5 Assessing the quality of different MCDA methods

Andrea de Montis, Pasquale De Toro,
Bert Droste-Franke, Ines Omann, Sigrid Stagl

Introduction

The awareness that environmental problems have a large impact on the economy and society has risen continuously during the last three decade. Lifting sustainable development to the political agenda nurtured the understanding that policy analysis is characterized by conflicting economic, environmental, societal, technical, and aesthetic objectives. Consequently numerous research projects on the interdependence of environment, society and economy were conducted in this period. With increasing knowledge about the economic and environmental processes and their interlinkages, it became clear that decisions about environmental problems are especially characterized by uncertainties, high stakes, urgency, and dispute (Funtowicz and Ravetz 1990). These characteristics influence the choice of the method.

In the field of environmental decision-making, multiple criteria decision aid (MCDA) has become a widespread tool, able to tackle problems involving more than one objective. A number of theoretical (e.g. Munda 1995; Castells and Munda 1999) and empirical studies (e.g. Janssen 1992; Andreoli and Tellarini 2000; Mendoza and Prabhu 2000) showed that MCDA provides a useful tool for decision aid in this field, as it allows for multiple objectives, for the use of different types of data and the involvement of different stakeholders. The choice of a particular method depends on the specific problem as well as on users' demands. As the number of existing methods is already quite large and is still increasing, the choice of the 'right' method is not an easy task.

The aim of this chapter is the comparison of frequently used multiple criteria methods. A number of authors have compared multiple criteria algorithms with regard to their consistency, relative accuracy, speed, etc. (e.g. Olson 2001; Salminen *et al.* 1997). Since the underlying philosophies and assumptions, and hence the applicability, vary significantly, we chose, in contrast, a less mechanistic approach and describe the methods with reference to a comprehensive list of quality criteria. The criteria were chosen with particular emphasis on their relevance for questions of sustainable development. For the evaluation we apply the list of quality criteria to seven different MCDA methods (MAUT, AHP, Evamix, ELECTRE III, Regime, NAIADE, and programming methods) and compare them. This list of criteria provides a basis to assess the usefulness of each method for a specific application. The chapter is structured as follows. First, the list of quality criteria are presented, then the seven methods are characterized and evaluated by use of the list of criteria. This evaluation leads to a comparison of the methods, which is followed by conclusions.

Criteria for the quality assessment of methods

To ease the selection of the appropriate method for a specific decision-making situation, a list of quality criteria was developed in De Montis *et al.* (2000) which can be used to reveal strengths and weaknesses of MCDA methods with respect to their application for environmental problems. These MCDA quality criteria can be grouped according to three main aspects:

- operational components of MCDA methods,
- applicability of MCDA methods in the user context, and
- applicability of MCDA methods considering the problem structure.

Table 5.1 lists the criteria together with a summarizing description.

The first group of criteria includes operational components and characteristics of the method resulting from its theoretical underpinnings. The first important point for the method characterization is whether interdependence of criteria can be addressed in its application (e.g. equity or efficiency). Different technical procedures used to derive the decision may lead to different results in the same decision-making situation, because they use different mechanisms to compare criteria and alternatives. Some methods, for instance, do allow for incomparability, while others do not. Additionally, with regard to the characteristics of the criteria and the weighing process, the methods are distinct from each other.

Beside the theoretical foundations and the technical components of the method, specific properties concerning the applicability of the method with respect to the user's/users' situation are important factors for the method characterization. Quality criteria concerning the project constraints (costs and time), and on the other hand criteria addressing the envisaged type of the structure of the problem-solving process (e.g. the possibility of stakeholder participation) are relevant factors.

Another type of criteria which are relevant in the context of the methods' applicability, are criteria which give information about the possibility for consideration of important issues within the specific problem-solving situation. One group of these criteria aim to examine the possibilities for the inclusion of indicators with specific characteristics necessary for the problem solution, e.g. the evaluation of indicators over different geographical scales. Another group of criteria characterizes the data situation, i.e. the type of data which can be used and the situation of data availability for the indicators' evaluation in the context of the specific problem tackled.

According to the milestone textbooks by Roy (1985; 1996), multiple criteria methods can be divided into three operative approaches (Roy 1996: 241): (a)

Table 5.1 List of quality criteria for MCDA methods (De Montis *et al.* 2000)

Quality criterion	*Description*
Operational components	
Criteria	
Interdependence	Allowance for the interdependence of different criteria
Completeness	Need for the completeness of the criteria
Non-linear preferences	Possibility to express non-linear valuation patterns
Weights	
Transparency of process, type of weights	Type of the procedure of deriving values for the weights
Meaning	Interpretation and role of weights in the evaluation process
Solution finding procedure	Type of procedure used for the comparison of alternatives
Issues addressed by results	Interpretation of the results generated by the use of method
Applicability of MCDA methods – user context	
Project constraints	
Costs	Implementation costs in the specific user situation
Time	Implementation time in the specific user situation
Structure of problem solving process	
Stakeholder participation	Possibility to include more than one person as decision maker
Problems structuring	Existence of mechanisms supporting the structuring of the problem
Tool for learning	Support of learning processes
Transparency	Promotion of transparency in the decision making process
Actor communication	Support of the communication between opposing parties
Applicability of MCDA methods – problem structure	
Indicator characteristics	
Geographical scale	Applicability of different geographical scales for one case
Micro-macro link	Applicability of different institutional scales for one case
Societal/technical issues	Possibility for the consideration of both societal and technical issues
Methods combination	Possibility of methods combination
Data situation	
Type of data	Type of data supported as values for the indicators
Risk/uncertainties	Possibilities for the consideration of evaluation risk and/or uncertainties
Data processing amount	Processing amount needed to compile the data required for the method
Non-substitutability	Possibility to consider sustainability standards and non-substitutability

methods based on the use of a single synthesizing criterion without incomparability, (b) methods based on the synthesis by outranking with incomparability, and (c) methods based on interactive local judgements with trial-and-error iteration. While the first two groups embody a clear mathematical structure, the third does not use any formalized or automatic procedure. Another possible division refers to two groups of multiple criteria methods, namely discrete and continuous, depending on whether the set of alternatives is finite or infinite (Voogd 1983). In the following we will first present two types of discrete methods (single synthesizing criterion and outranking methods) and then the continuous methods (programming methods).

Single synthesizing criteria

MCDA methods using a single-criterion approach are trying to convert the impacts concerning the different criteria into one criterion or attribute, which build the base for the comparison of the alternatives. A well-known example for a single-criterion approach is cost-benefit analysis (CBA), which is using the Net Present Value as the criterion. This requires the conversion of impacts into monetary values. Hence, CBA cannot be seen as an MCDA method, because by use of MCDA methods each impact is evaluated on its most suitable scale; this scale can be both cardinal (monetary and non-monetary) and ordinal (qualitative). In the case of MCDA methods using a single synthesizing criterion, the aggregation of all the evaluations concerning each criterion results in a single index associated with each alternative and it expresses the performance of the alternative with respect to all the evaluation criteria simultaneously. Usually the higher this index the better ranks the alternative. Three such methods are described in the following: multiple attribute value theory (MAUT), analytic hierarchy process (AHP) and evaluation matrix (Evamix).

Multiple attribute value theory (MAUT)

The foundations of MAUT were laid by Churchman, Ackoff and Arnoff (1957) who first treated a multiple criteria decision problem using a simple additive weighting method. It was developed further using a quasi-additive and multi-linear utility function. Its early development is attributed to Fishburn (1968, 1970, 1978). Keeney and Raiffa devoted much of their work (e.g. 1976) to MAUT. Since then the method has been developed further in the way that methodologies to elicit the utility functions are elaborated and to aggregate the single-criterion utility functions to a multi-attribute utility function, and now provides a formal mechanism for handling multiple objectives, intangible factors, risk, qualitative data and time-sequence effects in ex-ante appraisals based on the decision-maker's preferences (Dillon and Perry 1977). A frequently used software package, which is based upon the estimation of an additive utility function and on interaction with the decision-maker is the so-called PREFCALC software. Other software packages are 'Decide Right', which is based on SMART, the Simple Multi-Attribute Rating Technique (for explanation, see later in this chapter) and 'HIPRE 3+', which integrates two methods of decision analysis and problem solving, namely AHP – Analytic Hierarchy Process (see **cross references needed pp.**) and SMART.

As the name MAUT indicates, it is based on utility theory (von Neumann and Morgenstern 1947). It relies upon the basic von Neumann–Morgenstern axioms of preference and thus upon a utility function, which allows the comparison of risky outcomes through the computation of expected utility. Risk can arise in two ways, as risk about the outcomes or about the attribute values. MAUT can be used for the first form of risk[1] (Hwang and Yoon 1981). MAUT uses directly assessed preferences with general aggregation; this involves direct questioning of the decision-makers and the choice on the basis of an aggregate measure for each alternative (Dillon and Perry 1977). It can be seen as an extension of ordinary subjective expected utility procedures to the case of choice between multi-attributed alternatives.

To prepare a multiple attribute decision by use of MAUT requires the following steps (Dillon and Perry 1977):

1 specify the project alternatives (including combinations) as discrete entities,
2 elicit the decision-maker's set of probability distributions for outcomes associated with each project alternative in each attribute if there is risk,
3 elicit the decision-maker's utility function $u_i(x_i)$ for the range of outcomes on each attribute,
4 use the appropriate global multi-attribute utility function $U(x)$ to find the expected utility of each project alternative, and
5 choose the project or project combination with the highest expected utility; thus the function U should be maximized.

Step 1 comprises the definition of the alternatives, the objectives, the attributes and the outcomes of each alternative in each attribute. This step is not specific to MAUT; this is how all discrete MCDA methods start.

Step 2 concerns risk viewed as subjective probability. If there is risk concerning the impact of the actions on the attributes, probabilities p_{ij} can be assigned to the outcomes $x_1, y_1, \ldots$ of the alternatives on the attributes. It is assumed that the value of the action a for attribute i is given by the expected value of u_i (Vincke 1985; Dillon and Perry 1977): $u_i(x_i) = \mathrm{S}_j\, p_{ij} u_i(x_{ij})$, where p_{ij} is the probability of the j-th outcome in the i-th attribute or in continuous terms, and $u_i(_{xi}) = \int u_i(x_i)\, f_i(x_i) dx_i$, where $f_i(x_i)$ is the probability distribution of outcomes in the i-th attribute.

There are different methods to elicit subjective probability distributions, which are not without difficulties. Which method to choose depends on the problem and often on the experience of the analysts. Direct involvement methods are simple and allow the decision-makers to see and understand the built distribution. For these reasons, such direct involvement methods seem preferable to others, such as lottery-type questions (Dillon and Perry 1977). The simplest method of all is asking for a few point estimates of probability; this is crude but practicable.

Step 3 requires the elicitation of the decision-maker's utility function $u_i(x_i)$ for the range of outcomes on each attribute. A utility function $u_i(x_i)$ is developed for each attribute. This is done by asking the decision-maker a series of questions based on the axiom of continuity (e.g. Dillon and Perry 1977: 10). Additionally weights w_i are assigned to the attributes, presenting trade-offs, by asking for the decision-maker's order of importance of the attributes. The assessment of the appropriate utility function is complex and intricate. Thus it is a crucial step in MAUT. This procedure leads us to three difficult questions (Vincke 1985): (a) what must the properties of the decision-maker's preferences be so that U is a certain function of u_i?; (b) how to test the properties?; and (c) how to construct the function U? The method can only be applied if these questions can be answered in a meaningful way.

After the establishment of the utility functions $u_i(x_i)$ and the elicitation of the attribute probability distributions, the expected utility is aggregated across the attribute distributions for each alternative in Step 4. For each attribute a function u_i is built and the functions u_i are aggregated in a global criterion U, such that $U(u_1(x_1),\ldots, u_n(x_n)) > U(u_1(y_1),\ldots,u_n(y_n))$, if the action represented by $(x_1,x_2,\ldots,x_n)$ is better than the action represented by $(y_1,y_2,\ldots,y_n)$, when considering all the attributes simultaneously (Vincke 1985; Roy and Bouyssou 1985). In that way the multiple criteria problem is reduced to a single-criterion decision problem. In order to build the aggregate function, different aggregation procedures are possible; the additive or the multiplicative aggregations are most widely applied. The additive form is the simplest form of a utility function, requiring two strong assumptions, namely utility independence and preferential independence of any subset of criteria (Fandel and Spronk 1985).[2] SMART (simple multiple attribute rating technique) presents one method to build the additive form (Martel, in Climaco 1997). If the additive model holds, the expected utility of each attribute for each alternative is added. Each attribute must get a weighting factor w_i to get a common scale for the utility functions $u_i(x_i)$. The utility for a certain alternative with uncertain consequences is then given by: $U(x) = \Sigma w_i\ [u_i(x_i)]$, where in discrete terms, $u_i(x_i) = \Sigma_j\ p_{ij} u_i(x_{ij})$, p_{ij} being the probability of the j-th outcome in the i-th attribute; or in continuous terms, $u_i(x_i) = \int u_i(x_i)\ f_i(x_i) dx_i$, where $f_i(x_i)$ is the probability distribution of outcomes in the i-th attribute. If independence does not hold, the multiplicative form is usually applied. Another form is the quasi-additive model, which can be used if neither marginality nor preferential independence prevails; but this procedure becomes very complicated for cases with more than three attributes (Martel, in Climaco 1997). Step 5 consists of summarizing the results and interpreting them.

A sub-methodology of MAUT is multiple attribute value theory (MAVT). The two approaches deal with risk in different ways. While MAUT relies upon a utility function, which allows the comparison of risky outcomes through the computation of an expected utility, MAVT is unable to take risk into account. It uses a value function to represent the outcome of the alternatives, not allowing for risky outcomes. Value functions preserve deterministic ordering, whereas utility functions preserve the expected utility hypothesis. MAUT is suited for ex-ante evaluations for multiple objective problems with risky outcomes as long as one accepts the assumptions on which the method is based.

MAVT is based on four key assumptions: First, a complete system of preferences pre-exists, which is an adequate representation of the axioms in the decision-maker's mind. Second, the preference relation of the decision-maker is of a weak order, which means that all states are comparable, transitivity of the preferences and transitivity of the indifferences is given. The last one is particularly difficult as it can lead to a lack of realism; see for example the famous cup-of-coffee example by Luce (1956).[3] Third, the main hypothesis underlying MAUT is that in any decision problem there exists a real-valued function U defined on A (the set of actions), which will be examined. This function aggregates the utility functions u_i (Vincke 1985). Fourth, in the probabilistic case, when there is risk concerning the outcomes, the axioms of von Neumann and Morgenstern (1947) must hold. In this way the concept of rationality comes into MAUT, which can be seen as an extension of ordinary subjective expected utility procedures to the case of choice between multi-attribute alternatives. They imply an existing unique set of subjective probability distributions for outcomes in each of the attribute dimensions associated with any risk multi-objective alternative and a utility function whose expected value gives a measure of the decision-maker's preference for each of the risky multi-attribute alternatives. If the axioms are accepted, there exists both 'a unique set of subjective probability distributions for outcomes in each of the attribute dimensions associated with any risky alternative and a utility function $U(x_i)$ whose expected value (in terms of the decision-maker's subjective probabilities) gives a measure of the decision-maker's preference for each of the risky multi-attribute alternatives' (Dillon and Perry 1977: 9).

The development of MAUT was a milestone in the history of MCDA, as it presented a formal and transparent method able to deal with multiple objectives, risk concerning the consequences, the pattern of consequences over time, and a start for using qualitative data. But MAUT's main problem is the strong assumptions that need to be fulfilled to derive the global utility function. Basically, MAUT relies upon the same axioms as social welfare theory like CBA does and therefore a number of the points of critique towards CBA also apply to MAUT. A review of case studies shows that MAUT is used frequently for decision-making for economic, financial, and actuarial problems as well as problems concerning environmental issues, like water management, energy management, agriculture, but not really in the field of sustainable development, which is much broader as it covers an economic, a social, and an environmental dimension.

Concerning the first group of quality criteria (see Tables 5.1 and 5.2) for MCDA methods, interdependence of attributes is not allowed; on the contrary the independence assumption is crucial. This assumption makes it hard for use in complex decision problems, as is the case for problems in the field of sustainable development. The meaning and assignment of weights is clear (weights stand for the importance of criteria) and transparent. The solution is a complete ranking of actions, according to their expected global utility. This again is clear and transparent, but the derivation of the utility function is the most delicate part of the whole procedure. Different issues concerning the type of decision problem are allowed, as long as they can be addressed by attributes; a utility function is designed

Table 5.2 Summary of method comparison

		MAUT	*AHP*	*Evamix*	*ELECTRE III*	*Regime*	*NAIADE*	*MOP/GP*
Operational components of methods								
Criteria	Inter-dependence	Not allowed	Allowed	Not allowed	Not allowed	Not allowed	Not allowed	Not allowed
	Completeness	Allowed	Allowed	Allowed	Allowed	Allowed	Allowed	Required
	Non-linear preferences	Allowed	Not allowed	Not allowed	Not allowed	Not allowed	Allowed	Allowed for some variants
Weights	Transparency of process, type of weights	Cardinal weights assigned	Cardinal weights assigned	Ordinal weights assigned	Cardinal weights assigned	Ordinal weights assigned	No weights assigned	Weights assigned for some variants
	Meaning	Trade-offs	Importance	Importance	Importance	Importance	No weights	Importance
Solution finding procedure		Complete ranking	Complete ranking	Complete ranking	Non-dominated option/s calculated	Complete ranking	Non-dominated option/s calculated	Non-dominated option/s calculated
Applicability in user context								
Project constraints	Costs	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.	Depending on of the numbers attributes, stakeholders involved, etc.
	Time	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributed, stakeholders involved, etc.	Depending on the numbers of attributes, stakeholders involved, etc.
Structure of problem solving process	Stakeholder participation	Necessary	Necessary	Supported	Supported	Supported	Necessary	Supported by a few variants

	Problems structuring	Via the construction of utility functions	Via the construction of suitable hierarchies	Via the construction of the evaluation matrix	Sometimes cumbersome	Vie the construction of the Regime vector	Via the construction of the evaluation matrix	Model based
	Tool for learning	Appropriate	Appropriate	Not appropriate			Appropriate	interactive MOP
	Transparency	High	High	High	Medium	Medium	High	Low
	Actors communi-cation	Via the integration of stakeholders	Via the integration of stakeholders	Via the integration of stakeholders	Via the integration of stakeholders	Via the integration of stakeholders	Via the integration of stakeholders	Via the integration of stakeholders
		MAUT	AHP	Evamix	Electre III	Regime	NAIADE	MOP/GP
Applicability for problem structure								
Indicator characteris-tics	Geographical scale	Different can be treated	Different can be treated	Different can be treated	Different can be treated	Different can be treated	Different can be treated	Different can be treated
	Micro-macro link	Possible	Possible	Possible	Possible	Possible	Possible	Possible
	Societal/ technical issues	Different issues are possible	Different issues are possible	Different issues are possible	Different issues are possible	Different issues are possible	Different issues are possible	Different issues are possible
	Methods combinations	Possible	Possible	Possible	Possible	Possible	Not possible	Possible
Data situation	Type of data	Qualitative and quantitative possible	Qualitative and quantitative possible	Qualitative and quantitative possible	Mainly quantitative	Qualitative and quantitative possible	Qualitative and quantitative possible	Mainly quantitative
	Risk/ uncertainties	Risky outcomes: probabilities, sensitivity analysis	Via sensitivity analysis	Via ordinal criteria only	Via thresholds		Via stochastic and fuzzy numbers	Via sensitivity analysis and fuzzy numbers
	Data processing amount	High	Medium	Low	High	Medium	Medium	Medium
	Non-substitutability	Not possible	Not possible	Not possible	Not possible	Not possible	Not possible	Possible via constraints or lexicographic ordering

for each attribute. Thus the more issues are addressed the more time it takes and the higher are the costs.

Stakeholder participation is necessary and desirable. Stakeholders are involved twice during the decision-making process. First, when developing the utility function for each attribute, they have to elicit their preferences. Second, when calculating the expected utility, they have to assign probabilities to certain outcomes. Preferences are directly assessed and generally aggregated and used to develop the utility functions. Under MAUT, transparency of the process is given as well as a clear structuring. Concerning the indicators and data, MAUT can handle different scales as well as different issues. The required amount of data is usually quite high. The data itself can be of qualitative or quantitative character. Risks of the outcome are accounted for by the expected utility theory, while uncertainties are checked with a sensitivity analysis.

In sum, MAUT ensures in risky project choices the correspondence of the choice with the decision-maker's preferences and it is a mechanism for doing this in cases too complex to be handled satisfactorily by intuition, as long as the underlying assumptions are known and always in the minds of the analysts. Information about the preferences of the decision-maker are necessary to build the singular utility functions for each attribute and to build the aggregate utility function based on all attributes. These individual utility functions build the base for the decision. No social function is constructed, thus no interlinkages or interdependencies between the individuals are taken into account. The building of a group or social function is a problem in every MCDA method, as we learned from Arrow's impossibility theorem that it is not possible to aggregate individual preference functions to group opinions. The individual preferences and estimations can be seen as an acceptable proxy as long as the decision projects stay on a rather micro-related level, like the decision about an irrigation system. But if we have to deal with macro-related problems, additional influences gain importance, such as the impact on the different dimensions of sustainable development or on different groups in society. This decreases the appropriateness of individual utility functions as an acceptable proxy for social preferences. Thus MAUT, like CBA, is hardly suited for macro-related problems.

Analytic hierarchy process (AHP)

AHP was developed by Thomas Lorie Saaty. The method is implemented by the software package 'Expert Choice'. From a procedural point of view this approach consists of three steps: (1) construct suitable hierarchies; (2) establish priorities between elements of the hierarchies by means of pairwise comparisons; (3) check logical consistency of pairwise comparisons (Saaty 1980, 1988; Saaty and Alexander 1989; Saaty and Forman 1993).

Step 1 concerns the construction of suitable hierarchies. This step is based on findings indicating that when elaborating information, the human mind recognizes objects and concepts, and identifies relations existing between them. Because the human mind is not able to perceive simultaneously all factors affected by an action and their connections, it helps to break down complex systems into simple structures: this simplification is possible by means of a logical process which aims at the construction of suitable hierarchies. Hence, hierarchies can be useful in helping the human mind to make decisions by constructing a framework of different and numerous elements, which are separated and connected at the same time. 'A hierarchy is a particular type of system, which is based on the assumption that the entities, which we have identified, can be grouped into disjoint sets, with the entities of one group influencing the entities of only one other group, and being influenced by the entities of only one group' (Saaty 1980: 7).

The simplest model of hierarchy consists of three steps. The first step coincides with the main objective (called the 'goal') of the decision-making problem; the second and third steps include criteria and alternatives. It is, however, possible to develop more complex hierarchies (i.e. with more levels), which include a certain number of sub-criteria. This means that factors affecting the decision are organized in gradual steps from the general, in the upper level of the hierarchy, to the particular, in the lower levels. With reference to the same decision-making problem it is also possible to construct more than one hierarchy, for example, a first hierarchy for benefits, a second one for costs, and a third one for risks. In order to obtain satisfactory results, hierarchies should be large enough to capture all the major factors relevant for the decision-making problem but small enough to remain sensitive to changing crucial factors.

Step 2 establishes priorities between elements of the hierarchies by means of pairwise comparisons (i.e. comparing elements in pairs with respect to a given criterion). In the AHP approach, pairwise comparisons are used for establishing priorities or weights among elements of the same hierarchical level. They are compared in pairs with respect to the corresponding elements in the next higher level, obtaining a matrix of pairwise comparisons. For representing the relative importance of one element over another, a suitable evaluation scale is introduced (Saaty 1988, 1992), called 'Saaty's scale'. It defines and explains the values 1 to 9 assigned to judgements in comparing pairs of elements in each level with respect to a criterion in the next higher level. In particular, the following values which express the intensity of importance between the elements are used:

1 if two elements have 'equal' importance;
3 if 'moderate' importance of one element over another is recognized;
5 if 'strong' importance of one element over another is recognized;
7 if 'very strong' importance of one element over another is recognized;
9 if 'extreme' importance of one element over another is recognized.

The numbers 2, 4, 6, 8 can be used to express intermediate values between two adjacent judgements (i.e. between 'equal' and 'moderate', 'moderate' and 'strong', etc.). When element i compared with j is assigned one of the above numbers, then element j compared with i is assigned its reciprocal.

In particular, for each criterion C, an n-by-n matrix **A** of pairwise comparisons is constructed. The components a_{ij} $(i, j = 1, 2, \ldots, n)$ of the matrix **A** are numerical entries, which express (by means of Saaty's scale) the relative importance of the element *i* over the element *j* with respect to the corresponding element in the next higher level. Thus the matrix **A** has the form:

$$A \equiv \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

where:

$$a_{ii} = 1\ a_{ji} = \frac{1}{a_{ij}} \qquad a_{ij} \neq 0$$

In order to calculate relative priorities among the *n* elements considered, the 'principal eigenvector' (**v**, with usually Sv_i^{1} 1) of the matrix **A** is computed. Then this eigenvector is normalized obtaining the 'priority vector' (**x**, with $\Sigma x_i = 1$), which expresses the priorities among the elements belonging to the same node of the hierarchy. Each component of the vector **x** represents the 'local priority' of an element (i.e. a node of the hierarchy) of the pairwise comparisons; the 'global priority' of that element is the product of its local priority with the global priority of the upper node. Of course, the local priority of nodes at the first and second levels is equal to their global priority. Thus each element at each level has a weight (i.e. its global priority) assigned to it. The composite weight for each of the elements at the final level (i.e. the alternatives) is obtained by multiplying the weights along each path of the hierarchy from the apex to the final element and adding the resultant weights from all paths to the element. The result is a vector of final weights for the alternatives under consideration: the higher its weight the better the alternative is. Moreover, it can be examined how priorities among alternatives change according to the variation of weights assigned to the upper elements ('sensitivity analysis').

ks for the logical consistency of pairwise comparisons. In comparing elements, inconsistency of a certain degree can arise: in the AHP approach the 'maximum or principal eigenvalue' (called λ_{max}) of each matrix of pairwise comparisons is computed for checking the degree of inconsistency.[4] In particular the consistency of a positive reciprocal *n*-by-*n* matrix is equivalent to the requirement that its maximum eigenvalue λ_{max} should be equal to *n*; if the matrix is inconsistent $\lambda_{max} > n$ it is possible to estimate the departure from consistency by the difference ($\lambda_{max} - 1$) divided by ($n-1$) (Saaty and Vargas 1991; Saaty 1994). The closer λ_{max} is to *n* the more consistent is the matrix **A** and its deviation from consistency may be represented by the so-called 'consistency index' (*C.I.*):

$$C.I. = \frac{\lambda_{max} - n}{n-1}$$

Saaty proposes a 'consistency ratio' in order to appreciate the consistency of the matrix which is calculated by dividing the *C.I.* by *R.I.*, where *R.I.* represents an experimental 'random index' which increases as the order of the matrix increases.[5] A consistency ratio of 0.10 or less is considered acceptable; if this ratio is more than 0.10, it is necessary to reformulate the judgements by means of new pairwise comparisons.

The AHP method can take into account the interdependence existing between the different evaluation criteria by means of the construction of suitable hierarchies. The elements of a hierarchy (i.e. criteria, sub-criteria, alternatives) are not linked to all the others but they are grouped into disjoint sets; if an interdependence between some elements is recognized, they are placed in the same hierarchical level and connected to the same element of the next higher level. If the choice of the number of hierarchical levels is unlimited (although, as a rule of thumb, no more than seven elements for each node is best) it is possible to insert in the hierarchy all the criteria required by the decision-making problem. In this sense a complete (i.e. exhaustive) list of evaluation criteria is allowed. A cardinal weight is assigned to each element of the hierarchy by means of opportune pairwise comparisons. The weights express the importance of one element over the others. The final result offered by the method is a complete ranking of the evaluated alternatives.

AHP facilitates stakeholder participation. Different social actors are invited to assign weights to the evaluation criteria; in this perspective weights reflect different social views. In comparing elements in pairs the different judgements given by each actor can also be integrated – by means of an arithmetic average – obtaining only one weight for each criterion, which expresses synthetically the points of view of all the involved stakeholders. It is also possible to promote the actors' communication if they are called to work together in order to identify general objectives, evaluation criteria, sub-criteria, and deduce their importance.

The method has been applied to a range of decision-making problems (e.g. corporate policy and strategy, public policy, political strategy, environmental planning) at different geographical scales. Sometimes also using other evaluation methods it is possible to make a pairwise comparison between criteria as carried out in the AHP if normalized weights are required. In this sense a combination with other methods is possible. AHP allows use of qualitative and quantitative data; in particular, cardinal numbers can be directly normalized without the need of pairwise comparisons.

Evaluation matrix (Evamix)

Evamix was developed by Henk Voogd. The evaluation matrix in this method may include mixed (i.e. qualitative and quantitative) data. Both the method and the software are called Evamix. From a procedural point of view this approach consists of five steps: (1) make a distinction between ordinal and cardinal criteria; (2) calculate dominance scores for all ordinal and cardinal criteria; (3) calculate standardized dominance scores for all ordinal and cardinal criteria; (4) calculate overall dominance scores; and (5) calculate appraisal scores (Voogd 1981, 1983).

Step 1 makes a distinction between ordinal and cardinal criteria. The first step is the construction of an evaluation matrix **E**, which is an m-by-n matrix characterized by m evaluation criteria and n alternatives. Its components are qualitative or quantitative entries, which express by rows the performance of each alternative with respect to a certain criterion. Given a set of evaluation criteria j ($j = 1, 2, \ldots, m$) and a finite set of alternatives i ($i = 1, 2, \ldots, n$), the evaluation matrix **E** will be characterized by its qualitative and quantitative[6] components e_{ji}:

$$E \equiv \begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{pmatrix}$$

The set of criteria j is divided into two subsets, denoted O and C, where O is the set of the ordinal (qualitative) criteria and C the set of cardinal (quantitative) criteria, obtaining two distinct evaluation matrices: $\mathbf{E_O}$ (ordinal criteria/alternatives) and $\mathbf{E_C}$ (cardinal criteria/alternatives). This way the differences among alternatives can be expressed by means of two dominance measures: the first one based on ordinal criteria and the second one based on cardinal criteria. In particular, in order to construct the cardinal dominance score, the components $e_{ji} \in \mathbf{E_C}$ are standardized to a common unit; in this way all the quantitative criteria are expressed on a scale from 0 to 1. All the standardized scores should have the same direction; this means that a higher score implies a better score. In the case of criteria for which lower scores means better scores, the standardized scores are transformed by subtracting them from 1.

Additionally, a qualitative weight can be assigned to each criterion constructing a vector ω_j ($j = 1, 2, \ldots, m$); the associated cardinal weights w_j can be obtained – in an approximated way – by looking at the 'extreme weights sets' which delimit the values that metric weights may have.[7] Because weights represent subjective value judgements, it is preferable to use different priority sets so that the consequences of different political viewpoints can be illustrated.

Step 2 requires calculation of dominance scores for all ordinal and cardinal criteria. Differences between alternatives are expressed by means of two dominance scores, for ordinal ($\alpha_{ii\ell}$) and cardinal ($a_{ii\ell}$) criteria:

$$\alpha_{ii'} = f\left(e_{ji}, e_{ji'}, \omega_j\right) \quad \left(\forall j \in O\right)$$
$$a_{ii'} = g\left(e_{ji}, e_{ji'}, \omega_j\right) \quad \left(\forall j \in C\right)$$

where e_{ji} ($e_{ji'}$) represents the score of criterion j and alternative i (i'), ω_j represents the weight attached to criterion j, and f and g are two different functions. If each extreme (quantitative) weight set w_ℓ is substituted into the above formulae instead of the qualitative weights ω_j, it is possible to define – for each extreme weight set – a new dominance measure for ordinal ($\alpha_{ii'\ell}$) and cardinal ($a_{ii'\ell}$) criteria:

$$\alpha_{ii'} = f\left(e_{ji}, e_{ji'}, w_j\right) \quad \left(\forall j \in O\right)$$
$$a_{ii'} = g\left(e_{ji}, e_{ji'}, w_j\right) \quad \left(\forall j \in C\right)$$

The scores expressed by means of these formulae reflect the degree in which alternative i dominates alternative i' for ordinal and cardinal criteria respectively.

Step 3 calculates standardized dominance scores for all ordinal and cardinal criteria. The dominance scores $\alpha_{ii'\ell}$ and $a_{ii'\ell}$ are standardized into the same measurement unit in order to make them comparable. If h is a standardization function, the 'standardized dominance measures' for all ordinal $(d_{ii'})$ and cardinal $(d_{ii'})$ criteria have the following expression:

$$\delta_{ii'\ell} = h(\alpha_{ii'\ell})$$
$$d_{ii'\ell} = h(a_{ii'\ell})$$

The standardization of the scores $\alpha_{ii'\ell}$ and $a_{ii'\ell}$ can be done in three different ways: the 'subtractive summation technique', the 'subtracted shifted interval technique' and the 'additive interval technique'.[8] Also in this case the standardized dominance scores reflect the degree in which alternative i dominates alternative i' for ordinal and cardinal criteria. Different scores are obtained using the three different techniques.

Step 4 calculates overall dominance scores. The 'overall dominance measure' $m_{ii'\ell}$ for each pair of alternatives $(i,\ i')$ – giving the degree in which solution i dominates solution i' – is calculated by means of the following formula:

$$m_{ii'\ell} = w_{O\ell}\delta_{ii'\ell} + w_{C\ell}d_{ii'\ell}$$

where $w_{O\ell}$ represents the weight of the qualitative criterion set O and $w_{C\ell}$ represents the weight of the quantitative criterion set C:

$$w_{O\ell} = \sum_{j \in O} w_{j\ell} \qquad\qquad w_{C\ell} = \sum_{j \in C} w_{j\ell}$$

Step 5 calculates appraisal scores. The overall dominance measure $m_{ii'\ell}$ may also be considered as a function k of the 'appraisal scores' s_i and $s_{i'}$ of the alternatives i and i':

$$m_{ii'\ell} = k(s_{i\ell}, s_{i'\ell})$$

At the same time this means that the appraisal score of an alternative depends on the overall dominance score such as calculated in Step 4. In particular, three different appraisal scores of alternative i with respect to the n other alternatives are obtained if the subtractive summation technique, subtracted shifted interval technique or additive interval technique has been used in Step 3. Additionally, by using the appraisal score of all three techniques ($t = 1, 2, 3$) a standardized 'average appraisal score' $a_{i\ell}$ for alternative i can be calculated:

$$a_{il} = \sum_{t=1}^{3} \left(\frac{s_{ti\ell} - s_{ti\ell}^{-}}{s_{ti\ell}^{+} - s_{ti\ell}^{-}} \right)$$

where $s_{ti\ell}$ is the appraisal score of alternative i for each technique; $s_{ti\ell}^{-}$ is the lowest $s_{ti\ell}$ -score for each technique; $s_{ti\ell}^{+}$ is the highest $s_{ti\ell}$ -score for each technique. The result is a complete ranking of alternatives and the higher $a_{i\ell}$ is the better is alternative i.

Since the starting point of Evamix is the construction of an evaluation matrix, it is not possible with this method to take the interdependence between the different evaluation criteria into account. But if we consider that there is no limit on the choice of the number of criteria to insert in the evaluation matrix, it is possible to consider all the criteria required by the decision-making problem. In this sense a complete (i.e. exhaustive) list of evaluation criteria is allowed. An ordinal weight, expressing the importance between all the relevant criteria, is assigned to each criterion, and subsequently ordinal weights are transformed into cardinal ones.

Evamix supports stakeholder participation. Different social actors are invited to assign weights to the evaluation criteria, and in this case the different weights given by each actor are not integrated but they are used to show the different points of view (i.e. economic, social, environmental) expressed by the stakeholders involved. It is also possible to promote actor communication if they are called to work together in order to identify general objectives, points of view, evaluation criteria, sub-criteria, and deduce their importance. The method is used to cope with different decision-making problems at different geographical scales. Most frequently, Evamix been used in urban and regional planning. Evamix allows using qualitative and quantitative data. In particular, quantitative data are expressed in an ordinal scale and the cardinal numbers are standardized in a scale from 0 to 1.

Outranking methods

The development of the outranking methods was started in France in the late 1960s by Bernard Roy and his team (Roy 1985). Outranking methods were developed in an attempt to manage with less strong assumptions (about existence of utility function, additivity, etc.) and to require less information from decision-makers (especially preference intensities, rates of substitution) than the methods described above.

The outranking methods also differ fundamentally from other methods in the approach insofar as they start from the assumption that the decision is a process during which decision-makers may change their preferences in response to the information provided and/or after thorough reflection about the problem. For this reason the developers of outranking methods considered it important to allow at the beginning for incomparability of options. This avoids a complete ranking being identified too early; the aim is therefore to stimulate thorough consideration and to avoid premature conclusions. Hence, more than other methods, these methods encourage interaction between the model and the decision-maker/s. The outranking algorithms also account for the political reality that options which perform very poorly in one dimension are likely to face severe opposition from at least some of the stakeholders and are therefore deemed unacceptable.

ELECTRE III

Roy (1985) developed one of the most widely used software package, known as ELECTRE, which is available in four different main releases (ELECTRE I, II, III, IV) and two other versions (ELECTRE IS and TRI). A comprehensive comparison of these approaches can be found in Vincke (1992). The underlying principle of the ELECTRE approach stems from processing a system of pairwise comparisons of the alternatives. This procedure involves the use of a system of preference relations (SPR) that is based on the exploitation of the outranking relation. According to this relation, an action will outrank another if the first action is considered to be at least as good as the second. Roy (1996) defines a system like this as a basic system of outranking relations. Recently it was pointed out that release III is remarkable among the ELECTRE family, since it takes explicitly into account the indifference and preference thresholds but, furthermore, it is based on an outranking relation, which is less sensible to the variation of the input data (Scarelli 1997). The system solves ranking problems, dividing the set A of actions into equivalence classes (Scarelli 1997). In release III of ELECTRE, the process can be sketched in the following steps.

A cardinal impact table is prepared for a finite number of alternatives 'measured' according to a set of pseudo-criteria. The latter are criteria whose SPR includes weak preference, indifference and incomparability, and a coherent system of thresholds has to be set. The program allows to set thresholds for indifference and for preference, according to a linear relationship: $T = \alpha \bullet g(a) + \beta$, where T is the total threshold, α is a proportional threshold, β is a fixed threshold and g is the criterion function. This threshold is defined directly when it refers to the maximum value of the criterion, and indirectly when it refers to the minimum value. This allows constructing fuzzy outranking relations, given that it is possible to state how much an action outranks another. In fact, a series of pairwise comparisons is set through the use of concordance and discordance indices. A general framework for the calculation of these indices is given by Voogd (1983). The degree of dominance of alternative i over alternative $i¢$ is measured by the following index:

$$c_{ii'} = \left[\frac{\sum_{j \in \psi} w_j^{\alpha}}{\sum_j w_j^{a}} \right]^{1/\alpha}$$

where ψ is the set of criteria that indicate a preference for the alternative a with respect to the alternative a', w the criterion weight and α a scaling parameter greater than one. 'By means of this scaling parameter the analyst is able to vary the importance of the small weights and small divergences between the scores' (Voogd 1983: 123). In many applications, the value of the scaling parameter is set equal to one, which allows managing a simpler functional algorithm.

The degree of discordance between the same alternatives is measured by the following index:

$$d_{ii'} = \left[\frac{\sum_{j \in \varphi} \left(w_j \mid e_{ji} - e_{ji'} \mid \right)^{\alpha}}{\sum_{j} \left(w_j \mid e_{ji} - e_{ji'} \mid \right)^{\alpha}} \right]^{1/\alpha}$$

where φ is the set of criteria that do not agree about the preference of alternative i with respect to alternative i'.

In the generalized concordance analysis the values of the concordance and discordance indices individuate the subsets of alternatives that are accepted and rejected. The first group consists of the alternatives that fulfil the following requirements: $T_c e \geq \eta$ and $T_d \geq \mu$.

In ELECTRE III, the concordance and discordance indices are calculated with reference to the values that the criterion functions take for the consequences of each alternative. In the same pattern, the threshold functions are modified as linear functions of the pseudo-criterion g. For each criterion g_j a preference threshold p_j and an indifference threshold q_j are set.

Given the pseudo-criterion g and two variable alternatives i and i', two concordance indices $c_j(i, i')$ are calculated for the comparison of alternative i with respect to alternative i' and vice versa. For the comparison between alternative i and i', the concordance index takes values between 0 and 1.

The discordance index $d_j(i, i')$ is calculated with reference also to the veto threshold v_j. This threshold is associated with a criterion that is considered to be very important. If the difference between two functional values in the case of a number of alternatives is greater than the threshold value, this condition acts as a veto to the affirmation: the first alternative outranks the second.

The construction of the outranking relation is based upon the credibility index $\sigma(i, i')$, which ranges between 0 and 1 and is a function of the concordance and of the discordance index. The credibility index takes the following values:

$$\sigma (i, i') = c (i, i'), \text{ if } d_j (i, i') < c_j (i,);$$

otherwise

$$\sigma(i,i') = c_j(i,i') \prod \frac{1 - d_j(i,i')}{1 - c_j(i,i')}$$

The credibility index provides the information for the construction of two rankings of the alternatives. First, from the worst to the best (ascending distillation) and the second from the best to the worst (descending distillation). The intersection of these distillations yields the final distillation graph. This graph indicates the system of relationships among the alternatives, according to a system of oriented arcs. This framework reveals incomparability relationships that the other distillations do not consider.

Roy (1996) considers that in every decision-making situation, four problem areas have to be taken into account: P_α (choice), P_β (sort), P_γ(rank), and P_δ (description). With reference to this outline set forth by Roy (1996), the procedure does not deliver a complete ranking of the evaluated alternatives, leaving partially unresolved what Roy (1985) calls the problematic *P.* γ (ranking of the alternatives). On the other hand, the procedure of ELECTRE III is helpful for what Roy (1985) names the problematic *P.* α (choosing the best action).

ELECTRE III is based on the assumption that the system of individual preference relationships can be described by means of a sophisticated analysis of the relationships of preference, indifference, and incomparability. As a matter of fact, incomparability enters as a puzzling factor into the final ranking. On the one hand, this complex analysis involving the assessment of several thresholds and preference relationships may be judged as a remarkable effort. On the other hand, it may lead to cumbersome work. Often it is very difficult to express credible threshold values and the resulting ranking can end up being hardly understandable. As a matter of fact, in the situation where results were described by a Kernel graph, the complex meaning stemming from this kind of representation would be hardly transmissible to a group of laypersons, such as a local community. Local communities seems to be more keen to comprehend and discuss definitive positions and comparable outcomes than to give personal interpretation of relationships. The nature of the ranking, which is presented by a graph, allows also for incomparability among the alternatives: such an incompletely resolved outcome challenges the analyst who should then, if equipped with abilities of good advocacy, foster communication among stakeholders to discuss the incomparability. Politicians often are unwilling to accept the complexity of this structure and show a tendency to sceptical behaviour against an incomprehensible technique for decision-making. Often it is the expectation that the multiple criteria analysis will solve the problem at hand completely. The obvious fact that MCDA supports decision-making, but must not substitute decision-makers, becomes even more apparent with outranking methods than with the previously discussed ones. These are the main reasons why in Table 5.2 ELECTRE is rated highly with respect to criteria completeness and interdependence, while it is rated poorly with respect to transparency and to applicability of criteria. It should be noted that in terms of 'meaning of weights' ELECTRE received a zero score, since it does imply weighting procedures connected to the combinatory use of the so-called 'weights' and to the thresholds.

Regime

The Regime method was developed by Hinloopen *et al.* (1983), assessed and refined by Hinloopen (1985) and Hinloopen and Nijkamp (1990). This method can be considered as a multiple criteria evaluation framework belonging to the family of dominance analysis approaches, even though it has been pointed out that it is quite distinct from other methods in this family, mainly because Regime is a qualitative multiple criteria method. Qualitative multiple criteria methods aim to provide a tool for analysing complex situations, which are impossible to model by means of quantitative information. Hinloopen and Nijkamp (1990) list several such methods: (a) the extreme expected value method (Kmietowicz and Pearman 1981; Rietveld 1980), (b) the permutation method (Mastenbroek and Paelink 1977), (c) the frequency method (Van Delft and Nijkamp 1977), (d) the multidimensional scaling method (Nijkamp and Voodg 1981), and (e) the mixed data method (Voogd 1983).

Since human activities and settlements are characterized by phenomena that are often impossible to measure, studies have focused on the construction of techniques able to deal with 'soft' and 'fuzzy' information. 'These qualitative multi-dimensional judgement methods are essentially more in agreement with "satisfying behaviour" (based on Simon's bounded rationality principle) than with conventional "optimising decision making", as imprecise information precludes the attainment of an unambiguous maximum solution' (Hinloopen and Nijkamp 1990: 38).

As in Evamix, an evaluation table is given and composed by e_{ij} scores of a number n of alternative scenarios with respect to m criteria. In the case of ordinal information the weight can be represented by means of rank orders w_j in a weight vector w: $w = (w_1, w_2, \ldots, w_j)^T$. The higher the value of the weight, the better the correspondent criterion.

The following main features characterize the multiple choice regime method. First, with this method the evaluation table can contain cardinal as well as ordinal data; this is accomplished by treating cardinal information as ordinal, with reference to the ranking position of each alternative. Second, the basis of the method is the regime vector. For two alternative choice options i and i', the difference of the criterion scores is assessed by considering the difference between the scores $s_{ii'j} = e_{ij} - e_{i'j}$.

In case of ordinal information, the order of magnitude of $s_{ii'j}$ is not relevant, but only its sign $\theta_{ii'j}$. Therefore, if $\theta_{ii'j} = +$, then alternative i is better than alternative i' for criterion j; i, $\theta_{ii'j} = -$, then alternative i is better than alternative i' for criterion j. The extension of the pairwise comparison to all the alternatives leads to the following $\vartheta \times 1$ regime vector: $r_{ii'} = (\sigma_{ii'1}, \sigma_{ii'2}, \sigma_{ii'3}, \ldots \sigma_{ii'n})$. The regime vector, composed of + or – signs, or eventually zeros, 'reflects a certain degree of (pairwise) dominance of choice option i with respect to i' for the unweighted effects for all ϑ judgement criteria' (Hinloopen and Nijkamp 1990: 41). For all the combinations of comparisons, a regime matrix is drawn with the following pattern:

$$R = \lfloor r_{12} r_{13} \ldots r_{I1} r_{21} \ldots\ldots\ldots r_{Ia} \ldots\ldots\ldots r_{I(I-1)} \rfloor$$

Usually, the regime vector is not completely composed of a series of '+' or a series of '–'. Instead it is often composed of both signs. Additional information has to be processed by means of ordinal values, on the assumption that the ordinal nature of the weights reflects the inability to measure human preferences, which in principle are cardinal. This weight vector can be considered as a 'rank order representation of an (unknown) underlying cardinal stochastic weight vector' (Hinloopen and Nijkamp 1990: 41): $w^* = (w_1^*, w_2^*, \ldots w_j^*)^T$ with $0 < w_j^* < 1$. The consistency hypothesis implies that the weight vector is supposed to be consistent with the quantitative information incorporated in an unknown cardinal vector.

The next assumption is that the dominance relation can be represented by means of the following stochastic expression referred to the following weighted linear summation of cardinal entities:

$$\nu_{ii'} = \sum_j \sigma_{ii'j} w_j^*$$

If $\nu_{ii'}$ is positive, then a dominance relation of choice option i over i' is revealed. Since the information about w_j^* is unknown, the probability of the dominance of i with respect to i' is introduced in the following pattern: $p_{ii'} = \text{prob}\,(n_{ii'} > 0)$. Thus an aggregate probability measure (or success score) can be defined as follows:

$$p_i = \frac{1}{I-1} \sum_{i' \neq i} p_{ii'}$$

Hence, p_i is the average probability that alternative i is given a higher value than another alternative. The rank order of choice options is then determined by the rank orders (or the order of magnitude) of the p_i. Since the assessment of the $p_{ii¢}$ is the crucial point, a uniform density function is usually assumed for the probability distribution of the w_j^*. This argument is referred to as the 'principle of insufficient reason', and as the Laplace criterion, in case of decision-making under uncertainties (Taha 1976). According to this statement, without any prior information, 'there is no reason to assume that a certain numerical value of w^* has a higher probability than any other value' (Hinloopen and Nijkamp 1990: 42).

Regime differs in some important points from other outranking methods. While ELECTRE requires cardinal rankings, the Regime method allows using mixed data so that cardinal and ordinal criterion can be included. The Regime method adopts ordinal weights; this feature seems to depict real systems of preferences better than other multiple criteria weighting methods. This approach is based on more prudent assumptions as it refrains from the difficult task of attaching cardinal values to measure the intensity of criteria importance.

In terms of shortcomings and benefits, this method is in a sense the counterpart of ELECTRE III. It has a simpler structure of the preference model than ELECTRE III, since qualitative and quantitative data can be used for criteria. In other words, this method is able to process mixed data, because it can adopt an ordinal scale for the impact scores. The result of the evaluation process with Regime is a complete ranking of the alternatives. The outcome of the analysis is quite easy to communicate and discuss. However, the Regime method does not admit incomparabilities among alternatives. Also it is not clear how the system operates in the stochastic domain. This is mainly due to the random generation system adopted for the weight vector: the analyst, involved in the task of explaining the exact mechanism of the algorithm, may meet significant difficulties in communicating the meaning and in involving the decisional community. Hence, despite the apparently simple structure of the Regime method, the resulting decision-making process may not be so transparent.

Novel approach to imprecise assessment and decision environments (NAIADE)

a MCDA method developed by Giuseppe Munda (1995), whose impact or evaluation matrix may include quantitative and qualitative data; in particular the values assigned to the criteria for each alternative may be expressed in the form of either crisp, stochastic, fuzzy numbers or linguistic expressions. Hence it allows the use of information affected by different types of uncertainty. It is a discrete method, and no weighting of criteria is used explicitly. The method is implemented by a software application called NAIADE.

From a procedural point of view the method consists of three steps: (1) the pairwise comparison of alternatives; (2) the aggregation of all criteria; (3) the evaluation of alternatives (Munda *et al.* 1994; Munda 1995; Menegolo and Guimarães Pereira 1996).

Step 1 requires pairwise comparison of alternatives. The first step is the construction of an evaluation matrix **E**, which is an m-by-n matrix characterized by m evaluation criteria and n alternatives. Its components are qualitative or quantitative entries, which express by rows the performance of each alternative with respect to a certain criterion. Given a set of evaluation criteria j ($j = 1, 2, \ldots, m$) and a finite set of alternatives i ($i = 1, 2, \ldots, n$), the evaluation matrix **E** will be characterized by its qualitative and quantitative[9] components e_{ji}:

$$E \equiv \begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{pmatrix}$$

and the pairwise comparison of alternatives (with reference to each criterion) is carried out by means of the concept of 'semantic distance' between two fuzzy sets.[10] More precisely, if S_1 and S_2 are two fuzzy sets and μ_1 and μ_2 are respectively their membership functions, it is possible to define two new functions:

$$f(x) = k_1\mu_1(x) \qquad \text{and} \qquad g(y) = k_2\mu_2(y)$$

obtained by rescaling the ordinates of μ_1 and μ_2 through k_1 and k_2 such as:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} g(y)dy = 1$$

The semantic distance between the two fuzzy sets (that is the distance between all points of their membership functions) is computed. The same distance concept can be extended to stochastic measures considering $f(x)$ and $g(y)$ probability density functions of the measures.

The pairwise comparison of alternatives is based on six 'preference relations':

1 much greater than (>>)
2 greater than (>)
3 approximately equal to (≅)
4 very equal to (=)
5 less than (<)
6 much less than (<<)

expressed for each criterion starting from the distance between alternatives. The above preference relations are analytically defined by means of six functions that express (for each criterion on the base of the distance between alternatives) an index of credibility of the statements that an alternative is 'much greater', 'greater', approximately equal', 'very equal', 'less' or 'much less' than another.[11] This credibility index goes, increasing monotonically, from 0 (definitely non-credible) to 1 (definitely credible). In particular, given a criterion j and a pair of alternatives i and i', it is possible to define six membership functions (different for crisp criterion scores and for stochastic or the fuzzy criterion scores), which are respectively named[12]:

$\mu_{>>}(i,i')_j$ (for much greater than)
$\mu_{>}(i,i')_j$ (for greater than)
$\mu \cong (i,i')_j$ (for approximately equal to)
$\mu_{=}(i,i')_j$ (for very equal to)
$\mu_{<}(i,i')_j$ (for less than)
$\mu_{<<}(i,i')_j$ (for much less than)

Membership functions related to stochastic or fuzzy numbers consider in their mathematical expressions the semantic distance defined above.

Step 2 requires aggregation of all criteria. In order to take into account all criteria simultaneously, it is necessary to aggregate the evaluations related to the pairwise performance of alternatives according to each single criterion. A 'preference intensity index' of one alternative with respect to another is calculated:[13]

$$\mu_*(i,i') = f(\mu_*(i,i')_j)$$

and then an aggregate 'fuzzy preference relation' can be obtained:

$$\begin{cases} \mu_{>>}(i,i') & H(>>) \\ \mu_{>}(i,i') & H(>) \\ \mu_{\cong}(i,i') & H(\cong) \\ \mu_{=}(i,i') & H(=) \\ \mu_{<}(i,i') & H(<) \\ \mu_{<<}(i,i') & H(<<) \end{cases}$$

where $\mu_*(i,i')$ is the overall evaluation of a given fuzzy relation for each pair of actions and $H(*)$ is the associated entropy level.[14]

Step 3 is the evaluation of alternatives. The evaluation of the alternatives derives from the information provided by the above aggregate fuzzy preference relation, defining for each action two functions ϕ^+ and ϕ^-, from which separate rankings of alternatives are obtained. The function $\phi^+(i)$ is based on the 'better' and 'much better' preference relations and with a value from 0 to 1 indicates how the alternative i is 'better' then all other alternatives. The second function $\phi^-(i)$ is based on the 'worse' and 'much worse' preference relations, its value going from 0 to 1 which indicates how the alternative i is 'worse' than all other alternatives (Menegolo and Guimarães Pereira 1996). The final ranking of alternatives comes from the intersection of two separate rankings obtained by means of the functions $\phi^+(i)$ and $\phi^-(i)$, taking into account that it can also be an incomplete ranking because non-dominated alternatives are calculated.

Additionally, NAIADE allows for another type of evaluation. It analyses conflicts between different interest groups and the possible formation of coalitions according to the proposed alternative options. Besides the 'impact matrix', also an 'equity matrix' is constructed, which contains linguistic evaluations of different social groups for each alternative. In particular, 'equity analysis is performed by the completion of an equity matrix from which a similarity matrix is calculated. Through a mathematical reduction algorithm, it is possible to build a dendrogram of coalitions which shows possible coalition formation, and a level of conflict among the interest groups' (Menegolo and Guimarães Pereira 1996: 1).

Because the starting point of NAIADE is the construction of an evaluation matrix, it is not possible with this method to take into account the interdependence existing between some different evaluation criteria. Since there is no limit on the choice of the number of criteria that can be inserted in the evaluation matrix, it is possible to consider all the criteria required by the decision-making problem. In this sense a complete (i.e. exhaustive) list of evaluation criteria is allowed. No weights are assigned to the criteria. The final result offered by the method can be an incomplete ranking of the evaluated alternatives. In fact, the final ranking comes from the intersection of two separate rankings which allows to calculate also the non-dominated alternatives.

Stakeholder participation is explicitly supported in NAIADE. The method allows the construction of an equity matrix which reflect how the different social groups involved evaluate the alternatives. It is also possible to promote the actors communication if stakeholders are called to discuss their evaluation of the alternatives. The method is used to cope with different decision-making problems (it has especially been used for problems of unsustainability) at different geographical scales. NAIADE allows using qualitative and quantitative data. In particular, qualitative data are expressed by means of linguistic evaluations and quantitative data may be expressed either in crisp, stochastic or fuzzy numbers.

Programming methods

Multi-Objective-Programming (MOP), Goal Programming (GP) and their variants are continuous multiple criteria methods. In contrast to the methods presented so far, the programming methods do not rank or sort a finite number of alternatives, but the alternatives are generated during the solution process on the basis of a mathematical model formulation.

The roots of GP lie in the work by Charnes and Cooper in the early 1960s. After the mid-1970s it became widely applied due to seminal work by Lee (1972) and Ignizio (1976). A number of specific software tools were developed, but often general programming software packages (like GAMS) are applied. From a procedural point of view MOP consists of two steps: (1) identification of Pareto efficient (also called non-dominated) solutions; and (2) identification of the most preferred alternative together with the decision-maker(s).

Step 1 consists of finding the non-dominated solutions. The problem is formulated as a task of simultaneous maximization/minimization of several objectives subject to a set of constraints. In a so-called criterion space it can be defined as follows: 'max' q, s.t. $q \in Q$, where Q is a so-called feasible region in the criterion space. The set Q is not specified directly, but by means of decision variables as usually done in single optimization problems: 'max' $q = f(x) = (f_1(x),...,f_k(x))$, s.t. $x \in X$, where X is a feasible set. The functions f_i, $i = 1,2,...,k$ are objective functions (Korhonen 2001).

Since the simultaneous optimization of all objectives is impossible – given a certain level of conflict between them in most real problems – MOP tries to find the set of Pareto efficient solutions. According to Ballestero and Romero (1998) a necessary condition to guarantee the rationality of any solution to an MCDA problem is the following: 'A set of solutions in a MCDA problem is Pareto efficient (also called non-dominated), if their elements are feasible solutions such that no other feasible solution can achieve the same or better performance for all the criteria being strictly better for at least one criterion' (p.7). The goal is therefore to divide the feasible set into a subset of Pareto efficient solutions and a subset of Pareto inefficient solutions.

Step 2 requires choosing the most preferred solution. Any choice from among the set of efficient solutions is an acceptable and 'reasonable' solution, as long as we have no additional information about the decision-maker's preference structure. The final solution $q \in Q$ is called 'the most preferred solution'. It is a solution preferred by the decision-maker to all other solutions (Korhonen 2001). Since we are always facing a number of conflicting goals in multiple criteria problems, a solution can never be optimal (or ideal), but rather a compromise. With regards to how much we need to know about the decision-maker's preferences, different approaches were developed: there are (a) those that need knowledge of the decision-maker's preference; (b) those in which the decision-maker's preferences are pre-emptively determined; and (c) those that progressively reveal the preferences of the decision-maker through man–machine interactions (interactive methods) (Kalu 1999). The latter typically operate with decision-maker's aspiration levels regarding the objectives on the feasible region. The aspiration levels are projected via minimizing so-called achievement scalarizing functions (Wierzbicki 1980). No specific behavioural assumptions, e.g. transitivity, are necessary (Korhonen 2000).

At the beginning of MOP research the expectation was that computing all efficient extreme points and then asking the decision-makers to examine their criterion vectors to select the best one could solve the problem. However, after vector-maximum codes became available, it was found that the number of efficient extreme points generated by MOP was by far larger than had been expected. As a result, interactive procedures for exploring efficient sets for the best solution became popular. Many different non-reference point and reference point procedures are available now (Steuer and Gardiner 1990).

In GP the impossibility of building a reliable mathematical representation of the decision-maker's preferences due to conflicts of interest and incompleteness of available information is addressed in a different way. Two major subsets of GP models can be distinguished: In the first type it is assumed that within this kind of decision environment decision-makers attempt to achieve a set of relevant goals as close as possible to the set of targets established (Ballestero and Romero 1998). The algebraic representation is given as the sum of unwanted deviations from target values if a number of objectives are minimized (weighted GP):

$$\min z \sum_{i=1}^{k} (u_i n_i + v_i p_i), \text{ s.t. } f_i(x) + n_i - p_i = b_i, i = 1 \ldots Q, x \in C_5,$$

where $f_i(x)$ is a linear function (objective) of x, and b_i is the target value for that objective. n_i and p_i represent the negative and positive deviations from this target value. u_i and v_i are the respective positive weights attached to these deviations in the achievement function z (Tamiz *et al.* 1998: 570). The methodology rests on the following basic scheme (Vincke 1992): (1) set the values one wishes to attain on each criterion (the objectives); (2) assign priorities (weights) to these objectives; (3) define (positive or negative) deviations with respect to these objectives; (4) minimize the weighted sum of these deviations; and (5) perform a sensitivity analysis. In practice, the weights are sometimes derived from the pairwise comparison of AHP. Alternatively, they can be derived by use of an interactive MCDA method, like the one by Zionts and Wallenius (1976) (as shown in Lara and Romero 1992). Variants of this type (Chebyshev GP, MinMax GP, and Fuzzy GP) seek to minimize the maximum deviation from amongst the set of unmet objectives instead. In the case of another variant (Preference-Function GP) the relationship between the deviation from the target and the penalty imposed is not the standard linear one, but it may be piecewise linear, discontinuous, or even non-linear (Tamiz and Jones 1997).

In the other major subset of GP the deviational variables are assigned to a number of priority levels and minimized in a lexicographic sense (lexicographic GP). The assumption of non-continuity of preferences implies the impossibility of ordering the decision-maker's preferences by a monotonic utility function. Lexicographic minimization means a sequential minimization of each priority while maintaining the minimal values achieved by all higher priority level minimizations. In mathematical terms this means:

$$\text{Lex min } a = (g_1(n,p), g_2(n,p), \ldots, g_L(n,p)), \text{ s.t. } f_i(x) + n_i - p_i = b_i, \; i = 1, \ldots, Q$$

The given model has L priority levels, and Q objectives. a is an ordered vector of these L priority levels. n_i and p_i are deviational variable which represent the under- and over-achievement of the i-th goal, respectively. x is the set of decision variables to be determined. The g (within a priority level) function is given by

$$g_l(n,p) = u_{l_q} n_1 + \ldots + u_{l_q} n_q + v_{l_1} p_1 + \ldots + v_{l_Q} P_Q ,$$

where u and v represent inter-priority level weights (Tamiz *et al.* 1998: 570). The possibility to account for the non-continuity of preferences seems particularly important for the analysis of environment problems (Spash 2000).

Instead of optimization, GP is based on Simon's satisfying philosophy (Simon 1955, 1979). In comparison, GP is a more pragmatic approach. It was used in the 1960s and 1970s by practitioners in corporations and agencies and only slowly received more attention by theorists. While the techniques applied are the same programming techniques as for the other MOP methods, the underlying philosophy differs. The idea of optimizing is abandoned; only satisfying is attempted.

MOP and particularly GP are the most frequently applied multiple criteria methods. They have been applied to a wide array of problems, namely production planning, oil refinery scheduling, health care, portfolio selection, distribution system design, energy planning, water reservoir management, timber harvest scheduling, problems of wildlife management, etc. Their popularity can be explained by their flexibility; they can account for diverse variable types (continuous, integer, Boolean, etc.) as well as constraints and objective functions (linearity, convexity, differentiability, etc.) (Vincke 1992).

More recent additions to the field of MOP have been: (a) to allow for fuzziness in the data which is a way to address the problem of uncertainty to some degree; and (b) evolutionary algorithms which were developed as a response to the fact that for models which are large in size and/or significantly non-linear, traditional solution methods are often unable to identify the global optimum. In such cases, typically linear approximations are applied instead in order to find a solution. Genetic algorithms were shown to be applicable to examples of large non-linear models (e.g. Mardle *et al.* 2000). They follow the concept of solution evolution, by stochastically developing generations of solutions populations using a given fitness statistic.

MOP and GP allow addressing a completely different type of multiple criteria problems. Instead of choosing from a limited number of alternatives, a continuous set of alternatives is generated during the solution process. They are the only methodologies in the field which provide the model-based calculation of non-dominated alternatives. Another advantage is that restrictions which are for instance needed for the analysis of decision-making processes based on criteria of strong sustainability (non-substitutability) can be accounted for. These special features turn out to be important for the decision-making situation in many case studies.

GP has often been criticized for being analytically inferior to MOP. While GP is indeed a somewhat simpler method, it is often preferred because of its capacity to handle large-scale operational real-life problems. MOP works efficiently for moderate-size problems defined within a well-structured environment. When programming methods are applied, interdependence between criteria is not allowed. The completeness description of the phenomenon is desired, but not explicitly mentioned. While MOP is very appealing theoretically it has the disadvantage that for large problems and particularly if non-linear functions are included, often no optimal solution can be found.

The development of interactive methods, which allow to progressively reveal the preferences of the decision-maker through man–machine interactions, made MOP much less reliant on doubtful assumptions about the decision-maker's preferences. A number of useful software tools were developed which allow to give the decision-makers an important role in the decision process and to enable 'psychological convergence' (i.e. termination of the interactive process as and when the decision-maker chooses). Hence, interactive MOP leads not only to a good structuring of the problem but can also provide a useful tool for learning for the decision-maker involved. A problem remaining with interactive methods is that mostly they do not provide an objective basis upon which the decision-maker could base the decision to terminate the interactive process (Schniederjans 1995).

The next step in perceiving the role of the decision-makers in a more democratic way, namely accounting for stakeholder participation (group decision-making), is however explicitly accounted for in few tools and is therefore exhausted in only some case studies. Another drawback in the user context is that the process cannot be made very transparent. MOP allows for the use of different geographical scales, as well as different general scales (micro–macro levels). Concerning data, MOP is able to deal with quantitative as well as qualitative and fuzzy data and allows for risk via fuzziness. An interesting aspect with regard to the fuzzy version of MOP was reported in several case studies: the large amount of time which is normally needed to select all the data required can on some occasions be bypassed by using the fuzzy relations (e.g. Chang *et al.* 1997). GP is not able to deal with a lot of different data, such as qualitative and fuzzy sets.

Comparison of the methods

The evaluations of the individual methods using the quality criteria pointed to the differences in the characteristics of the presented methods. The summary in Table

5.2 gives a detailed characterization of each method. Without reference to the specific characteristics of the case study for which the methods should be applied, the selection of the best applicable method is not possible. Consequently the aim here is a characterization of the methods based on the quality criteria with regard to sustainability issues in general. From the list of quality criteria we find the following four groups most crucial:

- possibility to deal with complex situations (criteria, consideration of different scales and aspects, i.e. geographical scales, micro–macro link, societal/technical issues, type of data, uncertainties),
- possibility to consider non-substitutability (strong sustainability) issues,
- possibility to involve more than one decision-maker (stakeholder participation, actors communication, and transparency), and
- information of stakeholders in order to increase their knowledge and change their opinion and behaviour (problem structuring, tool for learning, transparency, type of weights).

Some criteria could not be considered here, because they are especially designed to help the user in the selection of a method for a concrete decision-making situation and not designed for a theoretical method comparison with regard to general sustainability issues, without any specific application in mind, as done in this comparison. Reference to the whole list of criteria was made in De Montis *et al.* (2000), which included selected case studies of sustainable water use.

In handling complex situations, all methods show similar performance with respect to the aspects and scales which can be considered; but they all show weaknesses in the characteristics of the evaluation criteria which should be used as operational components. Only AHP allows the interdependence of evaluation criteria, while only MAUT and NAIADE allow non-linear preferences. Evamix, ELECTRE III, and Regime do not permit either of these characteristics. In each case quantitative and qualitative data can be addressed. In ELECTRE III qualitative data have to be transformed first to a quantitative scale. Additionally, some of the methods like GP/MOP and NAIADE include features to deal with fuzzy data and stochastic numbers, which is the way that these methods deal with risk and uncertainties. MAUT allows to deal with risk (not uncertainty) concerning the outcomes of the alternatives by assigning probabilities to the utility functions (von Neumann–Morgenstern utility functions). All of the methods provide the possibility to carry out a sensitivity analysis or to apply qualitative data to consider uncertainties and risks. Non-substitutability aspects can only be properly analysed by using GP/MOP and ELECTRE III which allow to set constraints and thresholds explicitly. Within the remaining methods non-substitutability can only be considered by assigning flags to the respective criteria.

The performance to inform stakeholders in order to increase their knowledge and change their opinion and behaviour is satisfactory for all methods. In MOP/GP, ELECTRE III and Regime the transparency is only of medium quality, which hinders stakeholder participation. The other methods are satisfactory or even good concerning these aspects. NAIADE does not allow to assign weights explicitly but is the only methodology which provides an explicit structure for stakeholder participation. For MOP/GP there are tools for interactive decision-making as well as for group decision-making. While the majority of applications did not make use of them, these tools have been available for more than ten years now.

The methods also differ in the possibilities for involving more than one decision-maker in the process. For MOP/GP there are tools for group decision-making, but again, they have not been applied much. Some methods like AHP support consideration of preferences of several decision-makers. In practice, each decision-maker is asked individually. Concerning MAUT, if *n* persons are involved in the decision, each person is asked for their single utility function for any attribute; these single functions are then aggregated to *n* multi-attribute utility functions. If they are in conflict with each other, group solutions are possible, but are not really treated by the literature about MAUT. Concerning stakeholder information, MAUT, AHP, and Evamix provide good transparency and allow to give weights explicitly. In a very interesting way, NAIADE supports the involvement of more than one person in the decision-making process, namely with reference to their different interests which allows for an explicit conflict analysis. In general, an important aspect in group decision-making is how transparent the decision-making process is, as this enhances possibilities of participation and goal-oriented discussion within the group.

Evamix, ELECTRE III, and Regime provide only satisfactory performance for complex situations, so they are not the first choice in this context. Non-substitutability aspects could best be considered by using ELECTRE III or GP/MOP. MAUT and Evamix have acceptable performance in general, however they lack important aspects crucial for sustainability issues, namely the ability to account for uncertainty and non-substitutability. NAIADE is a specialist to involve more than one decision-maker, while GP/MOP and ELECTRE III allow dealing with non-substitutability. Thus, a clear ranking of the methods concerning the different quality criteria is impossible, as each has its own strengths and weaknesses in different areas. The characterization shows, however, that depending on the specific problem the choice of one method over the other is advisable and that some methods are largely useless for the analysis of sustainability problems.

Conclusions

The goal of more sustainable development has been studied extensively during the last fifteen years; but a key criticism has been vagueness as to recommendations and lack of operational applications. The acknowledgement of the limits of growth has led to the revision of the paradigms of economics regarding human and natural resources. Besides well-known goals like efficiency and income distribution, many other conflicting factors need to be included in the analysis. MCDA constitutes a powerful tool that is able to take into account several concerns and to foster participation and learning processes among politicians and citizens. Thus, MCDA is useful for decision-making processes in the context of sustainability (e.g. Munda

1995; Castells and Munda 1999; Janssen 1992; Andreoli and Tellarini 2000; Mendoza and Prabhu 2000).

In contrast to earlier studies which compare MCDA methods with respect to more technical issues (e.g. Olson 2001, Tecle *et al.* 1997) we used a list of quality criteria for the assessment (De Montis *et al.* 2000) and concentrated on the most relevant criteria for sustainability issues. The comparison shows that the indicator list is applicable and useful for the evaluation of MCDA methods in the context of sustainability. As a result, a detailed characterization of the different MCDA methods was derived. Some rough guidelines which can be given from the comparison of the methods are:

1 If the respective decision problem is such that relying upon social welfare theory and its assumptions is possible and if the data to build utility functions is available (risk and qualitative data are possible) then MAUT is a good choice.
2 If working with different conflicting interest groups is important for the case, NAIADE and AHP provide the best performance.
3 If the decision-makers involved should primarily learn from the application of the MCDA tool, it is advisable to use MAUT or AHP.
4 If thresholds and constraints are central for the problem under investigation, which means that there is non-substitutability of some criteria, ELECTRE III or GP/MOP should be chosen.
5 If the problem is a continuous one, i.e. there is no discrete number of alternatives which comes out of the specific situation, GP or MOP should be chosen.
6 If a complete ranking of the given alternatives as a result of the analysis is indispensable, MAUT, AHP, Evamix, or Regime should be applied.

Further research should focus on the relationship between sustainability and evaluation with respect to the process of consensus building. Recent research studies (Hewitt 1995) state that general principles of environmental action consist of environment management and of democratic decision-making. In this sense, the technical features are not that important; rather, each multiple criteria technique should be evaluated according to its capacity to foster learning processes and to allow users to become aware of society's strategic role in environmental decision-making.

Notes

1 Risk and uncertainty are sometimes used interchangeable in the MCDA literature. As the difference is very important in the context of sustainability, we would like to distinguish between them. We use 'risk', if a probability distribution can be assigned for the unknown effects; otherwise the term 'uncertainy' is used (Knight 1921).
2 Utility independence: $u_i(x_i)$ is unaffected by levels of the other attributes. Preferential independence: the preferences for some attribute levels do not depend on the levels fixed for the other attributes.
3 No individual will sense a difference between a cup of coffee with *i* milligrams of sugar and one with $i + 1$ milligrams of sugar, but will express a preference between a cup with a lot of sugar and one with none; this contradicts transitivity of indifferences (Vincke 1992).

4 'In general what we mean by being consistent is that when we have a basic amount of raw data, all other data can be logically deduced from it. In doing pairwise comparison to relate *n* activities so that each one is represented in the data at least once, we need *n* - 1 pairwise comparison judgements. From them all other judgements can be deduced simply by using the following kind of relation: if activity A_1 is 3 times more dominant than activity A_2 and activity A_1 is 6 times more dominant than activity A_3 then $A_1 = 3A_2$ and $A_1 = 6A_3$. It should follow that $3A_2 = 6A_3$ or $A_2 = 2A_3$ or $A_3 = \frac{1}{2}A_2$. If the numerical value of the judgement in the (2, 3) position is different from 2 then the matrix would be inconsistent. This happens frequently and it is not a disaster' (Saaty 1980: 18).

5 For more information on the 'random index' (*R.I.*) see Saaty (1980: § 1.4). In the following table values of the random index for matrix of orders from 1 to 11 are reported:

Order of the	1	2	3	4	5	6	7	8	9	10	11
Random index	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

6 Elements of qualitative criteria are expressed by means of symbols like +++, ++, +, 0, –, – –, – – –, where $\ldots \succ ++++ \succ ++ \succ + \succ 0 \succ - \succ -- \succ --- \succ \ldots$.

7 An example of the way to obtain cardinal weights from ordinal preferences is given by Voogd (1983: 175): 'if $\omega_1 \le \omega_2 \le \omega_3$ and it is assumed that weights meet the condition

$$\sum_{j=1}^{m} w_j = 1$$

then the following extreme sets $w_\ell, \ell = 1, 2, 3$ can be distinguished:

$$w_1 = (1, 0, 0), \quad w_2 = \left(\frac{1}{2}, \frac{1}{2}, 0\right), \quad w_3 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)'.$$

8 For more information about the mathematical aspects of Evamix, see Voogd (1983: ch. 10).

9 Using stochastic numbers, it is necessary to choose the probability density functions; using fuzzy numbers, it is necessary to define their membership functions; using qualitative evaluations, some pre-defined linguistic variables such as 'good', 'moderate', 'bad' and so on are considerate. It is not possible to assign different type of measurements (i.e. crisp, stochastic, fuzzy, linguistic) to the same criterion for different alternatives.

10 In the case of two crisp numbers, the 'distance' is defined as their difference; in the case of stochastic and fuzzy numbers, the 'semantic distance' measures the distance between two functions (i.e. probability density functions or fuzzy membership functions); the linguistic variables are treated as fuzzy sets.

11 For linguistic variables may be used the equivalent preference relations: much better than (>>); better than (>), approximately equal to (≅); very equal to (=); worse than (<); much worse than (<<). In this case the six associated analytic functions express an index of credibility of the statements that an alternative is 'much better', 'better', 'approximately equal', 'very equal', 'worse' and 'much worse' than another.

12 For more information about the mathematical aspects of NAIADE, see Munda (1995: ch. 7).

13 The symbol * stands for >>, >, ≅, < or <<.

14 'Entropy measure of fuzziness can be easily extended to the notion of fuzzy relations … to have information on the diversity among the assessments of the single fuzzy relations, according to each criterion' (Munda 1995: 138). More precisely 'entropy is calculated as an index varying from 0 to 1 that gives an indication of the variance of the credibility indexes that are above the threshold, and around the crossover value 0.5 (maximum fuzziness). An entropy value of 0 means that all criteria give an exact indication (either definitely credible or definitely non-credible), whereas an entropy value of 1 means that all criteria give an indication biased by the maximum fuzziness (0.5)' (Menegolo and Pereira 1996: 5). For the mathematic formulae of entropy, see Munda (1995: § 7.4.)

References

Andreoli, M. and Tellarini, V. (2000) 'Farm sustainability evaluation: methodology and practice', *Agriculture, Ecosystems and Environment*, 77 (1–2): 43–52.

Ballestero, E. and Romero, C. (1998) *Multiple Criteria Decision Making and its Applications to Economic Problems*, Boston/Dordrecht/London: Kluwer Academic.

Castells, N. and Munda, G. (1999) 'International environmental issues: towards a new integrated assessment approach', in M. O'Connor and C. Spash. (eds) *Valuation and the Environment – theory, method and practice*, Cheltenham, UK/Northampton, MA: Edward Elgar: 309–27.

Chang, N.-B., Wen, C.G. and Chen, Y.L. (1997) 'A fuzzy multi-objective programming approach for optimal management of the reservoir watershed', *European Journal of Operational Research*, 99: 289–302.

Churchman, C.W., Ackoff, R.L. and Arnoff, E.L. (1957) *Introduction to Operations Research*, New York: Wiley.

Climaco, J. (1997) *Multiple criteria Analysis.* Proceedings of the XIth International Conference on MCDM, Coimbra, Portugal: Springer, 1–6 August 1994.

De Montis, A., De Toro, P., Droste-Franke, B., Omann, I. and Stagl, S. (2000) 'Criteria for quality assessment of MCDA methods', paper presented at the 3rd Biennial Conference of the European Society for Ecological Economics, Vienna, 3–6 May 2000.

Dillon, J.L. and Perry, C. (1977) 'Multiattribute utility theory, multiple objectives and uncertainty in ex ante project evaluation', *Review of Marketing and Agricultural Economics*, 45 (1, 2): 3–27.

Fandel, G. and Spronk, J. (eds) (1985) *Multiple Criteria Decision Methods and Applications.* Selected readings of the First International Summer School Acireale, Sicily, September 1983, Berlin: Springer.

Fishburn, P.C. (1968) 'Utility theory', *Management Science*, 14: 335–78.

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

Fishburn, P.C. (1978) 'A survey of multiattribute/multiple criteria evaluation theories', in S. Zionts (ed.) *Multiple Criteria Problem Solving*, Berlin: Springer: 181–224.

Funtowicz, S.O. and Ravetz, J. (1990) *Uncertainty and Quality in Science for Policy*, Dordrecht: Kluwer.

Hewitt, N. (1995) *European Local Agenda 21 Planning Guide – How to engage in long term environmental action planning towards sustainability?* Brussels: European Sustainable Cities and Towns Campaign.

Hinloopen, E. (1985) *De Regime Methode*, MA thesis, Interfaculty Actuariat and Econometrics, Free University Amsterdam.

Hinloopen, E. and Nijkamp, P. (1990) 'Qualitative multiple criteria choice analysis', *Quality and Quantity*, 24: 37–56.

Hinloopen, E., Nijkamp, P. and Rietveld, P. (1983) 'Qualitative discrete multiple criteria choice models in regional planning', *Regional Science and Urban Economics*, 13: 73–102.

Hwang, C. and Yoon, K. (1981) *Multiple Attribute Decision Making*, Berlin: Springer.

Ignizio, J.P. (1976) *Goal Programming and Extensions*, Lexington: Lexington Books.

Janssen, R. (1992) *Multiobjective Decision Support for Environmental Management*, Dortrecht: Kluwer Academic.

Kalu, T.C.U. (1999) 'An algorithm for systems welfare interactive goal programming modelling', *European Journal of Operational Research*, 116: 508–29.

Keeney, R.L. and Raiffa, H. (1976) *Decisions with Multiple Objectives*, New York: John Wiley.

Kmietowicz, Z.W. and Pearman, A.D. (1981) *Decision Theory and Incomplete Knowledge*, Aldershot: Gower.

Knight, F. (1921) *Risk, Uncertainty, and Profit*, Boston: Houghton Mifflin.
Korhonen, P. (2001) 'Multiple objective programming support', in C.A. Floudas and P.M. Pardalos (eds) *Encyclopedia of Optimization*, Dortrecht: Kluwer: Vol.3, 566–574.
Lara, P. and Romero, C. (1992) 'An interactive multigoal programming for determining livestock rations: an application to dairy cows in Andalusia (Spain)', *Journal of the Operational Research Society*, 43: 945–953.
Lee, S.M. (1972) *Goal Programming for Decision Analysis*, Philadelphia, PA: Auerbach.
Luce, R.D. (1956) 'Semiorders and a theory of utility discrimination', *Econometrica*, 24: 178–91.
Mardle, S.J., Pascoe, S. and Tamiz, M. (2000) 'An investigation of genetic algorithms for the optimization of multi-objective fisheries bioeconomic models', *International Transactions in Operational Research*, 7(1): 33–49.
Mastenbroek, P. and Paelinck, J.H.P. (1977) 'Qualitative multiple criteria analysis – applications to airport location', *Environment and Planning A*, 9 (8): 883–95.
Mendoza, G.A. and Prabhu, R. (2000) 'Multiple criteria decision making approaches to assessing forest sustainability using criteria and indicators: a case study', *Forest Ecology and Management*, 131(1–3): 107–26.
Menegolo, L. and Guimarães Pereira, A. (1996) *NAIADE Manual*, Joint Research Centre – EC, Institute for System, Informatics and safety, Ispra (VA), Italy.
Munda, G. (1995) *Multiple Criteria Evaluation in a Fuzzy Environment – theory and applications in ecological economics*, Heidelberg: Physika Verlag.
Munda, G., Nijkamp, P. and Rietveld, P. (1994) 'Fuzzy multigroup conflict resolution for environmental management", in J. Weiss (ed.) *The Economics of Project Appraisal and the Environment*, Aldershot: Edward Elgar.
von Neumann, J. and Morgenstern, O. (1947) *Theory of Games and Economic Behaviour*, Princeton, NJ: Princeton University Press.
Nijkamp, P. and Voogd, H. (1981) 'New multiple criteria methods for physical planning by means of multidimensional scaling techniques', in Y. Haimes and J. Kindler (eds) *Water and Related Land Resource System*, Oxford: Pergamon Press: 19–30.
Olson, D.L. (2001) 'Comparison of three multicriteria methods to predict known outcomes', *European Journal of Operational Research*, 130(3): 576–87.
Rietveld, P. (1980) *Multi Objective Decision Making and Regional Planning*, Amsterdam: North Holland.
Roy, B. (1985) *Méthodologie Multicritère d¢Aide à la Décision*, Parigi: Economica.
Roy, B. (1996) *Multiple Criteria Methodology for Decision Aiding*, Dordrecht: Kluwer Academic.
Roy, B. and Bouyssou, D. (1985) 'An example of comparison of two decision-aid models', in G. Fandel and J. Spronk (eds) *Multiple Criteria-Decision Methods and Applications.* Selected readings of the First International Summer School Acireale, Sicily, September 1983, Berlin: Springer: 361–81.
Saaty, T.L. (1980) *The Analytic Hierarchy Process for Decision in a Complex World*, Pittsburgh, PA: RWS Publications.
Saaty, T.L (1988) *Decision Making for Leaders*, Pittsburgh, PA: RWS Publications.
Saaty, T.L. and Alexander, J.M. (1989) *Conflict Resolution – The Analytic Hierarchy Process*, New York: Praeger.
Saaty, T.L. (1992) *Multicriteria Decision Making – The Analytic Hierarchy Process*, Pittsburgh, PA: RWS Publications.
Saaty, T.L. (1994) *Fundamentals of Decision Making and Priority Theory with the Analytical Hierarchy Process*, Pittsburgh, PA: RWS Publications.
Saaty, T.L. and Forman, E. (1993) *The Hierarchon*, Pittsburgh, PA: RWS Publications.

Saaty, T.L. and Vargas, L.G., (1991) *Prediction, Projection and Forecasting*, Dordrecht: Kluwer Academic Publishers.

Salminen, P., Hokkanen, J. and Lahdelma, R. (1998) 'Comparing multicriteria methods in the context of environmental problems', *European Journal of Operational Research*, 104: 485–96.

Scarelli, A., (1997) *Modelli matematici nell¢analisi multicriterio*, Viterbo: Edizioni Sette Città.

Simon, H. (1955) 'A behavioral model of rational choice', *Quarterly Journal of Economics*, 69: 99–118.

Simon, H. (1979) 'Rational decision making in business organizations', *American Economic Review*, 69: 493–513.

Schniederjans, M.J. (1995) *Goal Programming Methodology and Applications*, Boston: Kluwer Publishers.

Spash, C.L. (2000) 'Multiple value expression in contingent valuation: economics and ethics', *Environmental Science & Technology*, 34(8): 1433–8.

Steuer, R.E. and Gardiner, L.R. (1990) 'Interactive multiple objective programming: concepts, current status, and future directions', in C.A. Bana e Costa (ed.) *Readings in Multiple Criteria Decision Aid*, Berlin/Heidelberg/New York: Springer: 413–44.

Taha, H.A. (1976) *Operations Research*, New York: Macmillan.

Tamiz, M. and Jones, D. (1997) 'An example of good modelling practice in goal programming: means to overcoming incommensurability', in R. Caballero, F. Ruiz and R.E. Steuer (eds) *Advances in Multiple Objective and Goal Programming*, Berlin/Heidelberg/New York: Springer: 29–37.

Tamiz, M., Jones, D. and Romero, C. (1998) 'Goal programming for decision making: an overview of the current state-of-the-art', *European Journal of Operational Research*, 111: 569–81.

Van Delft, A. and Nijkamp, P. (1977) *Multiple Criteria Analysis and Regional Decision Making*, The Hague/Boston: Martinus Nijhoff.

Vincke, P. (1985) 'Multiattribute utility theory as a basic approach', in G. Fandel and J. Spronk (eds) *Multiple Criteria-Decision Methods and Applications.* Selected readings of the First International Summer School Acireale, Sicily, September 1983, Berlin: Springer: 27–40.

Vincke, P. (1992) *Multiple Criteria Decision-Aid*, New York: John Wiley & Sons.

Voogd, H. (1981) 'Multicriteria analysis with mixed qualitative–quantitative data', Delft University of Technology, Department of Urban and Regional Planning, *Planologisch Memorandom*, 81: 86.

Voogd, H. (1983) *Multiple Criteria Evaluation for Urban and Regional Planning*, London: Pion.

Wierzbicki, A. (1980) 'The use of reference objectives in multiobjective optimization', in G. Fandel and T. Gal (eds) *Multiple Objective Decision Making, Theory and Applications*, New York: Springer.

Zionts, S. and Wallenius, J. (1976) 'An interactive programming method for solving the multiple objective programming problem', *Management Science*, 22(6): 652–63.