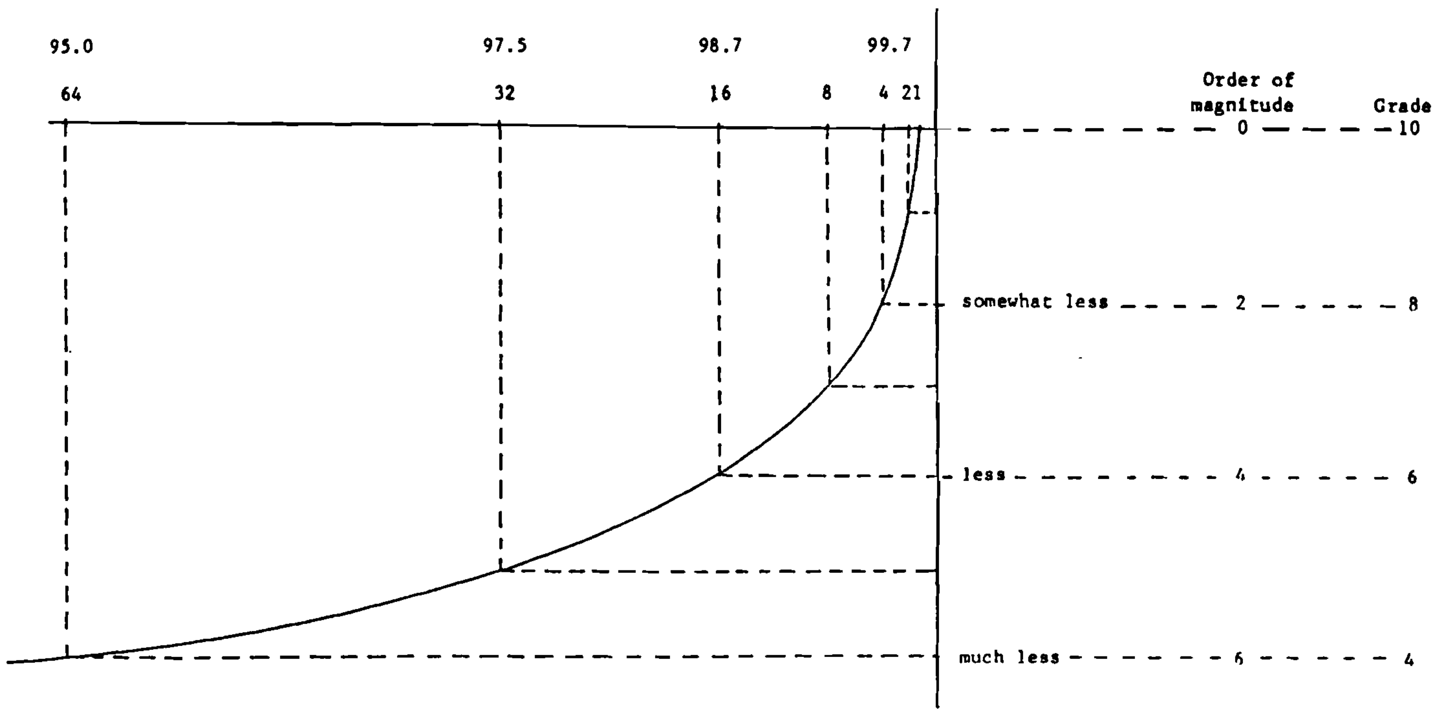


Figure 2. Categorization of an interval on the one-dimensional axis of reliability. Relationship between echelons on the dimension of reliability and orders of magnitude of subjective judgement.

between the echelons on the interval $(C_{\min}, C_{\max})$ and their order of magnitude $\mu$, a form which is well-known in psycho-physics. Mathematically, the relationship between $\mu$ and the consumer price $C$ can be written in the form

$$\mu = {}^{2}\log\left\{\frac{C - C_{\min}}{C_{\max} - C_{\min}} \star 64\right\}. \tag{2}$$

When the cars are judged under the reliability criterion, the orientation is downwards. Numerical data to estimate the reliability are usually available. Consumer organizations collect information about many types and models of cars which follow the prescribed maintenance procedures, and they publish the frequencies of technical failures in the first three or five years. Let us suppose that the decision maker only considers cars with a reliability of at least 95%, so that we are restricted to the interval $(R_{\min}, R_{\max})$ with $R_{\min} = 95$ and $R_{\max} = 100$. Following the mode of operation just described, we obtain the major categories

| | | |
|---|---|---|
| $R_{\max} - e_0$ | 99.9% | reliable cars, |
| $R_{\max} - e_2$ | 99.7% | somewhat less reliable cars, |
| $R_{\max} - e_4$ | 98.7% | less reliable cars, |
| $R_{\max} - e_6$ | 95.0% | much less reliable cars, |

because $e_0 = (100 - 95)/64 \approx 0.08$. Figure 2 illustrates the relationship between the echelons on the interval $(R_{\min}, R_{\max})$ and their order of magnitude $\mu$.

In summary, the alternatives are compared with respect to a certain target. The relative performance is inversely proportional to the distance from the target. The reader can easily verify this in the two examples just given. If we take $R_j$ and $R_k$ to denote the reliability of the alternative cars $A_j$ and $A_k$, for instance, and if we suppose that the respective reliabilities are of the order of magnitude $\mu_j$ and $\mu_k$, then the inverse ratio

$$e_{\mu_k}/e_{\mu_j} = (R_{\max} - R_k)/(R_{\max} - R_j) = 2^{\mu_k - \mu_j}$$

represents the relative performance of $A_j$ and $A_k$ under the reliability criterion. The qualifications "somewhat cheaper" and "somewhat more reliable" imply that the **inverse ratio** of the echelons (the distances to the target) is 4:1, at least under the assumption that the progression factor $(1 + \varepsilon)$ may be set to 2. The relationship between $\mu_j$ and $R_j$ takes the explicit form

$$\mu_j = {}^{2}\log\left\{\frac{R_{\max} - R_j}{R_{\max} - R_{\min}} \star 64\right\}. \tag{3}$$

## 3 Magnitude categories in the multiplicative AHP

So far, we have been working on two different dimensions: consumer price and reliability. Judgemental statements like "somewhat more expensive" and "somewhat more reliable" cannot be aggregated, however, unless we make a transition to a new, common dimension. That is the reason why we take the expression "somewhat more reliable" to stand for "somewhat more desirable" under the criterion of reliability. Similarly, we assume that the expression "somewhat more expensive" may stand for "somewhat less desirable" under the consumer-price criterion. We shall assume that the desirability of the alternatives varies over the same interval $(D_{\min}, D_{\max})$ under each of the respective criteria, at least during the decision process at hand. Moreover, we suppose that the interval $(D_{\min}, D_{\max})$ may be categorized in the same way as the intervals in the previous examples. Taking $D_j = D_{\max} - e_{\mu_j}$ and $D_k = D_{\max} - e_{\mu_k}$ to denote the desirability of $A_j$ and $A_k$ respectively, we model the preference for $A_j$ with respect to $A_k$ as the ratio of subjective stimulus values

$$\frac{V_j}{V_k} = \frac{e_{\mu_k}}{e_{\mu_j}} = \frac{D_{\max} - D_k}{D_{\max} - D_j} = 2^{\mu_k - \mu_j}. \tag{4}$$

The echelons $e_{\mu_j}$ and $e_{\mu_k}$ may be found on a geometric scale. In the multiplicative AHP, we therefore convert the gradations of the decision maker's comparative judgement into numerical values on a scale which is also geometric. Thus, we estimate the ratio $V_j/V_k$ by

$$r_{jk} = 2^{\delta_{jk}}, \tag{5}$$

where $\delta_{jk}$, an estimate of $(\mu_k - \mu_j)$, is an integer-valued index designating the gradation of the decision maker's judgement as follows:

-8 $A_j$ vastly less desirable than $A_k$,
-6 $A_j$ much less desirable than $A_k$,
-4 $A_j$ (definitely) less desirable than $A_k$,
-2 $A_j$ somewhat less desirable than $A_k$,
0 $A_j$ as desirable as $A_k$ (equally desirable),
+2 $A_j$ somewhat more desirable than $A_k$,
+4 $A_j$ (definitely) more desirable than $A_k$,
+6 $A_j$ much more desirable than $A_k$.
+8 $A_j$ vastly more desirable than $A_k$.

It is current practice to describe the gradations in terms of preference, so that the gradation index $\delta_{jk}$ can equivalently be used to designate the strength of preference in the following way:

-8 very strong preference for $A_k$ versus $A_j$,
-6 strong preference for $A_k$ versus $A_j$,
-4 strict (definite) preference for $A_k$ versus $A_j$,
-2 weak (mild, moderate) preference for $A_k$ versus $A_j$,
0 indifference between $A_j$ and $A_k$,
+2 weak (mild, moderate) preference for $A_j$ versus $A_k$,
+4 strict (definite) preference for $A_j$ versus $A_k$,
+6 strong preference for $A_j$ versus $A_k$.
+8 very strong preference for $A_j$ versus $A_k$.

Thus, we use the even values of the gradation index $\delta_{jk}$ to designate the major echelons (the major gradations) of comparative judgement, and the odd values for the threshold echelons (the threshold gradations).

The original AHP approximates the vector $V = (\ldots, V_j, \ldots, V_k, \ldots)$ of subjective stimulus values by the Perron-Frobenius eigenvector of the matrix $R = \{r_{jk}\}$. This approach has been criticized by various authors. Alternative proposals were also brought forward. Particularly logarithmic regression has been proposed, not only because it is an appropriate technique to deal with ratio information, but also on axiomatic grounds (Barzilai et al. (1987, 1991)). Thus, we approximate $V$ by the normalized vector $\overline{v}$ which minimizes the expression

$$\sum_{j<k} (\ell n \ r_{jk} - \ell n \ v_j + \ell n \ v_k)^2. \tag{6}$$

Minimization is carried out by solving the associated, linear system of normal equations with variables $u_j = \ell n \ v_j$. Obviously, the $u_j$ have an additive degree of freedom. The $v_j$ will accordingly have a multiplicative degree of freedom, which is used to single out the normalized vector $\overline{v}$ with components summing up to unity. Note that an unnormalized minimum solution is given by the geometric row means of the pairwise comparison matrix $R$.

## 4 Estimation of orders of magnitude in SMART

It frequently happens that the decision makers find it difficult to choose a gradation for their comparative judgement, particularly when the performance of the alternatives under the given criterion is expressed in physical or monetary units. The categorization of section 2 will help them to carry out the task properly. In many real-life applications we observed that the decision makers intuitively turn to such a procedure. They classify the alternatives in a small number of groups (the good ones, the bad ones, and an intermediate group) on a vaguely defined range of desirability, whereafter they judge them in pairs via inspection of the classification. The pairs are presented in random order, but the classification enables the decision makers to keep a somewhat holistic view on the set of alternatives within the context of the decision problem.

In doing so, the decision makers are not far away from the direct-rating procedure which is normally used in SMART (see Von Winterfeldt and Edwards (1986)). Thus, when they judge the performance of an alternative, they express their judgement by choosing an appropriate value between a predetermined lower limit for the worst (real or imaginary) alternative and a predetermined upper limit for the best (real or ideal) alternative. In schools and universities, such a procedure is well-known as the assignment of grades expressing the performance of the pupils or students on a category scale with equidistant echelons, between 1 and 5, between 1 and 10, or between 1 and 100 (the upper limit varies from country to country; sometimes the scale is upside-down so that the grade 1 is used to express excellent performance). Because everybody has once been subject to his or her teacher's judgement, the grades are numbers with a strong, qualitative connotation which can successfully be used in multi-criteria analysis. Concentrating on the scale between 1 and 10, we suppose that a unit step difference represents an order of magnitude difference in performance (a pupil who usually scores 9 is an order of magnitude better than a pupil scoring 8, etc). In the interval $(D_{\text{min}}, D_{\text{max}})$ we could accordingly assign the following grades to the subsequent gradations of desirability:

| | | |
|---|---|---|
| $D_{\text{max}} - e_0$ | 10 | excellent |
| $D_{\text{max}} - e_1$ | 9 | very good |
| ... | | ... |
| $D_{\text{max}} - e_6 = D_{\text{min}}$ | 4 | poor. |

In pass-or-fail decisions at schools, the grades 1,2, and 3 are normally used for a performance that cannot be compensated by high grades elsewhere. Hence, we mainly work here on a scale between 4 and 10 to rate the desirabilities between $D_{\text{min}}$ and $D_{\text{max}}$. What matters, however, is the difference between grades. Considering two alternatives $A_j$ and $A_k$ with the respective grades $g_j$ and $g_k$ assigned to them, we take the quantity

$$r_{jk} = 2^{g_j - g_k} \tag{7}$$

to estimate their relative desirability

$$2^{\mu_k - \mu_j},$$

because $g_j$ and $g_k$ are estimates of $10 - \mu_j$ and $10 - \mu_k$ respectively. Such a simple relationship between SMART and the multiplicative AHP enables us to carry out a cross-validation of the two methods. The experiments in question will be reported in later papers. Preliminary results may be found in Lootsma et al. (1993).

Since the $r_{jk}$ obtained by (7) are consistent, an unnormalized solution of (6) is given by the vector with components

$$2^{g_j}.$$

Normalization of this vector yields the desired vector of calculated weights of the alternatives.

## 5 Psycho-physical evidence

In the psycho-physical literature, the issue of how human beings judge the relationship between two stimuli, in a pairwise comparison on one single dimension, was brought up a few decades ago. Torgerson (1961) observed that human beings perceive only one quantitative relation, but they estimate differences of subjective stimulus values when they are requested to express their judgement on a category scale with arithmetic progression (equidistant echelons), and they estimate ratios of subjective stimulus values when the proposed scale is geometric. Thus, they interpret the relationship as it is required in the experiment. Which of the two interpretations is correct, cannot empirically be decided because they are alternative ways of saying the same thing.

Torgerson's observation is easy to understand if we assume that the subjective stimulus values are not identically used in the two types of experiments. In the ratio experiment with a geometric scale, human beings judge the ratio of two stimulus values. In the difference experiment with an arithmetic scale, they do not judge the ratio itself but its order of magnitude, which is essentially a logarithm of the ratio.

Psycho-physical research in the seventies and eighties (see Veit (1978) and Birnbaum (1982)) confirmed Torgerson's observation that pairwise comparative judgement of two stimuli uses one operation only in both types of experiments. Moreover, if subtraction is assumed to be the underlying operation, then ratio judgement is exponentially related to difference judgement. Since these two-stimuli experiments cannot discriminate between a one-operation and a two-operation theory, four-stimuli experiments were designed such as ratio judgement of differences of stimuli, difference judgement of differences etc., in order to analyze human perception of quantitative relations. An extensive discussion of the results so far is beyond the scope of the present paper. It is enough for our purposes to recall here that the basic operation in the comparison of the stimuli appears to be subtraction. Intuitively, this is easy to understand since orders of magnitude are easier to handle than the original magnitudes themselves (additive steps!), particularly if the magnitudes are large.

Figure 1 and Figure 2 show that the categorization of an interval is in fact a very simple procedure. In the horizontal direction one starts with a small initial step, which is repeatedly doubled until the endpoint of the interval is reached. An alternative way of looking at the procedure is to consider it as a series of bisections of the interval. In the vertical direction one just counts the number of such steps.

Let us summarize the significance of Torgerson's work in his own words: any model relying solely on ratios can be parallelled exactly by a model that solely relies on differences. In the field of decision analysis, these two models and the appropriate operations were independently established by Barzilai (1992). For the time being, a pragmatic conclusion is that the multiplicative AHP and SMART do the same thing albeit in alternative ways, that is, they rely on ratios and differences respectively, and they are exponentially related.

## 6 Criterion weights and aggregation

The assumption that the desirability of the alternative varies within the common interval $(D_{\min}, D_{\max})$ on the dimension of desirability enables us to operate with geometric means of preference ratios. Let us consider two alternatives $A_j$ and $A_k$ with their calculated weights $\overline{v}_{ij}$ and $\overline{v}_{ik}$ under criterion $C_i$. For each $i$ the preference ratio

$$\frac{\overline{v}_{ij}}{\overline{v}_{ik}},$$

expressing the relative preference for $A_j$ with respect to $A_k$ under $C_i$, is unique. Since we are dealing with ratios, it is natural to model the global preference for $A_j$ with respect to $A_k$ by the geometric mean

$$\prod_{i=1}^{m} \left(\frac{\overline{v}_{ij}}{\overline{v}_{ik}}\right)^{\overline{c}_i}, \tag{8}$$

where $\bar{c}_i$ denotes the calculated weight of the $i$-th criterion. Thus, we set the final score $\overline{f}_j$ of alternative $A_j$ to

$$\overline{f}_j = \alpha \prod_{i=1}^{m} (\overline{v}_{ij})^{\bar{c}_i}.$$

The final scores have a multiplicative degree of freedom $\alpha$. They can accordingly be normalized to sum up to unity.

The ratio (8) suggests the phrases to be used in the elicitation of the relative importance of two criteria $C_{i_1}$ and $C_{i_2}$. We may ask the decision maker to consider two real or imaginary alternatives $A_j$ and $A_k$ such that his preference for $A_j$ over $A_k$ under $C_{i_1}$ is roughly equal to his preference for $A_k$ over $A_j$ under $C_{i_2}$. These preferences are estimated by

$$2^{\delta_{jk}} \quad \text{and} \quad 2^{-\delta_{jk}}$$

respectively, where $\delta_{jk}$ designates the selected gradation of his comparative judgement. Next, assuming that these preferences do not depend on the performance of $A_j$ and $A_k$ under the remaining criteria, we ask him to estimate his preference for the two alternatives under $C_{i_1}$ and $C_{i_2}$ simultaneously. Let us take $\theta_{jk}$ to denote the selected gradation, so that his preference for $A_j$ over $A_k$ is estimated by

$$2^{\theta_{jk}}.$$

Now, taking $\omega = \omega_{i_1 i_2}$ to stand for the relative importance (the ratio of the weights) of $C_{i_1}$ and $C_{i_2}$, we obtain by the geometric mean aggregation rule that

$$\frac{\omega}{\omega+1}\delta_{jk} - \frac{1}{\omega+1}\delta_{jk} = \theta_{jk},$$

whence

$$\omega = \frac{\delta_{jk} + \theta_{jk}}{\delta_{jk} - \theta_{jk}}.$$

It will be clear that $|\theta_{jk}| < |\delta_{jk}|$ since the preference for $A_j$ over $A_k$ under the two criteria simultaneously cannot be greater than the preference for $A_j$ over $A_k$ under one of the criteria individually. It is easy to verify now that $\omega$ varies roughly between $\frac{1}{16}$ and 16 when $\delta_{jk}$ varies between -8 and 8 and $\theta_{jk}$ between $-|\delta_{jk}|$ and $|\delta_{jk}|$. The simplest geometric sequence of values between $\frac{1}{16}$ and 16, corresponding to indifference, weak, strict, strong, and very strong preference, is a sequence with progression factor 2. Thus, we obtain the following geometric scale for the major gradations in the pairwise comparison of the criteria $C_{i_1}$ and $C_{i_2}$:

| | |
|---|---|
| $\frac{1}{16}$ | $C_{i_1}$ vastly less important than $C_{i_2}$, |
| ........ | ........ |
| $\frac{1}{2}$ | $C_{i_1}$ somewhat less important than $C_{i_2}$, |
| 1 | $C_{i_1}$ as important as $C_{i_2}$, |
| 2 | $C_{i_1}$ somewhat more important than $C_{i_2}$, |
| 4 | $C_{i_1}$ more important than $C_{i_2}$, |
| ........ | ........ |
| 16 | $C_{i_1}$ vastly more important than $C_{i_2}$, |

and if we also allow threshold gradations to express hesitations between two adjacent qualifications in the above list, we have a geometric sequence with progression factor $\sqrt{2}$.

From the above considerations, we derive a simplified procedure to calculate the criterion weights in the multiplicative AHP. First, we ask the decision maker to compare the criteria in pairs and to estimate their relative importance in terms of the above gradations, possibly keeping in mind the interpretation that we have just given. Next, we use logarithmic regression to calculate criterion weights. If there is exactly one judgemental statement for each pair of criteria, the (unnormalized) weight of $C_i$ is the geometric mean of row $i$ in the pairwise-comparison matrix. Normalization so that the weights sum up to unity yields the weights $\bar{c}_i$ of the criteria $C_i, i = 1, \ldots, m$.

In SMART the calculations are even much simpler than in the multiplicative AHP. We ask the decision makers to express the importance of the criteria in grades or marks on the scale $4, \ldots, 10$ (grades lower than 4 are possible but they practically eliminate the corresponding criteria). Taking $m_i$ to stand for the mark assigned to criterion $C_i$, we obtain that the ratio of the weights of $C_{i_1}$ and $C_{i_2}$ is estimated by

$$(\sqrt{2})^{m_{i_1} - m_{i_2}},$$

so that an unnormalized weight of $C_i$ is given by

$$(\sqrt{2})^{m_i}, \tag{9}$$

whereafter normalization yields the calculated weight $\bar{c}_i$ of $C_i$.

The aggregation rules in SMART and in the multiplicative AHP are logarithmically related. If we let $g_{ij}$ stand for the grade assigned to alternative $A_j$ under criterion $C_i$, it must be true that

$$\frac{\bar{v}_{ij}}{\bar{v}_{ik}} = 2^{g_{ij} - g_{ik}},$$

so that

$$\prod_{i=1}^{m} \left(\frac{\bar{v}_{ij}}{\bar{v}_{ik}}\right)^{\bar{c}_i} = 2^{\Delta_{jk}} \tag{10}$$

where

$$\Delta_{jk} = \sum_{i=1}^{m} \bar{c}_i g_{ij} - \sum_{i=1}^{m} \bar{c}_i g_{ik}, \tag{11}$$

which demonstrates how the arithmetic-mean aggregation rule in SMART corresponds to the geometric-mean aggregation rule in SMART.

## 7 Numerical example

A well-known case study in the literature on multi-criteria analysis is the choice of a location for a nuclear power plant (Keeney and Nair (1977), Roy and Bouyssou (1983, 1992))

| Criteria | Scores | Alternative Locations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ | $A_6$ | $A_7$ | $A_8$ | $A_9$ |
| $C_1$ : Individuals affected | 8 | 57 | 40 | 25 | 48 | 44 | 23 | 52 | 11 | 18 |
| $C_2$ : River salmon lost | 7 | 6 | 6 | 6 | 0.83 | 2.55 | 0.75 | 0.45 | 4.3 | 3.65 |
| $C_3$ : Biological impact | 4 | 1.5 | 1.5 | 2 | 3.5 | 4.5 | 4.5 | 2.5 | 2 | 1 |
| $C_4$ : Socio-econ. impact | 6 | 2.5 | 2.5 | 2.5 | 3.5 | 3 | 4 | 3.5 | 4 | 3 |
| $C_5$ : High-voltage lines | 5 | 1 | 1 | 7 | 6 | 12 | 1 | 0 | 0 | 0 |
| $C_6$ : Incremental cost | 9 | 2.0 | 0 | 1.5 | 1.9 | 12.3 | 17.7 | 4.8 | 10.9 | 11.4 |

Table 1. Impacts of 9 alternative locations for a nuclear power plant under 6 performance criteria (case study by Keeney and Nair (1977)).

| Criteria | Weights | Alternative Locations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $A_5$ | $A_6$ | $A_7$ | $A_8$ | $A_9$ |
| $C_1$ : Individuals affected | 0.237 | 6 | 6.5 | 7 | 6 | 6 | 7 | 6 | 8 | 7.5 |
| $C_2$ : River salmon lost | 0.167 | 9.5 | 9.5 | 9.5 | 10 | 10 | 10 | 10 | 10 | 10 |
| $C_3$ : Biological impact | 0.059 | 8.5 | 8.5 | 8 | 6.5 | 5.5 | 5.5 | 7.5 | 8 | 9 |
| $C_4$ : Socio-econ. impact | 0.118 | 7.5 | 7.5 | 7.5 | 6.5 | 7 | 6 | 6.5 | 6 | 7 |
| $C_5$ : High-voltage lines | 0.084 | 9.5 | 9.5 | 7 | 7 | 6 | 9.5 | 10 | 10 | 10 |
| $C_6$ : Incremental cost | 0.335 | 8.5 | 10 | 8.5 | 8.5 | 5.5 | 5 | 7 | 6 | 6 |
| Final scores (SMART) | | 8.04 | 8.66 | 8.04 | 7.68 | 6.59 | 6.84 | 7.49 | 7.60 | 7.66 |
| Final scores (Mult. AHP) | | 13.7 | 21.1 | 13.7 | 10.7 | 5.0 | 5.9 | 9.3 | 10.1 | 10.5 |

Table 2. Direct scores of alternative locations for a nuclear power plant, criterion weights, as well as final scores according to SMART (arithmetic-mean aggregation rule) and the multiplicative AHP (geometric-mean aggregation rule).

| Expected utility, Keeney and Nair | | Multiplicative AHP and SMART | |
|---|---|---|---|
| $A_3$ | 0.926 | $A_2$ | 21.1 |
| $A_2$ | 0.920 | $A_3$ | 13.7 |
| $A_1$ | 0.885 | $A_1$ | 13.7 |
| $A_4$ | 0.883 | $A_4$ | 10.7 |
| $A_8$ | 0.872 | $A_9$ | 10.5 |
| $A_9$ | 0.871 | $A_8$ | 10.1 |
| $A_7$ | 0.862 | $A_7$ | 9.3 |
| $A_5$ | 0.813 | $A_6$ | 5.9 |
| $A_6$ | 0.804 | $A_5$ | 5.0 |

Table 3. Rank order of alternative locations for a nuclear power plant, calculated according to different methods for multi-criteria analysis.

in the North-West of the USA (Washington). The consultants evaluated the expected utility of 9 alternative sites on the basis of the following 6 criteria:

$C_1$: Health and safety, the annual number of human beings possibly affected by the power plant. This number is allowed to vary over a range between 0 and 200 ($\star 10^{-3}$).

$C_2$: Quantity of river salmon lost by thermal pollution. This quantity is allowed to vary over a range between 0 and 300 ($\star 10^{3}$).

$C_3$: Biological impact, measured on a qualitative scale between 0 (smallest impact) and 8 (worst possible impact), and averaged over a number of experts.

$C_4$: Socio-economic impact, also measured on a qualitative scale between 0 (smallest impact) and 7 (worst possible impact), and averaged over a number of experts.

$C_5$: Esthetical impact, measured by the length of the high-voltage lines connecting the plant to the electrical network. The length is allowed to vary over a range between 0 and 50 miles.

$C_6$: Cost, the incremental cost with respect to the cheapest possible location. The increment is allowed to vary over a range between US\$ 0 and US\$ 40 million.

Table 1 shows the impacts of the locations under the respective criteria. Moreover, we understand that the criteria have the following rank order in importance: $C_3 < C_5 < C_4 < C_2 < C_1 < C_6$, which is roughly modelled by the assignment of the scores $4, \ldots, 9$.

Since we have a range of permissible impact variations under each criterion (the ranges were required for the construction of the utility functions), we can easily convert the impacts into values on the SMART scale between 4 (poor performance that can still be compensated) and 10 (excellent performance) via the formulas in sec. 2 and sec. 4. Thus, an impact of 57 on the range (0, 200), when the objective is to minimize the impact (alternative $A_1$ under criterion $C_1$), yields the direct score

$$10 - {}^{2}\log\left(\frac{57}{200} \star 64\right) = 5.81.$$

We have arbitrarily rounded off this quantity to the nearest multiple of 0.5. Despite the choice of the step size 0.5 we maintain that we are still working with orders of magnitude, but we have somewhat reduced the coarseness of the original SMART scale. Table 2 exhibits the direct scores so obtained for all alternatives under each of the criteria, the criterion weights computed according to formula (9), and the final scores. Lastly, Table 3 shows the alternatives in the rank order obtained by Keeney and Nair as well as our rank order.

The similarity of the results is encouraging: it shows what we can obtain without asking the decision makers to contribute to the construction of utility functions. Simplification of the procedures for multi-criteria decision analysis is an urgent issue, and we have obtained practically the same rank order here, with a considerably reduced "information load" on the shoulders of the decision maker.

## 8 The discrimination thresholds in ELECTRE

In ELECTRE, a pairwise comparison of two alternatives $A_j$ and $A_k$ under criterion $C_i$ is based on the difference between the physical or monetary values expressing the performance of the respective alternatives under $C_i$. In what follows, we shall be assuming that increasing values designate a more desirable performance. In order to simplify matters even more, we shall consider two alternative cars $A_j$ and $A_k$ under the reliability criterion (see section 2). The respective reliabilities are denoted as $R_j$ and $R_k$. With $R_j$ as the reference value, the decision maker is requested to choose discrimination thresholds in order to judge $R_k - R_j$. First, there is an indifference zone around the reference value $R_j$, demarcated by a lower indifference threshold $R_j - q'(R_j)$ and an upper indifference threshold $R_j + q(R_j)$. Thus, the decision maker is indifferent between the two cars under the reliability criterion if

$$R_j - q'(R_j) \leq R_k \leq R_j + q(R_j),$$

where the decrement $q'$ and the increment $q$ are functions of $R_j$. Similarly, the decision maker has a weak preference for $A_k$ if

$$R_j + q(R_j) \leq R_k \leq R_j + p(R_j),$$

and a strict preference for $A_k$ if

$$R_j + p(R_j) \leq R_k \leq R_j + v(R_j),$$

where $R_j + p(R_j)$ and $R_j + v(R_j)$ stand for the upper preference threshold and the upper veto threshold respectively. The inequality

$$R_j + v(R_j) \leq R_k$$

implies dominance of $A_k$ over $A_j$ which cannot be compensated by the performance of $A_j$ under the remaining criteria. Usually, the increments $p$ and $v$ are also functions of $R_j$. Similarly, there are lower thresholds $R_j - p'(R_i)$ and $R_j - v'(R_j)$ in order to demarcate weak and strict preference for $A_j$ with respect to $A_k$.

Our experiences with the AHP and SMART prompt us to make some critical remarks. In the above definitions of thresholds, the context does not explicitly appear, although it could have been modelled as an interval $(R_{\text{min}}, R_{\text{max}})$. The target $R_{\text{max}}$ could have been used to define the increments $q, p$ and $v$ as functions of $R_{\text{max}} - R_j$, but the target has also been ignored. Finally, we will see that the increments $q, p$ and $v$ on the one hand, and the decrements $q', p'$ and $v'$ on the other, suggest a symmetry that does not really exist within the context of a decision problem.

When the decision maker is able and prepared to specify the context, we propose to structure the choice of discrimination thresholds as follows. In previous sections, we introduced the inverse ratio

$$\frac{R_{\max} - R_j}{R_{\max} - R_k}$$

to stand for the subjective value of $A_k$ with respect to $A_j$ under the reliability criterion, and we used the ratios 1:1, 4:1 and 16:1 to represent indifference, weak, and strict preference for $A_k$ with respect to $A_j$. Let us now take the ratio 2:1 to model hesitations between indifference and weak preference. Then $R_k$ is in the indifference zone around $R_j$ if

$$\tfrac{1}{2} \leq \frac{R_{\max} - R_j}{R_{\max} - R_k} \leq 2.$$

The upper indifference threshold, the upper endpoint of the zone, is accordingly given by

$$R_j + q(R_j) = R_j + \tfrac{1}{2}(R_{\max} - R_j),$$

and the lower indifference threshold is

$$R_j - q'(R_j) = R_j - (R_{\max} - R_j).$$

Similarly, using the ratio 8:1 to model hesitations between weak and strict preference, we find that $R_k$ is in the upper weak-preference zone if

$$2 \leq \frac{R_{\max} - R_j}{R_{\max} - R_k} \leq 8,$$

so that the upper preference threshold is given by

$$R_j + p(R_j) = R_j + \tfrac{7}{8}(R_{\max} - R_j).$$

The lower weak-preference zone is accordingly described by the inequalities

$$\tfrac{1}{8} \leq \frac{R_{\max} - R_j}{R_{\max} - R_k} \leq \tfrac{1}{2},$$

so that the lower preference threshold is given by

$$R_j - p'(R_j) = R_j - 7(R_{\max} - R_j).$$

The veto thresholds may be constructed in a similar way. Thus, $R_k$ is in the zone between the upper preference threshold and the upper veto threshold if

$$8 \leq \frac{R_{\max} - R_j}{R_{\max} - R_k} \leq 32,$$

since a ratio of 32:1 represents hesitations between strict and strong preference.

Of course, the thresholds so obtained are not necessarily situated in the range $(R_{\min}, R_{\max})$. In ELECTRE, however, the thresholds are only used to build up concordance and discordance information. Hence, they are also meaningful outside the range, even if the alternatives are only found between the endpoints.

To illustrate matters, we consider a car $A_j$ with reliability $R_j = 99\%$, in the context with $R_{\min} = 95\%$ and $R_{\max} = 100\%$. The indifference zone around $R_j$ is the interval (98, 99.5), the upper weak-preference zone is (99.5, 99.9), and the lower weak-preference zone is (92, 98).

Under the consumer-price criterion, in the context with $C_{\min}$= Dfl. 20.000 and $C_{\max}$ = Dfl. 40.000, the indifference zone around a price of Dfl. 22.000 would be between Dfl. 21.000 and Dfl. 24.000, the upper weak-preference zone would be between Dfl. 24.000 and Dfl. 36.000, etc.

The context-related choice of thresholds becomes much simpler on the SMART scale, which is commonly used in ELECTRE (the qualitative scale) when the performance of the alternatives cannot be expressed in physical or monetary terms. A ratio of 2:1 corresponds to a difference of 1 unit on the SMART scale, ratios of 8:1 and 32:1 are represented by differences of 3 and 5 units respectively. Consider again, for instance, a car $A_j$ with reliability $R_j = 99\%$ on the range between $R_{\min} = 95\%$ and $R_{\max} = 100\%$. By using formula (3), we obtain that the car would have a grade of 6.5 on the SMART scale. Then the indifference zone is between 5.5 and 7.5, the upper weak-preference zone between 7.5 and 9.5, etc. Note again that the thresholds are not necessarily situated in the range between 4 and 10.

The choice of the discrimination thresholds can accordingly be reduced to the choice of the intervals (the ranges) representing the context of the decision problem. This greatly simplifies the work of the decision maker. ELECTRE has many parameters to be set by the user (the discrimination thresholds, on each of the dimensions under consideration), so that it may be difficult to handle.

## 9 A comparative study

Both, the multiplicative AHP and SMART have been incorporated by L. Rog (Delft University of Technology) in the MCDA system REMBRANDT, using Ratio Estimation in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed. Furthermore, H. Schuyt (Delft University of Technology) compared the system with ELECTRE III at LAMSADE, Université de Paris-Dauphine. One of the test problems was the choice of a location for a nuclear power plant reported in sec. 7. Starting from base case, Schuyt employed 9 variations of the ranges and 18 variations of the criterion weights, so that there were $9 \star 18 = 162$ cases under consideration. The data of the base case may be found in sec. 4, with the following grades assigned to the criteria, however:

$C_1$: 9.5
$C_2$: 7.5
$C_3$: 4
$C_4$: 6.5
$C_5$: 5
$C_6$: 10.

ELECTRE III does not produce final scores of the alternatives. It only ranks the alternatives in a complete or incomplete order. Hence, we decided to compare only the rank orders produced by ELECTRE III and REMBRANDT, under the following range variations with respect to the base case:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | $C_1$: | 10 | - | 100, | | | | | |
| 3 | $C_2$: | 0 | - | 100, | | | | | |
| 4 | $C_6$: | -5 | - | 35, | | | | | |
| 5 | $C_6$: | -5 | - | 20, | | | | | |
| 6 | $C_6$: | 0 | - | 75, | | | | | |
| 7 | $C_1$: | 10 | - | 100, | $C_6$: | 0 | - | 75, | |
| 8 | $C_1$: | 10 | - | 100, | $C_6$: | -5 | - | 35, | |
| 9 | $C_1$: | 10 | - | 100, | $C_6$: | 0 | - | 20, | |

and with the following variations of the grades assigned to the criteria:

| | | | | |
|---|---|---|---|---|
| 2 | $C_1$: | 7.5 | | |
| 3 | $C_2$: | 9.5 | | |
| 4 | $C_2$: | 6.5 | | |
| 5 | $C_3$: | 5 | | |
| 6 | $C_4$: | 7.5 | | |
| 7 | $C_4$: | 5 | | |
| 8 | $C_5$: | 6.5 | | |
| 9 | $C_5$: | 4 | | |
| 10 | $C_6$: | 9.5 | | |
| 11 | $C_1$: | 8.5, | $C_2$: 8.5 | |
| 12 | $C_1$: | 8.5, | $C_6$: 8.5 | |
| 13 | $C_1$: | 8.5, | $C_4$: 5 | |
| 14 | $C_1$: | 8.5, | $C_2$: 6.5 | |
| 15 | $C_1$: | 8.5, | $C_2$: 6.5, | $C_6$ : 8.5 |
| 16 | $C_2$: | 6.5, | $C_4$: 5 | |
| 17 | $C_2$: | 6.5, | $C_6$: 9.5 | |
| 18 | $C_1$: | 10, | $C_2$: 8.5. | |

In all cases, the impacts of the alternatives were converted into scores on the SMART scale (the so-called qualitative scale in ELECTRE), in the same way as in sec. 7, whereafter the discrimination thresholds of sec. 8 could immediately be used to demarcate the transition from indifference to weak preference (1 unit on the scale), from weak preference to preference (3 units on the scale), etc.

This elaborate sensitivity analysis yields the frequencies of the rank order positions exhibited in the Tables 4 and 5 (in a case where two alternatives with the same final scores were competing for two consecutive rank order positions, each alternative was supposed to occupy 50% of the two positions, etc.). Obviously, alternative $A_2$ is leading for both

| | $A_2$ | $A_3$ | $A_1$ | $A_9$ | $A_8$ | $A_4$ | $A_7$ | $A_6$ | $A_5$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 69 | 7 | 0 | 1 | 22 | 0 | 0 | 0 | 0 |
| 2 | 9 | 53 | 0 | 18 | 18 | 0 | 1 | 0 | 0 |
| 3 | 12 | 22 | 31 | 10 | 17 | 6 | 1 | 0 | 0 |
| 4 | 10 | 14 | 9 | 31 | 5 | 27 | 3 | 0 | 0 |
| 5 | 1 | 1 | 35 | 14 | 15 | 12 | 15 | 8 | 0 |
| 6 | 0 | 2 | 16 | 24 | 14 | 30 | 14 | 1 | 0 |
| 7 | 0 | 0 | 6 | 1 | 8 | 19 | 65 | 1 | 0 |
| 8 | 0 | 0 | 3 | 0 | 0 | 7 | 0 | 90 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |

Table 4. Rank order of 9 alternative nuclear power plant locations calculated by the REMBRANDT program (multiplicative AHP and SMART) under 9 range variations and 18 variations of criterion weights. The entries represent the frequencies (in percentages of the $9 \star 18 = 162$ cases) of the rank order positions, so that the following rank order emerges: $A_2 > A_3 > (A_1 \simeq A_9 \simeq A_8 \simeq A_4) > A_7 > A_6 > A_5$.

| | $A_2$ | $A_9$ | $A_8$ | $A_3$ | $A_1$ | $A_4$ | $A_7$ | $A_6$ | $A_5$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 49 | 7 | 38 | 0 | 1 | 0 | 6 | 0 | 0 |
| 2 | 13 | 30 | 18 | 6 | 19 | 0 | 15 | 0 | 0 |
| 3 | 26 | 28 | 6 | 7 | 17 | 1 | 16 | 0 | 0 |
| 4 | 8 | 1 | 6 | 37 | 14 | 12 | 16 | 7 | 0 |
| 5 | 4 | 5 | 5 | 25 | 36 | 17 | 4 | 5 | 0 |
| 6 | 0 | 16 | 14 | 13 | 8 | 32 | 5 | 12 | 0 |
| 7 | 0 | 14 | 16 | 12 | 5 | 7 | 18 | 28 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 30 | 21 | 49 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |

Table 5. Rank order of 9 alternative nuclear power plant locations calculated by ELECTRE III under 9 range variations and 18 variations of criterion weights. The entries represent the frequencies (in percentages of the $9 \star 18 = 162$ cases) of the rank order positions, so that the following rank order emerges: $A_2 > (A_9 \simeq A_8) > (A_3 \simeq A_1 \simeq A_4 \simeq A_7) > A_6 > A_5$.

methods, $A_6$ and $A_5$ are at the bottom. Note that Roy and Bouyssou (1992), with their choice of the discrimination thresholds, obtained a rank order with $A_3 \simeq A_4 \simeq A_8$ at the top and $A_6 > A_1 > A_5$ at the bottom. REMBRANDT is more informative. The main diagonal of Table 4, with 5 entries above 50%, determines the position of 5 alternatives: $A_2 > A_3$ at the top, $A_7 > A_6 > A_5$ at the bottom. This brings up the question of whether it is rewarding for the users of ELECTRE to supply the thresholds, or whether psycho-psychical arguments are strong enough to justify the context-related setting of the thresholds. More experiments may further clarify the issue. The pertinent question at the end of the paper is, of course, whether sophisticated methods like the AHP and ELECTRE have an "added value" with respect to SMART which counterbalances the more complicated elicitation of human judgement. The discussion about the issue is still open.

**Acknowledgement** The research of the present report made a significant step forward when the author visited IIASA (Laxenburg, Austria, May - August 1992) in order to participate in the project "Technological and ecological risk management in Eastern Europe". It is a pleasure to acknowledge IIASA members for their stimulating discussions, as well as members of LAMSADE, Paris, for their critical questions and comments.

## References

[1] J. Barzilai. On the Use of the Eigenvector in the AHP. Proceedings of the $X$-th Int. Conf. on MCDM, Taipei, Taiwan, July 1992, Volume I, pp. 291-300.

[2] M.H. Birnbaum, Controversies in Psychological Measurement. In B. Wegener (ed.), Social Attitudes and Psycho-physical Measurement. Hillsdale, N.J., 1982, pp. 401-485.

[3] S. French, Decision Theory, an Introduction to the Mathematics of Rationality. Ellis Horwood, Chichester, 1988.

[4] R.L. Keeney and K. Nair, Evaluating Potential Nuclear Power Plant Sites in the Pacific Northwest using Decision Analysis. In D.E. Bell, R.L. Keeney, and H. Raiffa, Conflicting Objectives in Decisions. IIASA, Laxenburg, Austria, 1977, Chapter 14.

[5] F.A. Lootsma, Modélisation du Jugement Humain dans l'Analyse Multicritère au moyen de Comparaisons par Paires. R.A.I.R.O./Recherche Opérationelle 21, 241-257, 1987.

[6] F.A. Lootsma, Numerical Scaling of Human Judgement in Pairwise-Comparison Methods for Fuzzy Multi-Criteria Decision Analysis. In G. Mitra (ed.), Mathematical Models for Decision Support. Springer, Berlin, 1988, pp. 57-88.

[7] F.A. Lootsma, The French and the American School in Multi-Criteria Decision Analysis. R.A.I.R.O./Recherche Opérationnelle 24, 263-285, 1990. A short version appeared in A. Goicoechea, L. Duckstein, and S. Zionts (eds.), Multiple Criteria Decision Making. Springer, New York, 1992, pp. 253-268.

[8] F.A. Lootsma, Scale Sensitivity and Rank Preservation in a Multiplicative Variant of the AHP and SMART. Report 91-67, Faculty TWI, Delft University of Technology, Delft, The Netherlands, 1991.

[9] F.A. Lootsma, L. Rog, and M. Kok, Choice of Working Languages in a European Working Group for Multi-Criteria Decision Aid. To appear in the Journal of Information Science and Technology, 1993.

[10] B. Roy, Méthodologie Multicritère d'Aide à la Décision. Economica, Collection Gestion, Paris, 1985.

[11] B. Roy and D. Bouyssou, Comparison of two Decision-Aid Models applied to a Nuclear Power Plant Siting Example. Cahier 47, LAMSADE, Université de Paris-Dauphine, 1983.

[12] B. Roy et D. Bouyssou, Comparaison de Deux Études d'Aide à la Décision à propos d'un Cas de Localisation d'une Centrale Nucléaire. Document 72, LAMSADE, Université de Paris-Dauphine, 1992.

[13] Th. L. Saaty, The Analytic Hierarchy Process, Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York, 1980.

[14] A. Schärlig, Décider sur Plusieurs Critères. Presses Polytechniques Romandes, Lausannes, 1985.

[15] W.S. Torgerson, Distances and Ratios in Psycho-physical Scaling. Acta Psychologica XIX, 201-205, 1961.

[16] C.T. Veit, Ratio and Subtractive Processes in Psycho-physical Judgement. Journal of Experimental Psychology: General 107, 81-107, 1978.

[17] Ph. Vincke, L'Aide Multicritère à la Décision. Editions de l'Université de Bruxelles, Belgique, 1989.

[18] D. von Winterfeldt and W. Edwards, Decision Analysis and Behavioral Research. Cambridge University Press, Cambridge, UK, 1986.

[19] F. Zahedi, The Analytic Hierarchy Process - a Survey of the Method and its Applications. Interfaces 16, Nr. 4, 96-108, 1986.