European Journal of Operational Research 25 (1986) 200–215
North-Holland

Comparison of two decision-aid models applied to a nuclear power plant siting example

B. ROY and D. BOUYSSOU
LAMSADE, Université de Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, F-75775 Paris, France

Abstract: The aim of this paper is to examine on a non-theoretical ground to what extent outranking and MAUT decision-aid approaches differ.

For this purpose, we chose a study using utility theory conducted by Keeney and Nair, dealing with a nuclear plant siting problem. We had to determine what the study would have been if it had been conducted with the use of the ELECTRE III model.

In this 'experiment', we are not interested in the practical problem for its own sake but in:
- the way to build criteria;
- the representation of a decision-maker's preferences;
- the use of the model and the nature of the derived prescription.

Confronting the two models, we study:
- the differences that they induce when facing a real problem and building a set of data,
- their respective part of arbitrariness, weakness, realism, robustness,
- the convergence or divergence of their results,

and insist upon the differences between 'descriptive' and 'constructive' approaches.

Keywords: Decision theory, multi-attribute utility theory, outranking methods

1. Introduction

1.1. The two competing models

Let us consider a situation where a decision is necessary and where several criteria are involved. The analyst who has to help an actor in such a decision process by using as rigorous a method as possible, generally has the choice between several approaches, which involve several ways of viewing the real world and can lead to significantly different models. The objective of the present study is to compare two of these models that are frequently used, and thus to shed light on two different currents of thought that have been developing on either side of the Atlantic.

The first of the two models derives from multi-attribute utility theory. This theory is based on a set of axioms referring to a highly coherent and complete preference system, considered as an objective reality, not influenced by the analyst. His task is therefore supposed to consist merely of delimiting such a preference system and making it explicit. To this end he has to consider that probability distributions can always be used to analyse the uncertainty affecting the evaluations of the various consequences of each solution, alternative, programme or possibility, what we will call actions, relevant to the decision problem. The analyst may then assess using this probabilistic description, partial utility functions u_i (the subscript i referring to an attribute or to a 'point of view'), and aggregate them into a global utility function u. It is then a logical consequence of the set of axioms (cf. Von Neumann and Morgenstern (1947), Fishburn (1970) and, for a critical discussion, Allais (1954)) that the expected value of the global utility is a criterion representing the prefer-

Received October 1983; revised June 1984

ence system. More precisely, for any two actions a, a':

$$E(u(a')) > E(u(a)) \quad \Leftrightarrow \quad a' \, P \, a,$$
$$E(u(a')) = E(u(a)) \quad \Leftrightarrow \quad a' \, I \, a$$

(where P and I represent respectively strict preference and indifference relations). Accordingly, we will call this expected utility criterion a 'true-criterion'.

The second model does not claim to deal with an objective reality to be 'described', but with the relationship with reality that the actors of the decision process have or wish to have. This model is thus a construction designed to illuminate possible decisions by means of pragmatic ideas and intentional actions. It is therefore difficult to connect this model with a set of axioms. In addition to probability distributions, it uses dispersion thresholds and discrimination thresholds as a way of defining what is uncertain but also what is imprecise and ill-defined in the evaluation of the consequences of the actions. This model no longer refers to a complete and coherent preference system. It considers instead that, given any two actions a and a', and given their evaluations in terms of different criteria, each of the following two statements

"a' is to be considered as at least as good as a" ($a' \, S \, a$),

"a is to be considered as at least as good as a'" ($a \, S \, a'$)

can be either accepted, or refused, or, in ambiguous cases, appraised on a scale of credibility. Moreover, the acceptance or refusal of one of the two statements does not imply any information as to the acceptance or refusal of the other; if both statements are refused, the two actions are said to be incomparable.

The definition of such a relation S – which is called an outranking relation (see Roy (1971)) – involves not only the thresholds mentioned above, but also diverse variables ('indices of importance' and veto thresholds), whose function is to reflect the respective part to be played by each criterion. The formulas defining S are constructed in such a way as to respect certain qualitative principles, and, in particular, they rule out the possibility that a major disadvantage on one criterion could be compensated for by a large number of minor advantages on other criteria. They do not imply that S should necessarily be transitive or complete. The only justification for such formulas is the application of common sense to these principles.

In contrast with expected utility, S does not in general provide a clear ranking of the actions in the form of a complete preorder. In this approach, the systematic search for such a preorder cannot be justified, and, accordingly, the model only leads to the establishment of a *partial* preorder. A detailed 'robustness' analysis then allows one to determine which of the comparisons of actions are convincingly justified by the model in spite of the element of arbitrariness in the allocating of values to certain of the parameters (thresholds, indices of importance,. . .).

Further details of these models and their theoretical background can be found in Keeney and Raiffa (1976) and Roy (1977, 1978).

1.2. The methodology of the comparison

In order to compare the two models and, more generally, the two corresponding approaches, we examined a particular example, the siting of a nuclear power-plant on the North-West Coast of the United States. The Washington Public Power Supply System (WPPSS) requested Woodward–Clyde Consultants to carry out a study on this subject a few years ago. In many ways, this study seems to be a very good example of the application of the first of the above-mentioned approaches. It has been described in a number of papers, most notably by Keeney and Nair (1976) and Keeney and Robillard (1977).

After an initial stage of the study involving a large number of alternatives and attributes, the set of potential sites was reduced to 9. In order to judge and compare them, 6 points of view were chosen, leading to 6 partial utility functions (and consequently 6 criteria if one is arguing in terms of expected values). Our aim was to carry out the work that could have been done using the outranking model – henceforth model S – instead of the utility one – model U. The description below covers the different stages of the construction of model U, and for each one shows the corresponding stages in model S. The data of the situation will be given at the same time as the description, which will consist of three parts:

- the modelling of the partial preferences on each of the 6 points of view, in other words the construction of the criteria;
- the aggregation model defining the global preferences;
- the recommendations themselves.

Given that we could not obtain information either from experts or from the WPPSS management, we were often obliged to make deductions exclusively on the basis of the information available. As our aim was not to carry out another study but to compare the two models, this disadvantage had little influence on our work.

1.3. The objectives of the comparison

We had three objectives in comparing the two different models applied to the same decision situation:

(a) to emphasize the different ways in which the two models explored reality and drew on what are officially (and mistakenly) called 'data' (data are more often 'built' than 'given');

(b) to understand better the extent to which the two models are arbitrary, vulnerable, realistic or robust (all elements necessary for assessing their respective degrees of reliability);

(c) to appreciate better how and when the two models produce similar or different recommendations.

It would certainly have been interesting to attempt to place the comparison on another level: that of their contribution to the decision process, in other words, their acceptability to the different actors and their impact on the course of the process. However, this would have required an experimental study of a different nature from the present one.

The final section of this paper will be devoted to an assessment of the study in terms of these three objectives.

2. The criteria

2.1. Introduction

The designers of model U used 6 relevant points of view for comparing the sites, which we will accept for the purpose of the present study, assuming that the WPPSS was willing to impose them. The 6 points of view are:

1: the health and security of the population in the surrounding region;

2: the loss of salmonids in streams absorbing the heat from the power-station;

3: the biological effects on the surrounding region (excluding the salmonid loss);

4: the socio-economic impact of the installation;

5: the aesthetic impact of the power lines;

6: the investment costs and the operating costs of the power-station.

(Further details may be found in Keeney and Nair (1976)).

The description of the consequences of an action s (the installation of a power-station on site s) connected with any one of the 6 points of view is clearly not simple. Here again, we based model S on the description carried out by Keeney and Nair in the perspective of model U. We will give details of this description in the next paragraph. But first we must emphasize what such a description consists of, and how one deduces from it a representation of the preferences in model U vis-à-vis each point of view. We must also indicate how model S differs in these respects. We will thus see that, in each approach, a distinctive sub-model of preference is constructed. This sub-model constitutes what is usually called a criterion; it will be denoted g_i for the point of view i.

In model U, it is an *a priori* condition that the consequences of an action s be describable in terms of 6 random variables $X_i(s)$ $(i = 1, \ldots, 6)$.Each variable is regarded as an attribute linked to the action in question. The carrying out of this action must be accompanied by a realisation of $X_i(s)$ by means of a random draw according to its probability distribution. The particular value $x_i(s)$ thus realized must encapsulate on its own all the information to be taken into account concerning the point of view considered. The first step must therefore consists in determining this information in concrete fashion, in order to be able to define the attribute and then make the probability distribution explicit. But since the different distribution may be probabilistically dependent, the general case must be studied in terms of the joint distribution of the 6 random variables.

This explains why the preference system that the set of axioms refers to is based on the comparison of such multidimensional probability distributions. In the particular case we are considering, but also in general when dealing with real decision-aid problems, it is accepted in practice that:

- the random variables $X_i(s)$ are probabilistically independent;
- the preference system benefits from two simplifying hypotheses: preferential independence and utility independence (cf. Keeney and Raiffa (1976) and Keeney (1974)).

These two hypotheses[1] together with the classical axioms of utility theory renders the following procedure legitimate:

- the analyst questions the person who seems to possess the preference system to be represented, in order to assess a partial utility function $u_i(x)$ related to the point of view i;
- he makes explicit the marginal probability distribution of the attribute $X_i(s)$;
- he calculates the expected value of this partial utility for each of the actions: $g_i(s) = E[u_i(X_i(s)]$;
- in the preference system to be represented, the bigger $g_i(s)$ is, the better s is, other things being equal.

In this case, it is meaningful to compare two actions s and s' by referring only to point of view i. The comparison is carried out in terms of the numbers $g_i(s)$ and $g_i(s')$. The function g_i is then a true-criterion in the sense ascribed to this term in Section 1.1 (for further details, see Roy (1985, Chapter 9)).

This possibility of comparing any two actions – other things being equal – is a prerequisite for model S. The points of view i must indeed be designed in such a way that these ceteris paribus comparisons constitute an appropriate departure point for the relationships that the analyst must establish between the actors (possibly the decision-makers) and their vision of reality. Since the preference system of these actors is no longer regarded as pre-existing in this reality, the existence and the definition of the criteria g_i can no longer be a direct consequence of its observable properties. These criteria should, in particular, be defined with relation to the nature of the information available on each point of view and by taking into account as much as possible the elements of imprecision, uncertainty and indetermination which affect this information. Obviously, there is nothing to prevent a given criterion from taking the form of an expected utility criterion. However, in many cases, probability distributions may appear insufficient for taking into account the whole significance of these elements. In addition, the framework of true-criterion may seem too narrow to describe the conclusions of such comparisons. Model S therefore leads one to substitute pseudo-criteria for the true-criteria of model U.

The pseudo-criterion induces on the set of actions a structure generalising the semi-order one (see Luce (1956)) by introducing two discrimination thresholds: q_i (the indifference threshold) and p_i (the preference threshold). For the point of view of criterion g_i, we have:

- s' indifferent to s iff $|g_i(s') - g_i(s)| \leqslant q_i$;
- s' strictly preferred to s iff $g_i(s') > g_i(s) + p_i$;
- s' weakly preferred to s iff $q_i < g_i(s') - g_i(s) \leqslant p_i$.

In the general case, the thresholds q_i and p_i may be dependent on $g_i(s)$ (or on $g_i s')$). Further details may be found in Roy and Vincke (1984) and Jacquet-Lagrèze and Roy (1981).

In model U, the criteria g_i are defined as soon as one has assessed the utility functions u_i and chosen a probabilistic description for each of the attributes X_i. The procedure culminating in the determination of $g_i(s)$ and the two associated discrimination thresholds characterising each of the pseudo-criteria of model S is completely different (cf. Roy (1985, Chapters 8 and 9). It is based on an analysis of the consequences belonging to the point of view i and on our ability to model them, either as a single number constituting what we will call a 'single point evaluation' (which may or may not be allocated an imprecision threshold), or as several numbers constituting a 'non single point evaluation', each of these numbers possessing (potentially) an index of likelihood having the meaning, for example, of a probability. Since the only information available to us was the probabilistic description of model U, such a thorough analysis was not possible here. Consequently, we based the definition of the criteria involved in model S on common sense, although we tried to stay as close as possible to what we believe this part of study could have been in a real

---

[1] The theory includes tests designed to check their realism, but putting them into practice involves difficulties that make the results unconvincing.

context, with experts and decision-makers. The type of reasoning used in the next sections is therefore more important than the precise numerical values elicited.

2.2. Case of two criteria (no. 1 and 5) based on quantitative single point evaluation

Amongst the 6 attributes used to describe the consequences of the actions in model U, there were two, X_1 and X_5, which were not regarded as random numbers, but as numbers that were known with certainty. In other words, a site s is characterised in terms of these two points of view by two figures , $x_1(s)$, $x_5(s)$; and this is why we speak in this case of single-point quantitative evaluations. The evaluation on point of view no 5 being in many ways simpler, we will choose this one to start with.

The figure $x_5(s)$ represents the length of the high-tension wires (needed to connect the power-station to the grid) which will harm the environment if the power-station is constructed. For the 9 potential sites, it varies from 0 to 12 miles [2]. Although the measure of this attribute was not regarded as a random variable, it proved necessary to define a utility function $u_5(x_5)$ in order to take this attribute into account in the global preference model. The assessment of this function was carried out using the classical 50–50 lottery technique (cf. Raiffa (1968) and Keeney and Nair (1976)). The results obtained implied a linear expression:

$$u_5(x_5) = 1 - \frac{x_5}{50}.$$

It follows that the true-criterion g_5 of model U is simply

$$g_5(s) = 1 - \frac{x_5(s)}{50}.$$

Within model S, a criterion associated with this point of view could have been defined by letting $g_5(s) = x_5(s)$. Nevertheless, this number does not seem to be precise enough, for one to be able to say that, if two sites s and s' are characterized, respectively, by

$$x_5(s) = 10, \quad x_5(s') = 9,$$

then site s' can necessarily be regarded (other things being equal) as significantly better than site s. The difference of one mile may indeed not seem convincing, given the uncertainty in the situating of the powerlines and, especially, the arbitrariness inherent in the choice of the sections of line to be taken into consideration. We did not have access to the information necessary for evaluating the influence of these factors, and we consequently assumed that $x_5(s)$ was not known within an interval whose size grew with the distance involved but remaining no less than 1 mile for short distances. It seemed reasonable to choose a very low rate of growth: 3% (a rate of 10% would not have changed the results). This amounts to saying that $g_5(s) = x_5(s)$ is ill-determined over an interval of the form:

$$[g_5(s) - \eta_5(g_5(s));\ g_5(s) + \eta_5(g_5(s))]$$

with $\eta_5(s) = 1 + \frac{3}{100} g_5(s)$.

The function η_5 characterizes what is called a dispersion threshold (cf. Roy (1985, Chapter 8)). General formulas (cf. Roy and Bouyssou (1983, Appendix 4)) can be used to deduce the two discrimination thresholds which complete the definition of the pseudo-criterion g_5:

indifference threshold:

$$q_5(g_5(s)) = 1 + \tfrac{3}{100} g_5(s),$$

preference threshold:

$$p_5(g_5(s)) = 2.0618 + 0.0618\ g_5(s).$$

The certain number $x_1(s)$ is an official index: the 'site population factor'. This index provides a measure of the total population whose health and security might be affected by the construction of a power-station on the site, and is expressed as a function of the distance of the population from the power-station. The index varies in this case between 0.011 and 0.057. Still considering the 50–50 lottery technique, a linear form was again employed for the utility function. Given extreme values for x_1 of 0 and 0.2, we have:

$$u_1(x_1) = 1 - 5x_1,$$

and hence the true criterion of model U:

$$g_1(s) = 1 - 5x_1(s).$$

For model S, once again it would have been natural to set $g_1(s) = x_1(s)$. Even more than $x_5(s)$,

---

[2] All the numerical data used in models U and S can be found in Roy and Bouyssou (1983).

$x_1(s)$ seems to be imprecise and arbitrary. This number is the outcome of an 'aggregation operation' whose aim is to represent a distribution characterizing a set of people located at various distances from the power-station by means of a single number. The problem is that this distribution may change with time. The type of this 'aggregation operation' is not the only one that can be imagined; and indeed the very way in which it is applied can result in variations. Accordingly, it seemed to be reasonable to adopt a dispersion threshold equal to $\frac{10}{100}x_1$. The indifference and preference thresholds characterizing the pseudo-criterion $g_1(s)$ have, under these conditions, the following values:

$$q_1(g_1(s)) = 0.1\, g_1(s), \quad p_1(g_1(s)) = \tfrac{2}{9} g_1(s).$$

2.3. *The case of two criteria (no. 3 and 4) based on non-single point qualitative evaluations*

To define the attributes X_3 and X_4, Keeney and Nair introduced two qualitative scales having respectively 8 and 7 adjacent intervals. The nature of the biological or socio-economic impact, covered by each interval, was determined by means of relatively concrete and precise descriptions of the future situation. For each of the two attributes and for each site s, approximately 10 experts were asked to use such descriptions to characterize the outcome which, in their view, seemed most probable in the hypothesis of the power-station being constructed on that site. The proportion of votes received by each interval was used to define the (subjective) probability distributions of $X_3(s)$ and $X_4(s)$.

Two utility functions, $u_3(x_3)$ and $u_4(x_4)$ were then assessed (using a particular technique adapted to the qualitative nature of these scales, cf. Keeney and Nair (1976)), $g_3(x_3)$ and $g_4(x_4)$ corresponding respectively to the expected utility of $X_3(s)$ and $X_4(s)$.

Once again, it is important to point out that we would have used a similar method to evaluate the biological and socio-economic impacts on the potential sites. The evaluation obtained by Keeney and Nair (a distribution of the experts' opinions, involving in general more than one interval of the scale in question) is called a 'non single-point one'. In order to define $g_3(s)$ and $g_4(s)$, only one of the intervals considered by the experts must be chosen. We selected the interval nearest the centre, that is the one which divides the experts most equally into those who are at least as optimistic and those who are at least as pessimistic as this value. Given the nature of the scales in question, constant discrimination thresholds were adopted. After examining the distributions of the experts' opinions, we used

$$q_3 = 1, \quad p_3 = 2,$$
$$q_4 = 0, \quad p_4 = 1.$$

2.4. *Case of a first criterion (no. 2) based on non single point quantitative evaluations*

X_2 is more complex than the other attributes studied up till now. The total quantity Q of salmonids which might be destroyed following the construction of a power-station was not relevant on its own to the appraisal of the 'loss of salmonids'. Given the sensitivity of certain ecological equilibria, the destruction of 10 000 salmonids in a river containing 20 000 cannot be regarded as equivalent as the loss of the 10 000 in a river containing 300 000. It was therefore necessary to analyse the consequences in terms of two factors:

- the total number Y of salmonids living in the river;
- the percentage Z of salmonids destroyed.

An exhaustive study (cf. Keeney and Robillard (1977)) led the authors to distinguish between large rivers ($Y > 300\,000$) and small ones ($Y < 100\,000$) there were no medium-sized rivers in this particular study. For the large rivers, the attribute studied X_2 could be taken into account simply by using the absolute number $Q = Y \cdot Z$ by means of a utility function defined by:

$$u_2(X_2) = 0.568 + 0.432\, u_Q(Q)$$

with

$$U_Q(Q) = 0.7843(\mathrm{e}^{(0.00274(300-Q)} - 1)$$

(Q being expressed in thousands).

For the small rivers, on the other hand, it proved necessary to take Y and Z into account separately, by means of two partial utility functions $u_Y(Y)$ and $u_Z(Z)$ (cf. Roy and Bouyssou (1983, Appendix 3)), the utility of X_2 being deduced from them by:

$$u_2(X_2) = u_Y(Y) + u_Z(Z) - u_Y(Y) \cdot u_Z(Z).$$

To calculate the expected value $g_2(s)$, the authors of model U assumed that:

- for each site s, Y took on a value $y(s)$, known with certainty;
- Z was a normal random variable with a standard deviation equal to half its expected value.

In order to implement model S, we would probably not have undertaken so complex a study to define criterion g_2. Doubts about the results of this work may be all the more justified given that:

- the probability distributions of variables Y and Z were not defined with as much care as the utility function, and
- the expected utility $g_2(s)$ (which orders the 9 sites in exactly the same way as the numbers $E(Q(s))$) does not seem to reflect very faithfully the qualitative principles adopted at the beginning of the utility analysis.

We would instead have tried to analyse why, given two rivers containing exactly y and y' salmonids, it was more damaging to destroy q of them in the first – assumed here to contain the least fish – than a slightly larger number q' in the second. Then we would have explored qualitative considerations to try to connect q' with q, y and y' in such a way that the damage done in the two rivers was of the same magnitude. One could, for instance, have examined whether a simple formula such as $q' = q \cdot (y'/y)^\alpha$ was capable – with α appropriately chosen between 0 and 1 – of representing the experts' opinions on such cases of equivalent amounts of damage. On the sole basis of the analysis done for model U, we considered it possible to define criterion g_2 from the above formula, by adopting two different versions of this criterion corresponding respectively to:

$$\alpha = \tfrac{1}{2}: \quad g_2'(s) = \frac{q}{\sqrt{y}} = z\sqrt{y},$$

$$\alpha = 0: \quad g_2''(s) = q = z \cdot y.$$

(The values of the criteria g_2 are calculated, in model S, by setting $z = \bar{z}(s)$).

The above reasoning was effected without taking into account the difficulties of evaluating y and predicting z for each river. The large value adopted for the standard deviation of Z and the necessity of coping with the imprecision affecting y led us to adopt a broad dispersion threshold which we fixed as 0.5 $g_2'(s)$ and 0.5 $g_2''(s)$. We thus have:

$$q_2' = 0.5\, g_2'(s), \qquad p_2' = 2g_2'(s),$$
$$q_2'' = 0.5\, g_2''(s), \qquad p_2'' = 2g_2''(s).$$

2.5. *Case of a second criterion (no. 6) based on non-single point quantitative evaluations*

The authors of model U considered that the investment and operating costs of a power-station located on a site could be appraised relatively to the costs of the cheapest site s_2. The attribute $X_6(s)$ therefore reflects a differential cost. It was supposed that the insufficient knowledge affecting this cost could be modelled by treating $X_6(s)$ as a normal random variable with a standard deviation equal to a quarter of its expected value [3]. This expected value was estimated by the values $\bar{x}_6(s)$ varying from 0 to 17.7 (in millions of dollars per year, cf. Roy and Bouyssou (1983, Appendix 3)). Let us point out that it is sure that $X_6(s_2) = 0$.

The criterion $g_6(s)$ of the model U is the expected utility of this random differential cost. Again invoking the lottery technique, the utility function $u_6(x_6)$ was defined as

$$u_6(x_6) = 1 + 2.3(1 - e^{0.009\, x_6}).$$

Once again, we would probably have constructed model S in different way. Since it is not the same actors who are responsible for the investment and running costs, we would perhaps have introduced a criterion for each of them. But because we cannot analyse these costs in detail in the present study, we will merely set

$$g_6(s) = \bar{x}_6(s).$$

Lacking a more objective foundation, we can use the following reasoning to determine dispersion thresholds. Firstly, the values of $\bar{x}_6(s)$ which were suggested contain the assumption that the investment and running costs that are not included in the differential cost will actually lead to the same expenses on site s_2 as on any other site s. This is obviously a source of sufficient error to

---

[3] The costs are supposed to correspond to a standard type of construction which is considered fixed. No trade-offs with criterion 1 are explicitely considered.

cast into doubt the whole idea that a site s' is more economical than a site s when $\bar{x}_6(s) - \bar{x}_6(s')$ is small. We decided, on the basis of this single hypothesis, that the 'real' differential cost had to be regarded as ill-determined on an assymetrical interval: $[\bar{x}_6(s) - 1;\ \bar{x}_6(s) + 2]$.

Secondly, the calculation of $\bar{x}_6(s)$ follows on from the evaluation of multiple factors which all involve specific expenses for site s. But the study carried out on each site remains brief until the construction is actually decided. In other words, these costs are not necessarily the only ones: they are relatively imprecise and possibly too optimistic. The margin of error resulting is asymmetric and its size is proportional to $\bar{x}_6(s)$ itself. The factors involved here seem to have no connection with the ones taken into account previously. We shall therefore assume that the effects can be added together. We have the following dispersion threshold:

$$\left[\bar{x}_6(s) - 1 - 0.1\,\bar{x}_6(s),\ \bar{x}_6(s) + 2 + 0.5\,\bar{x}_6(s)\right].$$

Thus

$$q_6(g_6(s)) = 1.1 + 0.11\,g_6(s),$$

$$p_6(g_6(s)) = 3.33 + 0.67\,g_6(s).$$

3. Aggregation of the criteria and global preference

3.1. Introduction

Having in this way defined the true-criteria of model U and the pseudo-criteria of model S, we will now present the part of the model dealing with their aggregation. In the present section, we will briefly describe the parameters involved in the aggregation phase of each model. The following two sections will be devoted to the evaluation of these parameters.

Assuming that the WPPSS's preference system is a pre-existing entity, that it conforms to the axioms of utility theory, that the hypotheses of independence mentioned in Section 1.2 are acceptable, and that the responses to the questions posed in order to assess the partial utility functions were governed by this preference system implies (using a general theorem – cf. Keeney and Raiffa (1976)) that this preference system is representable by means of a true-criterion $g(s)$ defined in terms of the criteria $g_i(s)$ by one of the following two expressions:

$$g(s) = \sum_{i=1}^{i=6} k_i \cdot g_i(s) \quad \text{with} \sum_{i=1}^{i=6} k_i = 1, \tag{1}$$

$$g(s) = \frac{1}{k}\left[\prod_{i=1}^{i=6} (1 + k \cdot k \cdot g_i(s)) - 1\right] \tag{2}$$

with

$$k \neq 0, \quad k \geqslant -1, \quad k = \prod_{i=s}^{i=6} (1 + k \cdot k_i) - 1. \tag{3}$$

This last expression of $g(s)$ was the one chosen by Keeney and Nair (we will see the reasons why in Section 3.2). In order to complete the characterization of model U, it is consequently sufficient to assess the coefficients k_i (whose values increase with the relative importance attached to criterion i, once the utility functions have been defined) and to deduce the value of k from them by solving equation (3), which normally has only non-zero root greater than -1 (cf. Keeney and Nair (1976)).

In model S – which corresponds to ELECTRE III (cf. Roy (1978)) – the aim is no longer to use the pseudo-criteria $g_i(s)$ to determine a true-criterion, or even a pseudo-criterion. The more modest aim is to compare each site s to site s' [4] on the basis of their values on each g_i, taking into account the threholds q_i and p_i, and hence to adopt a position on the acceptance, the refusal or, more generally, the credibility of the proposition:

"site s is at least as good as site s' ".

As we pointed out in Section 1.2, this credibility depends on pragmatic rules of simple common sense, rules which are mainly based on notions called concordance and discordance. These notions allow one:

– to characterize a group of criteria judged concordant with the proposition studied, and to assess the relative importance of this group of criteria within the set of the 6 criteria;

---

[4] This pairwise comparison remains feasible even when hundreds of alternatives are taken into account. See for example Roy et al. (1983).

– to characterize amongst the criteria not compatible with the proposition being studied, those which are sufficiently in opposition to reduce the credibility resulting from the taking into consideration of the concordance itself, and to calculate the possible reduction that would result from this.

In order to be able to carry out such calculations, we must express in explicitly numerical fashion:

– the relative importance k_i accorded by the decision-maker to criterion i in calculating the concordance; let us merely indicate here that these numbers have virtually no influence except for the order that they induce (because of their addition) on the groups of criteria involved in the calculations of concordance;

– the minimum level of the discordance giving to criteria i the power of withdrawing all credibility from the proposition being studied, in the case when this criterion is the only one of the 6 which is not in concordance with the proposition: this minimum level is called the veto threshold of criterion i; it is not necessarily a constant, and therefore we will denote it $v_i[g_i(s)]$.

It is important to emphasize that model S is different from model U in that the indices of importance (and also the veto thresholds) are not values stemming from the observation of a pre-existing variable but values designed to convey deliberate positions adopted by the decision-maker, positions which are mainly of a qualitative nature. It follows that the techniques to be applied in order to evaluate the parameters we have just discussed for both models reflect two different attitudes towards reality (cf. 5.1) even more than the criteria do.

In each model, there is a considerable amount of arbitrariness affecting the value chosen. The recommendations must consequently take into account the robustness of the results towards these factors. They nevertheless depend strongly on the underlying model.

3.2. *Modulation of the importance of the criteria*

Within model U, the most classical method to assess the scaling constants k_i consists in comparing lotteries (see Raiffa (1969) for a review of other available methods).

Let us denote $\tilde{x}_i$ and $\underset{\sim}{x}_i$ the respective values used to scale the partial utility function u_i between 0 and 1. We have $u_i(\underset{\sim}{x}_i) = 0$ and $u_i(\tilde{x}_i) = 1$. Let us consider the following two multidimensional lotteries. The first one, L_1, is a degenerate lottery resulting for sure in an 'imaginary site' [5] which receives the worst evaluations on all the criteria except j, where its evaluation is $\tilde{x}_j$. The second lottery, L_2, give rise to another imaginary site whose evaluation is either the best possible on all the criteria with probability p, or the worst possible on all the criteria with probability $(1-p)$.

The expected utility of L_2 is p; and the utility of L_1 of k_j in the multiplicative representation (2) – and indeed also in the additive one (1). If the decision-maker is able to determine that particular probability p which guarantees indifference between the two lotteries, we can state $k_j = p$.

By iterating this procedure, one can therefore – in principle – assess the 6 coefficients k_i, and hence k, the solution to equation (3).

The lotteries to be compared here are multidimensional, unlike the ones used to assess the partial utility functions. Even with the help of sophisticated interview techniques to assess the probability p, it is difficult to escape the conclusion that this sort of comparison of imaginary sites is extremely complex, and that the decision-maker may be unable to reply to such questions in a reliable fashion. In order to try to avoid this obstacle, the designers of model U used a more indirect assessment technique comprising:

– an ordering of the coefficients k_i;

– an estimation of trade-offs between attributes;

– an estimation of the coefficients k_i.

This procedure, which is described in detail in Roy and Bouyssou (1983, Appendix 6) and Keeney and Nair (1976), is still based on lottery comparisons of type L_1 and L_2. It is therefore vital not to attribute an illusory precision to the values of the k_i estimated in this way.

The designers of model U used in the end:

$$k_1 = 0.358, \quad k_2 = 0.218, \quad k_3 = 0.013,$$
$$k_4 = 0.104, \quad k_5 = 0.059, \quad k_6 = 0.400.$$

One can observe that $\Sigma_{i=1}^{6} k_i = 1.152 \neq 1$, which justifies the choice of the multiplicative structure (cf. Keeney (1974)).

---

[5] This imaginary site is also 'idealized' since its consequences are supposed to be perfectly determined by the probability distribution.

Solving equation (3) then gives $k = -0.3316$ [6].

In model S, the only influence of the indices of importance is the ranking they impose on the different criteria or groups of criteria. If we had carried out the study, we would probably have tried to assess such a ranking interactively with the decision-makers of the WPPSS. We would then have tried to find various sets of indices of importance compatible with these merely ordinal considerations.

Without access to the decision-makers, we had to try to 'translate' the information conveyed by the utility function concerning the relative importance of the criteria into indices of importance, to attempt to produce a comparable system of values and hence to ensure that the comparison of the results of the two methods was still meaningful. The technique used is detailed in Roy and Bouyssou (1983, Appendix 7). Let us simply point out that the k_i in model U do not have an immediate interpretation in terms of the relative importance of the criteria (cf. Keeney and Raiffa (1976) and Zeleny (1981)). The magnitude of the scale and the shape of the partial utility function both affect the k_i values. This relative importance seemed to us to be reflected more accurately by the range of variation of the different ratios:

$$R_{ij} = \frac{\partial g/\partial g_i}{\partial g/\partial g_j}, \quad i,j = 1,\ldots,6, \tag{4}$$

where g is given by formula (2) and the g_i are as defined in Part 2.

One can qualitatively interpret the value of R_{ij} as the gain needed on criterion j to compensate a loss on criterion i. For example, if R_{ij} is always greater than 1 for all possible values of g_j and g_i, it seemed reasonable to us to consider that criterion i was intrinsically more important than criteria j within model S. We examined the variation ranges of the ratios R_{ij} which led us to employ eight sets of indices of importance (cf. Roy and Bouyssou (1983, Appendix 7)) covering collectively the same value system as the one convoyed by model U. In fact, we considered that the k_i were so imprecise in model U and that this translation was so inherently arbitrary that it became unrealistic to try to maintain a single set of indices.

3.3. The veto thresholds

As veto thresholds convey deliberate and 'intentional' positions, they cannot be 'assessed'. This explains why we would probably have produced the same kind of work as the one reported here had the study been a real one. Once the decision-maker is satisfied with the qualitative principles underlying the partially compensatory character of model S, one can then ascribe numerical values to the different thresholds in empiric fashion, taking into account the relative importance of the criteria, the distribution of the site evaluations over the criteria, and the size of the various preference thresholds. Given an inevitable arbitrariness in the choice of these numerical values, one generally then carries out a systematic robustness analysis on these coefficients.

Model U being compensatory, it was not possible to deduce from the available information qualitative considerations that would have helped to determine the veto thresholds. Therefore, it is principally our particular perception of the problem which is reflected in this choice. However, the robustness analysis showed that the values chosen had little influence on the results within a fairly wide range of variation. It seemed reasonable in all cases to take the thresholds $v_j(g_j(s))$ as multiples of the preferences thresholds $p_j(g_j(s))$ (not that there is necessarily any fixed link between these two figures). We imagined that the less important the criterion the larger the value of the coefficient α_j such that $v_j(g_j(s)) = \alpha_j p_j(g_j(s))$. In particular, the veto thresholds for criteria 3 (biological impact), 5 (aesthetic impact) and 4 (socio-economic impact) were chosen so as to have no influence. On the first level of analysis, we used the following values:

$$v_1(g_1(s)) = 6p_1(g_1(s)),$$
$$v_2(g_2(s)) = 2.5\, p_2(g_2(s)),$$
$$v_3(g_3(s)) = 4p_3(g_3(s)),$$
$$v_4(g_4(s)) = 4p_4(g_4(s)),$$
$$v_5(g_5(s)) = 20p_5(g_5(s)),$$
$$v_6(g_6(s)) = 1.7\, p_6(g_6(s)).$$

---

[6] The results in this paper are the ones we obtained by calculating from the data published in the articles quoted. They are slightly different from those given by Keeney and Nair (1976).

4. Contents and presentation of the recommendations

4.1. Introduction

We have, in model U: $g(s) = [\prod_{i=1}^{6}(1 + kk_i g_i(s)) - 1](1/k)$.

The values of k and of the k_i were given in Section 3.2 and the form of the $g_i(s)$ in Part 2. One can therefore obtain the number $g(s)$ and using the principles of the true-criterion, rank the sites on the following basis:

$$s' \text{ preferred to } s \quad \Leftrightarrow \quad g(s') > g(s)$$
$$s' \text{ indifferent to } s \quad \Leftrightarrow \quad g(s') = g(s),$$

and hence deduce the recommendations.

In model S, the situation is different. As mentioned above, this model seeks to establish a fuzzy outranking relation between the actions, that is to evaluate the proposition "s' is at least as good as s" on a credibility scale. A distillation procedure is then used to rank the actions on the basis of this fuzzy relation (see Roy (1978)). Two total preorders thus emerge, which behave in opposite ways when confronted with those actions which are hard to compare with another group of actions (one of the preorders tends to put them before this group, and the other after).

The intersection of these two preorders leads to a partial preorder emphasizing the actions which have an ill-defined situation in the ranking. This incomparability must be accepted, since model S explicitly acknowledges the imprecise, and even arbitrary, nature of some of the data used. The quality and reliability of the recommendations depend therefore to a considerable extent on a systematic robustness analysis.

4.2. The results

One can summarize the results of model U as in Table 1[7]:

The ranking obtained is therefore a complete ordering.

Table 1

Rank	Site	$g(s)$
1	S_3	0.926
2	S_2	0.920
3	S_1	0.885
4	S_4	0.883
5	S_8	0.872
6	S_9	0.871
7	S_7	0.862
8	S_5	0.813
9	S_6	0.804

The authors of model U carried out a sensitivity analysis on this ordering. Nevertheless, the fact that they disposed of an axiomatic basis and that they had obtained the various data (shapes of utility functions, values of the k_i) by questioning persons supposed to represent the decision-maker[8], led them to effect an analysis only of 'marginal'[9] modifications of the data. This resulted in a virtually complete stability of the ordering vis-à-vis these modifications (cf. Keeney and Nair (1976)).

The robustness analysis is a crucial part of model S. We present in Roy and Bouyssou (1983, Appendices 9 and 10) the overall robustness analysis (which involves more than 100 different sets of parameters) and the results obtained. Knowing the arbitrariness of the evaluation of some of the parameters, we considered that an entire subset of the space of the parameters was in fact plausible, a subset which we checked systematically in order to make our conclusions as reliable as possible.

We will merely observe here that, of all the possible sources of variation, the form of criterion 2 selected (g_2' or g_2'') has the greatest influence. In Roy and Bouyssou (1983, Appendix 10), we showed that, with the exception of the form of criterion 2, the stability of the results is good when confronted with variations that cannot be considered marginal. The robustness analysis bore principally on the indices of importance (8 sets), the discrimination thresholds (criteria 2 and 6) and the veto threshold (criteria 2, 3 and 6) (cf. Roy and Bouyssou (1983, Appendix 9).

The totality of these results may be presented, in very brief and qualitative form, as two graphs, corresponding respectively to the g_2' form and the g_2'' form of criterion 2 (the influence of the other

---

[7] See footnote 6.

[8] In fact, most frequently the research team themselves.

[9] Marginal, by opposition to a cross-linked variation of all the parameters in the model. Here, each parameter varies separately, within a variation range which is not necessarily small.

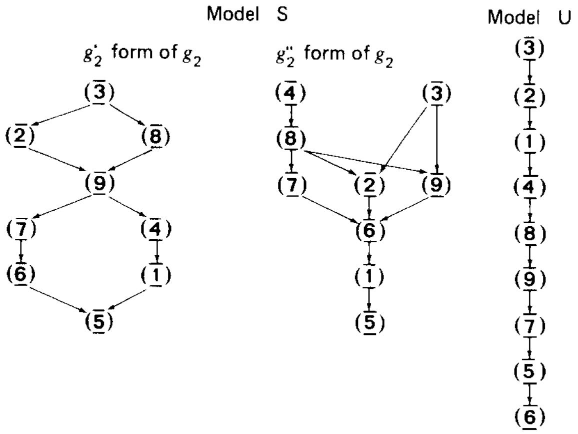

Figure 1. Representation outranking graphs. The transitivity arcs have been omitted; two sites not connected by an arc (not considering the transitivity ones) are incomparable. The graph given for model U is a diagrammatic representation of Table 1.

parameters being less important). Figure 1 shows representative outranking graphs.

4.3. The recommendations

It should be emphasized that the reason the WPPSS requested this study was to select which of the 9 sites were most likely to be chosen by the administration for the construction of the power-station. The WPPSS was interested in two sorts of information:

- the sites which could be totally eliminated at this stage of the decision process, from any further considerations;
- the sites among those remaining that would be the most likely to be considered the best in future, more detailed studies.

The study of the ranking provided by model U shows that S_5 and S_6 can safely be eliminated from further stages of the study, and that S_3 and S_2 are in the leading positions with S_1 and S_4 just behind (cf. Table 1 and Figure 1).

The analysis of the results of model S (cf. Figure 1 and Roy and Bouyssou (1983, Appendix 10) shows that there is a remarkable stability at the bottom of the ranking, with S_5, S_6 and S_1. Site S_3 is in the leading place, whatever form of criterion 2 is chosen. S_2, S_8 and S_4 are just behind, whereas S_7 and S_9 are to be found in a zone of instability in the middle.

Like the authors of model U, we would have recommended S_3, if the WPPSS had required that only one site be chosen. On the other hand, there is a major divergence between the two models concerning the position of S_1 and, to a certain extent, S_8 (we will come back to this point in Section 5.3).

Underlining the fact that the case has not been studies here for its sake, we will now try to give partial answers to the three questions mentioned in Section 1.3.

5. Conclusions

5.1. The origin and the treatment of the data

In model U, the procedures used to assess the different parameters involved in the definition of the global utility function (partial utility functions $u_i(s)$, coefficients k_i) follow logically from the set of axioms underlying the analysis. These axioms imply that lottery comparisons can always be used to carry out this estimation.

This position is unassailable on the formal level, but the number of questions raised – and their complexity – imply that the decision-maker (or his representative - cf. Section 4.2) is obliged to collaborate closely with the analyst. The legitimacy of these techniques is inseparable from the hypothesis that a complete system of preference pre-exists in a form which is implicit but which is nevertheless in line with the axioms in the decision-maker's mind[10]. It must also be assumed that the replies given by this decision-maker or his representatives are in fact governed by such an implicit system, and that this system is not likely to be fundamentally altered during the dialogue with the analyst. The urgency of the decision problem to be solved and the analyst's experience then create the necessary conditions for the disclosure of these attitudes which are represented in terms of a utility function. When certain opinions brought up are in contradiction with the axioms defining the coherence, it is assumed that the normative character of the axioms (completeness, transitivity, independence) is sufficiently obvious for the decision-maker to *adapt* his views to them

---

[10] In actual studies, the decision-maker is supposed to be able to express a set of fundamental attitudes compatible with the axioms. Comparing complex actions is then equivalent to an extrapolation of those attitudes, whose validity is guaranteed by the set of axioms (cf. Bouyssou (1984)).

(cf. Morgenstern (1979)). In such a perspective – unlike that prevailing in most of the other social sciences – the axioms of the formal model are also behavioural axioms – and, when necessary, normative axioms. This attitude underlies most of the studies based on model *U*. It explains why analysts place such great confidence in the data they gather and why they virtually never fundamentally question them when the sensitivity analysis is carried out.

The same is true when evaluating the consequences of the actions. The probability distributions provided by the experts are thus rarely questioned, even when they are clearly imprecise and/or arbitrary (cf. criteria 2 and 6 of the power-station study). One again, 'marginal' sensitivity analyses are carried out that imply generally a high level of stability in the ranking obtained.

Model *S* has no axiomatic basis, and consequently it is often difficult to interpret certain parameters used in it (veto thresholds, indices of importance). Only considerations based on common sense allow the decision-maker and the analyst to give them a numerical value. This explains why the results produced by model *S* are significant only when the analyst has carried out a major robustness analysis, systematically exploring the numerical values of the parameters compatible with the qualitative 'data' he started with. This procedure should not be considered as merely a palliative for the lack of axiomatic foundations and the lack of sophisticated techniques for assessing the parameters, but constitutes instead one of the original features of the approach, which consists of trying to design a preference system and not of trying to represent an existing system in the most accurate way possible.

The difference observed between those two prescriptive approaches in the way they obtain the data are in fact connected with a much deeper division: the one between a model drawing validity from a 'descriptive' aim of representing a pre-existing relation and a model whose validity is based on a 'constructive' aim of designing an acceptable preference relation in collaboration with the decision-maker [11]. Sophisticated assessment procedures only draw meaning with relation to a given reality, which must be adhered to as closely as possible.

In order to be in a position to apply utility theory, it must also be assumed that all the imprecise, uncertain or arbitrary elements in the evaluation of actions on the various consequences can be taken into account by means of probability distributions. Such a hypothesis is necessary for the expected value of this distribution on a utility scale to be regarded as a true-criterion.

In those cases where the principal aim is to help the decision-maker cope with a risk, a probability distribution can afford a satisfactory modelling of the evaluation of an action. When analysing the losses of salmonids in a river (criterion 2), one might try above all to study the risk of these species totally disappearing from it. If a well-established probability distribution is available for describing the phenomenon, expected utility may appear an adequate criterion.

In contrast, even if, a priori, it is possible to use probabilistic tools the model the cost of a power-station (the definition of which is not free from ambiguity—cf. Section 2.5) by closely modelling each of those of its elements (rate of inflation, cost of construction material and fission material, etc.) that might influence the cost of the project, this information is probably not very useful to the decision-maker. What is important is not to know a probability distribution on cost with a possibly misleading precision, but be able to say whether one action can be considered as significantly cheaper (or more expensive) than another. In this situation, arguing in terms of dispersion thresholds would seem necessary, as in all cases, where one is dealing more with conceptual looseness and imprecision than with a really random phenomenon. Model *S* does not assume any a priori restrictions on the nature of the imprecision and uncertainty affecting the evaluation of actions, and seeks to translate these phenomena as a pseudo-criterion.

However, these approaches are not exclusive, and indeed one can imagine using model *S* with a criterion based on an expected utility surrounded by thresholds. Model *S* substitutes pseudo-criteria for true-criteria, and this is as much the result of a refusal to restrict 'nondeterminism' to randomness as the result of the role played by the idea of criterion in designing the preference relation. This model is intended to 'construct' rather than 'de-

---

[11] It is important to stress that the terms 'descriptive' and 'constructive' do not apply to the models themselves but to their justification and the attitude in which they are elaborated.

scribe', and therefore starts from a criterion that allows one to compare two actions – other things being equal – unlike model *U*, where the fact of referring to a pre-existing reality (theoretically) obliges one to test hypotheses of independence associated with the preference structure before being able to talk of a criterion (cf. Section 2.1).

Because of this, the pseudo-criteria base the comparison of actions in model *S*, whereas the true-criteria of model *U* represent it.

The use of a pseudo-criterion follows on from the caution, and even the skepticism, with which the analyst using model *S* regards his methodology. He cannot use existing preferences as fixed points, and can only deduce that there is a convincing preference when the often approximative tools he is using leave him in no doubt–hence the use of a 'buffer-zone' embodied in the discrimination thresholds. As for model *U*, it supposes that a preference relations pre-exists, and that the information gathered using the function $u_i(x_i)$ is sufficiently reliable to allow it to be 'extrapolated' to more complex lotteries in exact fashion (for example, in the case of the cost, the function $u_6(x_6)$ is assessed from even-chance lotteries whereas the calculations are carried out using normal laws.).

2. Robustness and fragility of the approaches

The distinction between a 'constructive' attitude and a 'descriptive' one illustrates the relative advantages and disadvantages of models *U* and *S*. If the decision-maker is clearly identified and possesses a sufficiently precise and stable preference structure, one can certainly adopt a purely descriptive attitude. Nevertheless, we consider that in most real decision-aid problems, an attitude of a constructive nature is inevitable.

Every decision forms part of the social structure of the organisation, which if often complex and conflictual,meaning that often the only single decision-maker one can talk about is a fictional entity (see Roy (1985, Chapter 2)). It is then difficult to assume a collective group of decision possesses a pre-existing and coherent preference.

In fact, the designers of model *U* did not assess some parameters included in the global utility function by questioning the decision-maker(s) of the WPPSS (cf. Section 4.2), but by using judgements provided by the study team itself. This practice, which does not seem unusual in studies based on model *U* (given the difficulty and the number of the questions asked) can cause reasonable doubt as to the reliability of the assessment procedures of the utility function: it implies that sensitivity analyses of the same scope as for model *S* have to be carried out.

Once one has accepted the advantages – and even the necessity – of a constructive approach, one can understand better the implications of an axiomatic basis for decision-aid models. For many people, the attraction of an axiomatic basis is the legitimacy it apparently confers to their work. But this legitimacy is valid only for the 'theory', and not for the 'model' which is an 'interpretation' and a putting into practice of the 'theory'. Model *U* is based on a formal theory for representing an existing preference system. It is hard to imagine what a design theory of a preference system could be – a theory that would underly model *S*. If the axiomatic basis legitimises the theory, it does not follow that it does the same for the model. The legitimacy of the model must be sought in the effectiveness with which it enables the actors to arrive at convictions (possibly upsetting preconceptions) and to communicate with other people. A decision-aid model must not be merely a formal theory, but must form the basis for an interaction with reality and for an action on reality.

Finally, let us point out that model *U* can conceivably be used in a constructive perspective. This is in fact what is really done in most studies. However, model *U* should be considered in this case independently of its axiomatic basis: one should study the reliability of the assessment procedures of the partial utility functions and of the constants k_i as tools designed to construct and/or enrich the decision-maker's preference relation between the actions.

Many of the misunderstandings in comparing models *S* and *U* seem to stem from the fact that model *U* is designed in terms of a constructive attitude but only draws a particular legitimacy from its axiomatic basis if it derived from a descriptive attitude.

We do not believe that normative conclusions can be drawn from this study concerning models *S* and *U* as potential tools for decision-aid. Each model has advantages in certain domains – the usefulness of both has already been pointed out in numerous studies.

It should also be recognised that the choice of a model very often depends on 'cultural' factors and 'decision-making customs' which cannot be analysed in a formal way.

More generally, our study shows that the problem of the validation and the legitimacy of decision-aid models requires a major re-thinking. The concept of 'predictive power' cannot apparently act as the basis for validity tests in this domain – unlike the situation in many other disciplines.

3. Agreement amongst recommendations

In Section 4.3, we observed that, if there was a certain agreement in the recommendations on site S_3, there were also differences: the positioning of site S_1, in particular, was controversial. Model U ranked S_1 as amongst the best sites studied, while model S recommended that it be dropped from later stages of the study. In the same way, site S_8 is considered as a 'good' site in model S, but appears in the middle of the ranking in model U.

These disagreements in the two models reflect the contrasts in the qualitative principles underlying them, especially concerning the reliability of the differences between the evaluations on the different criteria and the more or less compensatory nature of their aggregation. Site S_1 (cf. Roy and Bouyssou (1983, Appendices 3 and 5)) is evaluated very highly on most of the criteria (g_3, g_4, g_5, g_6), but receives the worst possible evaluation on health and security (g_1) and salmonid loss (g_2). Model S, being partially compensatory, ranks such a profile near the bottom whereas model U (perfectly compensatory) places the site among the best, because of its very good scores on many criteria.

Inversely, site S_8 may be interpreted an an average 'compromise' site (cf. Roy and Bouyssou (1983, Appendices 3 and 5)), and is well-placed in model S; but in model U, it appears lower down, behind other sites where good performances on certain criteria compensate very bad ones on others.

In addition, conclusions of too great a generality should not be drawn from the good agreement of the recommendations on site S_3. An intuitive examination of the evaluations of this action shows that it seems to be a good site in terms of the information available. It is therefore 'normal' for S_3 to be in the first place in both methods. A good part of the agreement obtained in thus peculiar to the problem studied (in another problem, a site of type S_1 could have appeared at the top in model U).

Given such a fundamental opposition in the qualitative principles underlying the two models, it is not all surprising that they culminate in dissimilar recommendations.

In our view, these inevitable disagreements do not imply that decision-aid is useless but simply that a single problem may have several valid responses. Given that two different decision-aid models cannot be implemented in the same decision process, the decision-maker must be conscious of the qualitative choices implied by the different models – often conveying the analysts' own ethical choices – before coming to personal conclusions on the choice to be made. In this domain, the many different approaches reflect in our view the complexity of the researcher's task much more than a scientific weakness.

References

Allais, M. (1953), "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'Ecole Américaine", *Econometrica* 21 (4) 503–546.

Bouyssou, D. (1984), "Decision-aid and expected utility theory: A critical survey", in: O. Hagen and F. Wenstøp (eds.), *Progress in Utility and Risk Theory*, Theory and Decision Library, Reidel, Dordrecht

Fishburn, P.C. (1970), *Utility Theory for Decision Making*, Wiley, New York.

Jacquet-Lagreze, E. and Roy, B. (1981), "Aide à la décision multicritère et systèmes relationnels de préférences", in: P. Batteau, E. Jacquet-Lagreze and B. Monjardet (eds.), *Analyses et Agrégation des Préférences*, Economia, Paris.

Keeney, R.L. (1974), "Multiplicative utility functions", *Operations Research*, 22, 22–34.

Keeney, R.L., Nair, K. (1976), "Evaluating potential nuclear power plant sites in the pacific northwest using decision analysis", IIASA Professional Paper no. 76–1; also in Bell, D.E., Keeney, R.L. and Raiffa, H. (eds.), (1977), Conflicting Objectives in Decisions, Chapter 14, Wiley, and in : Keeney, R.L. (1980), *Siting Energy Facilities*, Chapter 3, Academic Press, New York.

Keeney, R.L. and Raiffa, H. (1976), *Decision with Multiple Objectives – Preferences and Value Tradeoffs*, Wiley, New York.

Keeney, R.L., Robillard, G.A. (1977), "Assessing and evaluating environmental impacts at proposed nuclear power plant sites", *Journal of Environmental Economics and management* 4, pp. 153–166.

Luce, R.D. (1956), "Semiorders and a theory of utility discrimination", *Econometrica* 24, 178–191.

Morgenstern, O. (1979), "Some reflections on utility", in: Allais, M. and Hagen, O. (eds.), *Expected Utility Hypotheses and the Allais Paradox*, D. Reidel Publishing Company, dordrecht, 175–183.

Raiffa, H. (1968), *Decision Analysis*, Addison-Wesley, New York.

Raiffa, H. (1969), "Preferences for multiattributed alternatives", RM 5868, DOT/RC, The Rand Corporation, April.

Roy, B. (1971), "Problems and methods with multiple objectives functions", *Mathematical Programming* 1, 239–266.

Roy, B. (1977), "Partial preference analysis and decision aid: The fuzzy outranking relation concept", in D.E. Bell, R.L. Keeney and Raiffa H. (eds.): *Conflicting Objectives in Decisions*, Wiley, New York, 40–75.

Roy, B. (1978), "ELECTRE III: Un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples", *Cahiers du CERO* 20 (1) 3–24.

Roy, B. and Bouyssou, D. (1983), "Comparison, sur un cas précis, de deux modèles concurrents d'aide à la décision", Université de Paris-Dauphine, Document du LAMSADE no 22 (102 p.).

Roy, B., Present, M. and Silhol, D. (1983), "Programmation de la rénovation du métro parisien: Un cas d'application de la méthode ELECTRE III", Universitéde Paris-Dauphine, Document du LAMSADE no 24.

Roy, B. and Vincke Ph. (1984), "Relational systems of preference with one or several pseudo-criteria: Some new concepts and new results", *Management Science* 30 (11), 1323–1335.

Roy, B. (1985), *Méthodologie Multicritère d'Aide a la Décision*, Economica, Paris.

Von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior, 2nd ed.*, Princeton University Press, NJ.

Zeleny, M. (1981), *Multiple Criteria Decision Making*, McGraw-Hill, New York.