Giuseppe Munda

Social Multi-Criteria Evaluation for a Sustainable Economy

Springer

Social Multi-Criteria Evaluation
for a Sustainable Economy

Giuseppe Munda

Social Multi-Criteria Evaluation for a Sustainable Economy

Springer

Giuseppe Munda
Joint Research Centre of the European Commission (JRC)
Institute for the Protection and Security of the Citizen (IPSC)
Econometrics and Applied Statistics Unit (EAS)
TP 361
21020 Ispra (VA)
Italy
giuseppe.munda@jrc.it
and
Universitat Autonoma de Barcelona
Department of Economics and Economic History
Institute of Environmental Sciences and Technologies
08193 Bellaterra (Barcelona)
Spain

ISBN-13: 978-3-540-73702-5 e-ISBN-13: 978-3-540-73703-2

Library of Congress Control Number: 2007938411

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

To my wife Michela
and my children Sofia and Enrico

“Dal primo giorno ch’io vidi il suo viso - in questa vita, infino a questa vista, - non m’è il seguire al mio cantar preciso; …”

Dante Alighieri – *La Divina Commedia*, “Paradiso, XXX, 28–33”.

Preface

The real world is characterized by deep *complexity*. May be a rather unremarkable observation, yet it has important implications on the manner policy problems are represented and decision-making is framed.

Is contemporary democracy compatible with science in real-world policy-making? This book gives answers in the affirmative. It also asserts that this congruence can have positive implications not only in terms of economic prosperity but also when dealing with the difficult sustainability policy problems of our millennium.

To address contemporary issues economic science will have to expand its empirical relevance by introducing more and more realistic assumptions to its models. One of the most interesting research orientations in recent times in the field of public economics is the explicit attempt to take account of political constraints, interest groups and collusion effects.

One of the main novelties of this book is its establishment of a clear relationship between social and public choice theories on one hand, and multiple criteria decision analysis on the other. The pioneering research developed by Arrow and Raynaud (1986) has shown that the relationships between multi-criteria decision theory and social choice are clear and relevant. The main directions of cross-fertilization between these research fields are twofold:

1. Multi-criteria decision theory can be an adequate framework for applied social (and public) choice.
2. Social choice can produce interesting theoretical results for ensuring the axiomatic consistency needed by multi-criterion aggregation conventions.

The first direction was not pursued at all by Arrow and Raynaud. These authors' explicit interest is in the so-called "*industrial outranking problem*", whose aim is to help decisions of business-people.

In empirical evaluations of public policies, multi-criteria decision theory is an adequate policy tool, since it allows us to take a wide range of assessment criteria into account (e.g. environmental impact, distributional equity, etc.), and not simply profit maximization, as a private economic agent would do. Thus the essential meaning of multi-criteria evaluation in a social context is simply *tolerance and democracy*. This is a very important feature when dealing with sustainability issues, since conflict among differing but equally legitimate values and interests is a normal state of affairs in these kinds of complex policy problems. Given that in a sustainability context both market and government failures can occur, economic optimization cannot be the only evaluation criterion, nor can mythical, benevolent policy-makers offer any unilateral, optimal solution.

For these reasons the new concept of social multi-criteria evaluation (SMCE) is proposed here as a tool to integrate different scientific languages in a public choice framework, when "civil society" and ethical concerns about future generations have to be considered, along with policy imperatives and market conditions.

In sum, we can say that this book attempts to combine both public and social choice traditions with multi-criteria decision analysis in order to deal with sustainability paradoxes in a complex world with multiple dimensions, values and scales.

In the light of the previous observations, the book is organized into three main parts:

Part A

Methodological Foundations and Operational Consequences of Social Multi-Criteria Evaluation (SMCE)

The main objective of this part is the proposal of the concept of "social multi-criteria evaluation" (SMCE) as a potentially useful framework for the application of public choice to the complex policy problems of our millennium, in which, as described by Funtowicz and Ravetz, "facts are uncertain, values in dispute, stakes high and decisions urgent". The following main questions are dealt with:

1. Why "social" multi-criteria evaluation?
2. How should such an approach be developed?

The foundations of SMCE are based on concepts coming mainly from microeconomics, complex systems theory and philosophy, such as behavioural assumptions, cost–benefit analysis, reflexive complexity, post-normal science and incommensurability. Lessons from real-world case studies are also dealt with.

Part B

Consistency in Social Multi-Criteria Evaluation

Mathematical algorithms have the important objective of guaranteeing consistency between the problem structuring and the ranking of feasible policy options. In Part B, desirable properties for formal procedures in SMCE are studied. The basic concepts of multi-criteria decision analysis and the most commonly used "multi-criteria methods" are critically reviewed. Given that the discrete multi-criterion and the social choice problem have many characteristics in common, some results of social choice are used to improve the axiomatic consistency of multi-criterion mathematical procedures. The treatment of technical uncertainty of both stochastic and fuzzy nature is also dealt with in detail.

Part C

Mathematical Procedures to Search for Technical and Social Compromise Solutions

A new mathematical aggregation convention explicitly designed for social multi-criteria evaluation problems is developed. This algorithm is combined with an eclectic procedure, based on concepts coming from land-use planning, fuzzy cluster analysis and social choice. The objective of this procedure is to illuminate distributional issues.

All properties respected by the two proposed procedures are clearly illustrated by means of their formal, descriptive and normative meanings. Musgrave's distinction among negligibility assumptions, domain assumptions and heuristic assumptions is also used. Annex gives an empirical example of combining multi-criteria evaluation with sensitivity analysis.

Throughout the whole book various examples of real-world applications in fields such as publicly provided goods, land-use planning, water and renewable energy policies and composite indicators are discussed.

Acknowledgements

First of all, I would like to thank all the staff of the Department of Economics and Economic History and the Institute of Environmental Sciences and Technologies of the Universitat Autonoma de Barcelona for having welcomed me to Catalunya, moltes gracies! Warmest thanks are also addressed to my more recent colleagues at the Joint Research Centre of the European Commission in Ispra in Italy. All of them have provided me with a priceless resource: the time to complete the writing of this book in the most peaceful and at the same time most stimulating working environment one could desire.

Many thanks are due to the numerous people (really too many to list) whose willingness to spend time sharing and discussing ideas with me has been fundamental to the development of many of the concepts presented in this book.

The presentation of the content of this book has benefited greatly from experience accumulated with the wide variety of students I had the privilege to teach in Barcelona and in many different parts of the world. To them all, my warmest thanks.

Different institutions have provided financial support for most of the research needed to write this book. Thanks are due to DG Research, DG Environment, DG Energy and Transport and DG JRC, the European Environment Agency, the Inter-American Development Bank, the Generalitat de Catalunya, the BBVA Foundation and the Spanish Ministry of the Environment.

The Symbol of Un-Sustainability: *Potosí City and the Cerro Rico Mountain (Photo by G. Munda)*

The city of Potosí (in Bolivia) was founded in 1545 following the discovery of silver in the Cerro Rico mountain. The veins proved so rich that the mines quickly became the world's most prolific. A popular boast was that the Spanish could have constructed a silver bridge to Spain and still we had some left to carry across it. By the end of the 18th century Potosí had grown into the largest and wealthiest city in Latin America. The price was the lives of hundred of thousands of Indian forced labourers and thousands of African slaves, as a result of accidents, mercury poisoning and silicosis pneumonia. Nowadays Potosí is a decaying place and Bolivia one of the poorest countries in the world…

Contents

Part A
Methodological Foundations and Operational Consequences of Social Multi-Criteria Evaluation (SMCE)

"It was the best of times, it was the worst of times,
it was the age of wisdom, it was the age of foolishness,
it was the epoch of belief, it was the epoch of incredulity,
it was the season of Light, it was the season of Darkness,
it was the spring of hope, it was the winter of despair,
we had everything before us, we had nothing before us,
we were all going direct to Heaven, we were all going direct
to the other way – in short, the period was so far like the
present period,
that some of its noisiest authorities insisted
on its being received, for good or for evil, in the
superlative degree of comparison only".

Charles Dickens – *A Tale of Two Cities, Signet Classic, New American Library, New York, 1960, p. 13.*

Chapter 1
Introduction

1.1 What is Multi-Criteria Evaluation?

Let us start the discussion of multi-criteria evaluation by means of a simple everyday life example. Let us imagine staying in front of a shop and admiring a selection of jackets. What is the next step? Probably entering the shop and asking for the *price*. At this point, we have two possibilities: leaving the shop because we think the price is too high or accepting the price as reasonable. In the second case, we still need to choose the jacket we want from the original set (of, for example, ten jackets). We are therefore, probably going to try on the jackets and see which *one suits us best aesthetically*. Let us assume that we are then still undecided between four of them, although we definitely do not like the other six. How do we choose between the remaining four? Perhaps at this stage we will use the criterion of *colour*. Let us imagine we are then still undecided between two jackets. We will now look at the *quality of the fabric* and we may finally choose the one with the better quality.

This is an example of the final selection of an option by using the *lexicographic model* (see Fig. 1.1). This method refers to the procedure used to order the words in a dictionary, the first letter playing the role of the first criterion, the second letter, the second criterion, and so on. To use this method, the decision-maker must give a total strict order to the criteria:

$$1 > 2 \ldots > i > \ldots > m \tag{1.1}$$

where g_l would be the most important criterion and g_m the least important. In the lexicographic model, all actions are first ranked by means of the first criterion, then if some equivalent actions exist, these are further explored by means of the second criterion, and so on. Lexicographic orders usually lead to a straightforward selection of the preferred option; however, most of the information collected on alternatives will not play a role in the decision process.

Let us discuss this example a little further and draw some conclusions. First of all, do we have any experience of a decision-making process like this? We probably do, even if not with these criteria or in this order. Thus it seems that

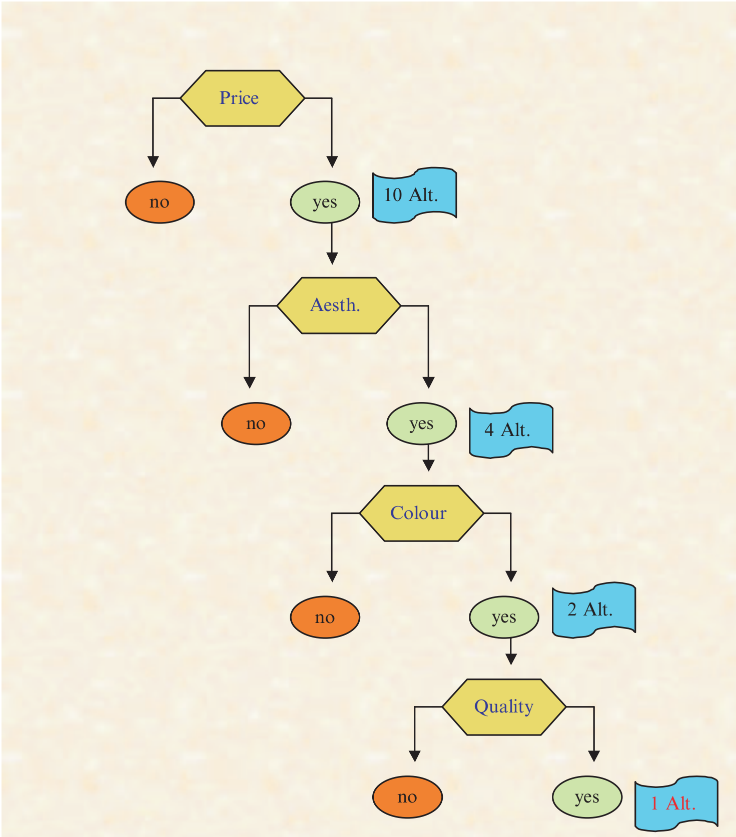


Fig. 1.1 A Lexicographic Decision Process

human beings use multi-criteria evaluation without any formal knowledge of it. We could then say that it is a behavioural assumption with a high degree of descriptive content. Secondly, does the order of criteria have any influence on the final alternative selected? Yes, of course. If we start with the criterion of quality instead of price, the jacket selected will probably be the most expensive one. This shows that when using various criteria human beings do not necessarily accord them the same weight (here I refer to the concept of weight as a coefficient of importance). In principle in the case of the lexicographic model, the first criterion alone might sometimes be enough to make the final selection (e.g. if only one jacket has the price we are willing to pay), this implies that its weight is much greater than that of any other criterion used in the selection process. This is the reason the first criterion is sometimes called the *"dictator"*.

Clearly then, the order of consideration of criteria determines their relative weights.

Thirdly, what happens in our example if we do not like the overall properties of the selected jacket? Probably we will start the process again, e.g. changing the order of criteria (i.e. their weight) or accepting to pay a higher price. Again this is something which we have probably experienced and which shows that what really matters is the *learning process* and not the final decision. This latter one is *constructed* by means of the decision process and not "discovered" as the global optimum.

Finally, does the lexicographic method allow for any compensability among the various criteria considered? Intuitively, compensability refers to the possibility that some bad criterion scores can be compensated for by other very good criterion scores. For example, an overall student evaluation may be based on the principle that a very bad score in mathematics (let us say a 2 on a 0–10 scale) can be compensated for by a 10 in literature and thus the student can pass the final evaluation. This evaluation system is completely compensatory. On the contrary, a system could be based on the principle that a student has to be "enough good" in all the subjects so that a 2 in mathematics cannot be offset by any other score, however high. This second evaluation system would be a non-compensatory (or partially compensatory) one. Compensability then requires that the various criteria scores can interact among themselves; if no interaction is possible, no compensability can exist. Since in a lexicographic method the evaluation criteria are not considered simultaneously, this procedure is completely non-compensatory.

Compensability is a very important concept when multi-criteria evaluation is applied to public policies. In fact in evaluating a project, if we consider that a 2 in mathematics could represent a very bad environmental impact and a 10 in literature a very good economic impact, it is clear that allowing or not for compensability, and to what degree, is the real issue in sustainability policies. To look for compromises then implies that a dictator cannot exist. That is, all the dimensions relevant to a policy problem have to be used simultaneously and not in a lexicographic order, since otherwise some social dimensions will a priori have a much greater weight. For example, a legislative system which accepts that a financial analysis of projects should be carried out before the evaluation of their environmental impacts, is indeed prioritizing the economic dimension with respect to the environmental.

In empirical evaluations of public projects and publicly provided goods, multi-criteria decision theory seems to be an appropriate policy tool, since it makes it possible to take into account a wide range of assessment criteria (e.g. environmental impact, distributional equity, etc.), and not simply profit maximization, as a private economic agent would do (Arrow and Raynaud, 1986; Martinez-Alier et al., 1998).

From an operational point of view, the major strength of multi-criteria methods is their ability to revolve questions characterized by various conflicting evaluations, thus allowing for an integrated assessment of the problem at hand. Multi-criteria decision theory builds on the following basic concepts[1].

The dimension is the highest hierarchical level of analysis and indicates the scope of objectives, criteria and criterion scores. For example, sustainability policy problems generally include economic, social and environmental dimensions.

The objective indicates the direction of change desired. For example, within the economic dimension GDP has to be maximized; within the social dimension social exclusion has to be minimized; within the environmental dimension CO_2 emissions have to be minimized.

The evaluation criterion[2] is the basis for evaluation in relation to a given objective (any objective may imply a number of different criteria). It is a *function* that associates each alternative with a variable indicating its desirability according to expected consequences related to the same objective, e.g. GDP, saving rate and inflation rate inside the objective "growth maximization".

The criterion score[3] is a constructed measure stemming from a process that represents, at a given point in space and time, a shared perception of a real-world state of affairs consistent with a given criterion. To give an example, when comparing two countries, within the economic dimension, one objective could be "maximization of economic growth"; the criterion might be R&D performance, the criterion score could be "number of patents per million of inhabitants". Another example: an objective connected with the social dimension might be "maximization of residential attractiveness". A possible criterion could then be "residential density". The criterion score might be the ratio of persons per hectare.

The constraint is a limit on the values that criterion scores may assume; it may or may not be stated mathematically.

A goal (synonymous with target) is something that can be either achieved or missed (e.g. reducing nitrogen pollution in a lake by at least 10%). If a goal cannot be or is unlikely to be achieved, it may be converted to an objective.

An attribute is a measure that indicates whether goals have been met or not, given a particular decision that provides a means of evaluating the levels of various objectives.

A multi-criteria method is an aggregate of all dimensions, objectives (or goals), criteria (or attributes) and criterion scores used (in the framework of composite indicators this can be considered the definition of an *index*). This implies that what formally defines a multi-criteria method is the *set of properties underlying its aggregation convention*.

---

[1]These definitions have been developed by elaborating the standard definitions in multi-criteria decision literature by means of concepts coming mainly from complex systems theory. Discussions with M. Giampietro and M. Nardo have been essential.

[2]In the framework of composite indicators a criterion is synonymous with *"individual indicator"* (see Munda and Nardo, 2007).

[3]In the framework of composite indicators, a criterion score is synonymous with *"variable"*.

Table 1.1 Example of an impact matrix

Criteria	Units	Alternatives			
		$\boldsymbol{a_1}$	$\boldsymbol{a_2}$	$\boldsymbol{a_3}$	$\boldsymbol{a_4}$
$\boldsymbol{g_1}$		$g_1(a_1)$	$g_1(a_2)$	•	$g_1(a_4)$
$\boldsymbol{g_2}$		•	•	•	•
$\boldsymbol{g_3}$		•	•	•	•
$\boldsymbol{g_4}$		•	•	•	•
$\boldsymbol{g_5}$		•	•	•	•
$\boldsymbol{g_6}$		$g_6(a_1)$	$g_6(a_2)$	•	$g_6(a_4)$

The *discrete multi-criterion problem* can be described in the following way: *A* is a finite set of *N* feasible actions (or alternatives); *M* is the number of different points of view or evaluation criteria *gm i* = 1, 2, …, *M* considered relevant in a policy problem, where the action ***a*** is evaluated to be better than action ***b*** (both belonging to the set *A*) according to the *m*-th point of view if *gm(a)* > *gm(b)*. In this way a decision problem may be represented in a tabular or matrix form. Given the sets *A* (of alternatives) and *G* (of evaluation criteria) and assuming the existence of *N* alternatives and *M* criteria, it is possible to build an *N* × *M* matrix *P* called an *evaluation or impact matrix* whose typical element *pij (i* = 1, 2, …, *M; j* = 1, 2, …, *N)* represents the evaluation of the *j*-th alternative by means of the *i*-th criterion. The impact matrix may include quantitative, qualitative or both types of information (see Table 1.1).

In general, in a multi-criterion problem, there is no solution optimizing all the criteria at the same time (the so-called *ideal or utopia solution*) and therefore *compromise solutions* have to be found. Indeed this sad truth is very consistent with the basic principle of scarcity in economics (called the *sad science* for exactly this reason).

1.2 Social Multi-Criteria Evaluation

Various authors claim that modern economics needs to expand its empirical relevance by introducing more and more realistic (and of course more complex) assumptions to its models. The issue of "distributional coalitions" has recently been considered of key importance in determining growth factors (Olson, 1982). One of the most interesting research directions in the field of public economics is the attempt to account for political constraints, interest groups and collusion effects explicitly (see e.g. Laffont, 2000). In this context, *transparency* becomes an essential feature of public policy processes (Stiglitz, 2002). Social multi-criteria evaluation (SMCE) has been purposely designed to enhance transparency, the main idea being that the results of an evaluation exercise depends on the way a given policy problem is *represented* and thus the assumptions used, the interests and values considered have to be declared (Munda, 2004). To illustrate this issue of problem representation and transparency I will present a couple of examples.

1.2.1 A Land-Use Conflict in the Netherlands

This illustrative application concerns a study of a public choice problem in the Netherlands (the southern part of Limburg). The problem can be briefly outlined as follows (for more information see Munda et al., 1994b).

A company enjoying almost absolute dominance in the Dutch cement industry has a concession to extract marl on one of the hills in south Limburg, but this concession may expire in the near future; therefore alternative areas have to be explored. Of the possible new locations, the most appropriate is the Plateau van Margraten; this is a rather flat area used for agriculture and for some recreation. It has a unique physical structure, although it is characteristic of the landscape of the region. Zoning of this area for marl mining would fundamentally affect its social and ecological value; on the other hand, if the authorities refused to grant permission for marl mining to the company, this would lead to an almost total collapse of the national cement industry and to serious unemployment effects in an economically weak region. This situation clearly demonstrates the sharp opposition between *environmental* and *economic* interests. A first meaningful step toward an evaluation analysis for this land-use problem is to identify a set of feasible and relevant options. These are:

(1) Implementation of the original plans of the company (i.e. a concession for the total area). This would guarantee the future position of the national cement industry and also favours employment and welfare in the region. Agriculture would suffer some negative impacts, while the negative social impacts (for recreation, etc.) would be rather strong. Finally, the environmental damage would be very great.

Clearly, this would be the best option for *interest groups supporting the cement company*.

(2) The use of an alternative area (the Rasberg area, in the same region) for marl mining. But this area is much smaller and the physical condition of the soil would hamper a profitable cement production at current prices. On the other hand, the ecological damage would be less serious.

This option could be interpreted as a *compromise between the authorities and the company* (in fact a concession is given, but the location is decided by the political authorities and not by the company).

(3) The provision of a concession for one half of the area (Plateau van Margraten). This would lead to fewer agricultural losses, while the environmental damage

would also be less severe. The economic impacts would be less favourable than those of the first option.

This is a possible basis for compromise. However, it *could be very dangerous for the political authorities*, since it might maximize the conflict precisely by making nobody happy.

(4) A concession from political authorities for marl mining in the present area. This would be only a short-term solution (since it is not possible from a technical point of view, to continue the extraction of marl for an indefinite period), which is a less attractive option from an economic perspective.

Clearly, this option is not consistent with the others with respect to the time horizon. It reflects a *typical attitude of politicians* all over the world, i.e. the inclination to pass the problem on to somebody else (the succeeding government).

(5) Importing marl from the Plateau van Vroenhoven, an area in Belgium.

This solution might be attractive from *a national standpoint*. But if we take into account the global environmental consequences (the environmental impact of extraction in Belgium and transport of marl) it is clearly the worst option from an environmental point of view. This option also shows the importance of the *hierarchical level* we use to describe a problem (region, country, European Union, etc.).

(6) A restructuring of the company so that it becomes a trade and research organization for cement instead of a production unit. This would lead to a certain loss of employment, while the future need for such an organization is unclear.

This option might be attractive for *anyone without an economic interest* in the company.

(7) A closedown of all productive activities of the company.

This might be *favourable from the viewpoint of environmentalists* and people who use the area for recreation, but would lead to serious economic problems for the region. These options are to be judged on the basis of various evaluation criteria. Three main groups of criteria can be distinguished, viz. economic, social and environmental. These three classes can be subdivided into various components.

(A) *Economic criteria*

1. Employment in agriculture
2. Employment in the cement industry (including marl mining)
3. Agricultural production
4. National production of marl
5. Value added in the cement industry

(B) *Social criteria*

6. Residential attractiveness
7. Recreational attractiveness (daily)

8. Tourism attractiveness
9. Congestion created in transportation infrastructure

(C) *Environmental criteria*

10. Quality of geo-physical structure
11. Diversity and scarcity of eco- and bio-components
12. Consistency with existing landscape components
13. Consistency with existing cultural–historical components.

Of course, the options and criteria described belong to only one of the possible approaches to the problem. They are the result of a decision process developed on that particular occasion. However, one should note that the degree of *transparency is really very high*. It is easy to identify the interests represented by each option.

One could decide to measure the environmental impact of the various policy options in the Netherlands alone, or in both the Netherlands and Belgium: according to the choice made, the "constructed solutions" would be different. In any case, it is clear that such a choice implies a *value judgement*, i.e. the implementation of a weak or strong sustainability philosophy.

1.2.2 Assessing Urban Sustainability

Let us consider an example involving four cities, two belonging to highly industrialized countries (Amsterdam and New York) and two belonging to transitional economies (Budapest and Moscow). The indicators used are typical of the literature on urban sustainability (see e.g. Barbiroli, 1993 or the Urban Indicator Programme). The profiles (i.e. the score of each city according to each indicator) of these four cities are described in Table 1.2.

Table 1.2 Impact matrix for the four chosen cities according to the selected indicators

Criteria \ Alternatives	Budapest	Moscow	Amsterdam	New York
Houses owned (%)	50.5	40.2	2.2	10.3
Residential density (pers./hectare)	123.3	225.2	152.1	72
Use of private car(%)	31.1	10	60	32.5
Mean travel time to work (minutes)	40	62	22	36.5
Solid waste generated per capita (t/year)	0.2	0.29	0.4	0.61
City product per person (Uss/year)	4750	5100	28251	30952
Income disparity (Q5/Q!)	9.19	7.61	5.25	14.81
Households below poverity line (%)	36.6	15	20.5	16.3
Crime rate per 1000 (theft)	39.4	4.3	144.05	56.7

A standard approach is to rank these cities by constructing a composite indicator. A typical composite indicator, *I*, is built as follows (OECD, 2003, p. 5):

$$I = \sum_{i=1}^{N} w_i x_i \tag{1.2}$$

where x_i is a normalised variable and a weight attached to x_i, with $\sum_{i=1}^{N} w_i = 1$ and $0 \le w_i \le 1$, $i = 1, 2, \ldots, N$.

It is clear that from a mathematical point of view a composite indicator entails a weighted linear aggregation rule applied to a set of variables. The main technical steps needed for its construction are the following:

1. Standardization of the variables to allow comparison without scale effect.
2. Weighted summation of these variables.

The standardization step is a very delicate one. The main sources of uncertainty and imprecise assessment are:

- *The normalization technique* used for the different measurement units involved.
- *The scale adjustment* used, for example population or GDP of each country considered.
- *The common measurement unit* used (money, energy, space, etc.).

Several techniques can be used to standardize variables (Saisana and Tarantola, 2002; OECD, 2003). However, although each normalisation technique entails different absolute values, the ranking produced remains constant. In our example, the "*distance from the best and worst performers*" technique is applied, where positioning is in relation to the global maximum and minimum and the index takes values between 0 (laggard) and 100 (leader):

$$100\left(\frac{\text{actual value - minimum value}}{\text{maximum value - minimum value}}\right) \tag{1.3}$$

By applying (1.3) to the values contained in Table 1.2, the results presented in Table 1.3 are obtained. The indicators "houses owned" and "city product per person" have to be maximized. All others have to be minimized. To apply (1.2) it is thus necessary to transform the indicator scores of these indicators by using the simple equation (100 – normalized indicator score).

By applying this transformation to the values contained in Table 1.3, the results presented in Table 1.4 are obtained.

By applying (1.2) to the values contained in Table 1.4, the following results are obtained:

Budapest = 512.986
Moscow = 533.373
Amsterdam = 463.169
New York = 492.052

Table 1.3 Normalized impact matrix

100	78.674	0	16.770
33.485	100	52.28	0
42.2	0	100	45
45	100	0	36.25
0	21.95	48.78	100
0	1.335	89.691	100
41.213	24.686	0	100
100	0	25.462	6.018
25.116	0	100	37.495

Table 1.4 Normalized impact matrix accounting for minimization objectives

100	78.674	0	16.770
66.515	0	47.72	100
57.8	100	0	55
55	0	100	63.75
100	78.05	51.22	0
0	1.335	89.691	100
58.787	75.314	100	0
0	100	74.538	93.982
74.884	100	0	62.505

Thus the final ranking puts Amsterdam into bottom position (worst than all other cities considered); Moscow is in top position, Budapest ranks second and New York third.

At this point two questions need to be answered: *Where are these (somewhat surprising) results coming from and what do they mean?*

Let us start with the first question. The results obtained depend on:

1. *Quality of the information available* (In our case, for example, the data values for Amsterdam on the use of private cars and on criminality are suspiciously high, while criminality in Moscow and residential density in New York are suspiciously low).
2. *Indicators chosen* (i.e. which representation of reality we are using, and therefore whose interests are taken into account).
3. *Direction of each indicator* (i.e. the bigger the better or vice versa, in our example, the principle is that house ownership should be maximized, but this could be quite debatable and culturally dependent).
4. *Relative importance of these indicators* (in our case all the indicators are considered to have the same importance, i.e. no weighting coefficient is used).
5. *Ranking method used* (in this case the linear aggregation rule).

All these uncertainties have to be taken into account when we claim that any given city is "better" than another. At this stage, it also seems clear why it is claimed in multi-criteria evaluation that what is really important is the "*decision process*" and not the final solution, since this solution has a value only as a product of the decision process and is not an "ultimate Truth" (in Herbert Simon's words, we could say that we should move from "*substantive to procedural rationality*").

Historically the first stage of the development of multi-criteria decision theory was characterized by the so-called methodological principle of *multi-criteria decision making* (MCDM), the main aim of which is to elicit clear subjective preferences from a mythical decision-maker (DM) and then try to solve a well-structured mathematical decision problem by means of a more or less sophisticated algorithm. In this way a multi-criterion problem can still be presented in the form of a classical optimization problem (Keeney and Raiffa, 1976). The limitations of the classical concept of an optimum solution and the consequential importance of the decision process was emphasized in the context of decision sciences by authors such as H. Simon and B. Roy.

According to Simon (1976), a distinction must be made between the general notion of rationality as an adaptation of available means to ends, and the various theories and models based on a rationality which is either substantive or procedural. This terminology can be used to distinguish between the rationality of a decision considered independently of the manner in which it is made (in the case of substantive rationality, the rationality of evaluation refers exclusively to the results of the choice) and the rationality of a decision in terms of the manner in which it is made (in the case of procedural rationality, the rationality of evaluation refers to the decision-making process itself). "A body of theory for procedural rationality is consistent with a world in which human beings continue to think and continue to invent: a theory of substantive rationality is not."

Roy (1985) states that in general it is impossible to say that a decision is a good or bad one by referring only to a mathematical model: all aspects of a decision process which leads to a given decision also contribute to its quality and success. Thus, it becomes impossible to establish the validity of a procedure either on a notion of *approximation* (i.e. discovering pre-existing truths) or on a mathematical property of *convergence* (i.e. does the decision automatically lead, in a finite number of steps, to the optimum a^*?). The final solution is more like a *"creation"* than a discovery. In *Multiple-Criteria Decision Aid* (MCDA) (Roy, 1985), the principal aim is not to discover a solution, but to construct or create something which is viewed as liable to help "an actor taking part in a decision process either to shape, and/or to argue, and/or to transform his preferences, or to make a decision in conformity with his goals" (Roy, 1990) (*constructive or creative approach*).

The need for public participation has been increasingly recognized in a multi-criteria decision-aid framework. Two recent proposals have been participatory multi-criteria evaluation (Banville et al., 1998) and social-multi-criteria evaluation (Munda, 2004). Social multi-criteria evaluation accords with the need to extend MCDA by incorporating the notion of the stakeholder; for this reason a social multi-criteria process must be as *participative* and as *transparent* as possible; although, it is further argued that participation is a *necessary* but not a *sufficient* condition. This is the main reason why the concept of "Social Multi-criteria Evaluation" (SMCE) is proposed in place of "Participative Multi-criteria Evaluation" (PMCE) or "Stakeholder Multi-criteria Decision Aid" (SMCDA).

In my opinion, one should not forget that even a participatory policy process could always be influenced by strong value judgements. Do all the social actors have the same importance (i.e. weight)? Should a socially desirable ranking be obtained on the grounds of the majority principle? Should some veto power be allowed to minorities? Are income distribution effects important? In the light of these questions, the objective of the present book is to discuss in depth the methodological foundations, the mathematical axiomatization and the operational consequences of SMCE in the context of public choice in sustainability issues.

1.3 Social Multi-Criteria Evaluation and Sustainability Issues

In the eighties, the awareness of actual and potential conflicts between economic growth and the environment led to the concept of "*sustainable development*". Since then, all governments have declared, and continuously proclaim, their willingness to pursue economic growth under the flag of sustainable development although often development and sustainability are contradictory terms. In the last decade, there have been various attempts to develop theoretical definitions and systems for the assessment of sustainability, but so far no consensus emerged on pros and cons of any definition and its implementation (see e.g. Barbier and Markandya, 1990; Horwarth and Norgaard, 1990, 1992; Chichilnisky, 1996; Musu and Siniscalco, 1996; Pearce et al., 1996 Munda, 1997a; Faucheux and O'Connor, 1998).

The concept of sustainable development has a wide appeal, partly because, in contrast with the "zero growth" idea by Daly (1977), it does not set economic growth and environmental preservation in sharp opposition. Rather, sustainable development carries the ideal, of a harmonization or simultaneous realization of economic growth and environmental concerns. For example, Barbier (1987, p. 103) writes that sustainable development implies: "*to maximize simultaneously*[4] the biological system goals (genetic diversity, resilience, biological productivity), economic system goals (satisfaction of basic needs, enhancement of equity, increasing useful goods and services), and social system goals (cultural diversity, institutional sustainability, social justice, participation)". This definition correctly points out that sustainable development is a *multidimensional* concept, but as everyday life teaches us, it is generally impossible to maximize different objectives at the same time, therefore as formalized by multi-criteria decision theory, compromise solutions must be found.

Let us try to clarify some fundamental points of the concept of "sustainable development". In economics by "development" is meant "the *set of changes* in the economic, social, institutional and political structure needed to implement the transition from a pre-capitalistic economy based on agriculture, to an industrial *capitalistic* economy" (Bresso, 1993). Such a definition of development has two main implications:

---

[4]Emphasis added.

1. The changes needed are not only quantitative (like growth in gross domestic product), but also qualitative (social, institutional and political).
2. There is only one possible model of development, i.e. that of western industrialized countries. This implies that the concept of development is viewed as a process of cultural fusion towards the *best* knowledge, the *best* set of values, the *best* organization and the *best* set of technologies (and that these are all western …).

Adding the term "sustainable" to the "set of changes" (the first point) means adding an ethical dimension to development. The issue of *distributional equity*, both within the same generation (intra-generational equity, e.g. the North-South conflict) and between different generations (inter-generational equity) becomes crucial (Munda, 1997a). Going further, a legitimate question could be raised: sustainable development *of what and for whom*? (Allen et al., 2002). Norgaard (1994, p. 11) writes: consumers want consumption sustained, workers want jobs sustained. Capitalists and socialists have their "isms", while aristocrats and technocrats have their "cracies".

Martinez-Alier and O'Connor (1996) have proposed the concept of ecological distribution to synthesize sustainability conflicts. The concept of *ecological distribution* refers to the social, spatial, and temporal asymmetries or inequalities in the use by humans of environmental resources and services. Thus, the territorial asymmetries between SO_2 emissions and the burdens of acid rain are an example of *spatial ecological distribution*. The inter-generational inequalities between the benefits of nuclear energy and the burdens of radioactive waste are an example of *temporal ecological distribution*. In the USA, "environmental racism", meaning locating polluting industries or toxic waste disposal sites in areas where poor people live, is an example of *social ecological distribution*. We can then conclude that sustainability management and planning is essentially a question of *conflict analysis*. As a tool for conflict management, multi-criteria evaluation has demonstrated its usefulness in many environmental policy and management problems (see e.g. Romero and Rehman, 1989; Nijkamp et al., 1990; Janssen, 1992; Munda, 1995; Beinat and Nijkamp, 1998; Munda et al., 1998; Ringius et al., 1998; Janssen and Munda, 1999; Hayashi, 2000; Bell et al., 2001). As a consequence the use of multi-criteria decision theory seems very relevant for tackling sustainability conflicts.

The second characteristic of the term "development" refers to the western industrialized production system as a symbol of a successful development process. However, serious environmental problems may proceed from this vision. For example, according to actual social values in western countries, to have a car per two/three persons could be considered a reasonable objective in less developed countries. This would imply a number of cars ten times greater than the existent one, with possible consequences for global warming, reserves of petroleum, loss of agricultural land and noise. The contradiction between the terms "development" and "sustainable" may not be reconcilable unless alternative models of development are considered.

One such model is offered by the so-called *co-evolutionary paradigm*. According to this view of social evolution, borrowed from biology (Ehrlich and Raven, 1964),

The maket Economy

Science

Civil Society

Policy Makers

Fig. 1.2 Main actors of a sustainability policy process

there is a constant and active *interaction* between organisms and their environment. Organisms are not simply the results but they are also the causes of their own environments (Gowdy, 1994; Norgaard, 1994). Economic development can be viewed as a process of adaptation to a changing environment while being itself a source of environmental change. In real societies, "people survive to a large extent as members of groups. Group success depends on culture: the system of values, beliefs, artefacts, and art forms which sustain social organization and rationalize action. Values and beliefs which fit the ecosystem survive and multiply; less fit ones eventually disappear. And thus cultural traits are selected much like genetic traits. At the same time, cultural values and beliefs influence how people interact with their ecosystem and apply selective pressure on species. Not only have people and their environment coevolved, but social systems and environmental systems have coevolved" (Norgaard, 1994, p. 41). From the co-evolutionary paradigm the following lessons can be learned:

(1) A priori, different models of co-evolution are possible, and then no unique optimal development path exists. *The spatial dimension* is a key feature of sustainable development.
(2) Respect for cultural diversity is of fundamental importance. In environmental management local knowledge and expertise (being the result of a long co-evolutionary process) are sometimes more useful than experts opinions. *Social participation* is then essential for successful sustainability policies.

From this brief discussion the following conclusions can be drawn (see Fig. 1.2):

1. A proper evaluation of sustainability options needs to deal with a plurality of legitimate values and interests found in a society. From a societal point of view,

economic optimization cannot be the only evaluation criterion. As is well known, not all goods have a market price, or this price is often too low (*market failures*). Environmental and distributional consequences (intra/inter-generational and for non-humans) must also be taken into account. In this framework multi-criteria evaluation is a very consistent approach.

2. If from a sustainability point of view, it is accepted that society as a whole has an indefinite lifespan, a much longer time horizon than is normally used on the market is required. A contradiction then arises: politicians usually have a very short time horizon (often four–five years depending on the electoral system) and this has the effect that sustainability is rarely among their priorities (thereby causing a *government failure* (for an overview of different perspectives on the role of governments in the economic sphere see e.g. Buchanan and Musgrave, 1999). For this reason I think that evaluation of public projects should take into account the entire "*civil society*" (including ethical concerns about *future generations*) and not only mythical benevolent policy-makers. This is why I am developing the concept of "social" multi-criteria evaluation, the main objective of which is to integrate scientific knowledge with social participation in the framework of sustainability public choice.

Chapter 2
Dealing with a Complex World: Multiple Dimensions, Values and Scales

2.1 Complexity and Post-Normal Science

The world is characterized by deep *complexity*. This apparently unremarkable observation has important implications for the manner in which policy problems are represented and decision-making is framed. One may decide to adopt a reductionistic approach by tackling only one of the many possible dimensions or one may try to deal with the complexity of the real-world. This book adopts the latter approach. My firm conviction is that any representation of a complex system reflects only a sub-set of its possible representations. *A system is complex when the relevant aspects of a particular problem cannot be captured using a single perspective* (Rosen, 1977; O'Connor et al., 1996; Funtowicz et al., 1999).

To make things more difficult, systems involving humans are *reflexively* complex. Reflexive systems display two peculiar characteristics: "*awareness*" and "*purpose*", both requiring an additional "jump" in describing complexity. The presence of self-consciousness and purpose (*reflexivity*) means that these systems can continuously add new relevant qualities/attributes to be considered when explaining, describing or forecasting their behaviour (i.e. human systems are learning systems).

Moreover, the existence of *different levels and scales* on which a hierarchical system can be analysed implies the unavoidable existence of non-equivalent descriptions of it (Giampietro, 1994, 2003). Even a simple "objective" description of a geographical orientation is impossible without taking an arbitrary subjective decision on the relevant system scale. In fact, as shown in Fig. 2.1, the same geographical place, for example, in the USA, may be considered to be in the north, south, east or west according to the scale chosen as a reference point (the whole USA, a single state, etc.)[1](Giampietro and Mayumi, 2000a,b). Therefore, the problem of *multiple identities* in complex systems cannot be interpreted solely in terms of *epistemologi-*

[1] These multiple-identity/multiple-scale systems can be defined as "*Learning Holarchies*". A "*holon*" is a whole made of smaller parts (e.g. a human being made of organs, tissues, cells, atoms) which at the same time forms a part of a larger whole (an individual human being is a part of a household, a community, a country, the global economy) (Koestler, 1969).

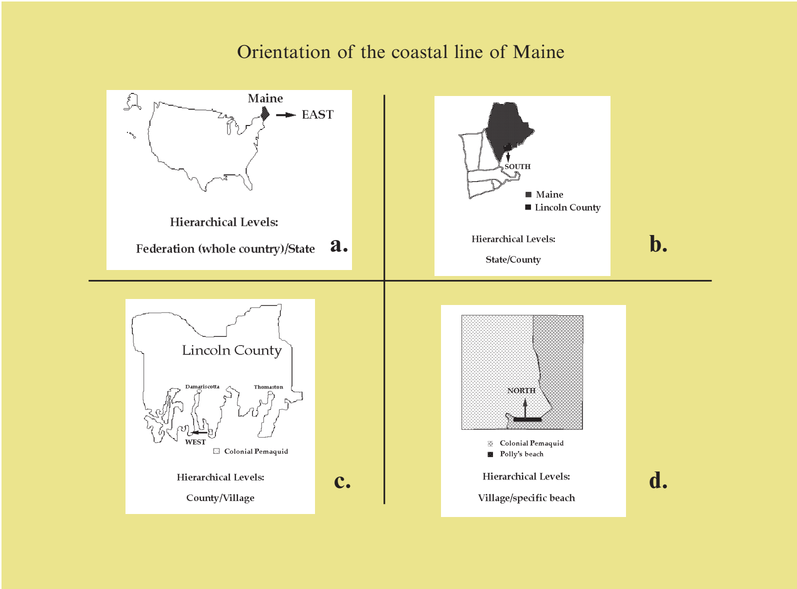


Fig. 2.1 An example of multiple identities according to the system scale (Source: PowerPoint Presentation by M. Giampietro)

cal plurality (non-equivalent observers), but also necessarily in terms of *ontological characteristics* of the observed system (non-equivalent observations).

The implications of scale for multi-criteria evaluation are very important. For example, in generating evaluation criteria (e.g., when evaluating the impacts building a skiing infrastructure in a mountain region, who are the relevant social actors? The inhabitants of the mountain region, the potential users in urban areas or even the ecological preservationists all around the world?), in computing impact scores (e.g., should a contamination indicator be computed locally, or at a larger scale? The use of hydrogen cars in cities is clearly good at local level, but not necessarily at global level, where the emissions depend on the technology by which hydrogen is produced –since hydrogen is an energy carrier and not an energy source.).

A consequence of the extreme subjectivities involved is that in any normative exercise connected with a social decision problem, one has to choose an operational definition of "*value*", in spite of the fact that social actors with different interests, cultural identities and goals all have different definitions of "value" (O'Neill, 1993). That is, to reach a ranking of policy options, there is a prior need to decide *what is important* for different social actors as well as *what is relevant* for the representation of the real-world entity described in the model.

Sustainability policies deal with reflexive phenomena. Because an effective assessment, in order to be realistic, should consider not merely the measurable and contrastable dimensions of the simple parts of the system, which even if complicated, may be technically simulated, it should also deal with the higher dimensions of the system: those dimensions in which power relations, hidden interests, social participation, cultural constraints, and other "soft" values become relevant, and unavoidable variables that strongly, but not deterministically, affect the possible outcomes of the strategies to be adopted.

No mathematical model, even if legitimate in its own terms, can be sufficient for a complete analysis of the reflexive properties of a real-world problem. These reflexive properties include the human dimensions of e.g. the ecological change and the transformations of human perceptions along the way. The *learning process* that takes place while analyzing the issue and defining policies will itself influence perceptions and alter significantly the decisional space in which alternative strategies are chosen. At the other end, *institutional and cultural representations* of the same system, while also legitimate, are on their own insufficient to define what should be done in any particular case.

The various dimensions are not totally disjointed; thus the institutional perspective can be a basis for the study of the social relations of the scientific processes. To take any particular dimension as the true, real or total picture amounts to *reductionism*, whether physical or sociological. As a consequence, any attempt to fit the real world into a closed model leads to a simplification, which does violence to the description of reality. In most cases the dimensions sacrificed are precisely the reflexive properties of the systems. These characterize the problem in a fundamental way but are difficult to identify or measure.

In general, these concerns were not considered very relevant by scientific research as long as time was considered an infinite resource. On the other hand, the new nature of the problems faced in this third millennium (e.g. mad-cow disease, genetically modified organisms, etc.) implies that, when dealing with problems that may have long term consequences, we are confronting issues "*where facts are uncertain, values in dispute, stakes high and decisions urgent*" (Funtowicz and Ravetz, 1991, 1994).

Scientists cannot therefore provide any useful input without interacting with the rest of society while the rest of the society cannot make any sound decision without interacting with scientists. That is, the question of "how to improve the quality of a policy process" must be put, rather quickly, on the agenda of "scientists", "decision-makers" and indeed of society as a whole. This extension of the "peer community" is essential for maintaining the quality of the process of decision-making when dealing with reflexive complex systems. In relation to this objective Funtowicz and Ravetz have developed a new epistemological framework called "*Post-Normal Science*", with which it is possible to deal better with two crucial aspects of science in the policy domain: *uncertainty* and *value conflict*. The term "post-normal" signals a divergence from the puzzle-solving exercises of normal science, in the Kuhnian sense (Kuhn, 1962).

Post-Normal Science can be characterized in relation to other, complementary, scientific strategies according to the diagram in Fig. 2.2, which is based on two axes: "*systems uncertainties*" and "*decision stakes*". When both uncertainty and stakes are low, we are in the realm of "*normal academic science*",[2] where it is safe

[2] Funtowicz and Ravetz use the term "applied science". I prefer the more general category of "academic science".

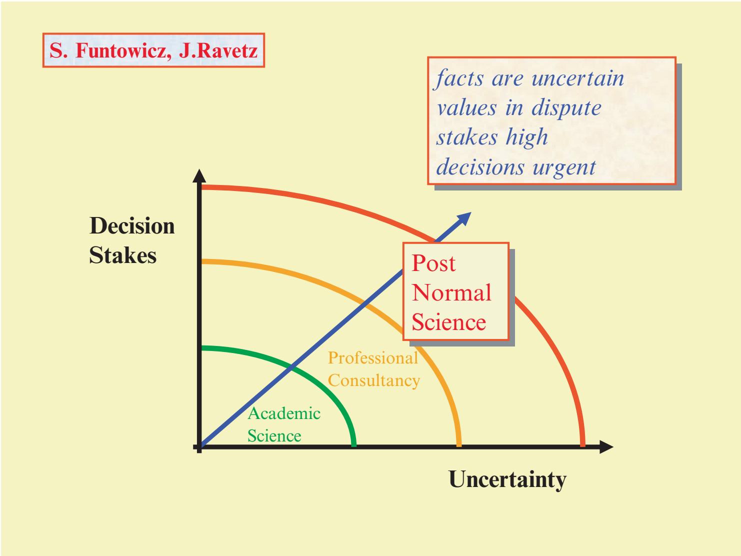


Fig. 2.2 Graphical representation of post-normal science

to rely on "codified expertise". When either uncertainty or stakes are in the medium range, then the application of routine techniques and standardized and generalized knowledge is no longer enough. In these cases, skill, judgement, and sometimes even courage, are required to adjust the "general knowledge" available to the "special situation". Funtowicz and Ravetz call this "*professional consultancy*", with the examples of the surgeon or the senior engineer facing a critical situation. Finally we arrive at cases, in which conclusions are not completely determined by scientific facts; inferences will (naturally and legitimately) be conditioned by the values held by the agents. When the stakes are very high (as when an institution is seriously threatened by a policy) then a defensive tactic will involve challenging every step of a scientific argument (this applies even to those cases in which system uncertainties are actually small). Such a tactic should be considered wrong only when is conducted covertly, as by scientists who present themselves as impartial judges when, in reality, they are actually committed advocates of a particular view. When legitimate contrasting views are openly used to challenge scientific arguments, we are in the realm of "*Post-Normal Science*".

There are now many initiatives constantly increasing in number and significance, for involving wider circles of people in decision-making and implementation on environmental issues. Examples of "Post-Normal Science" are to be found wherever local communities engage in scientific research and dialogue on the state of their personal and environmental health; one important case is the "popular epidemiology" movement connected to the Environmental Justice movement in the United States (Novotny, 1994). A number of grass-roots movements struggled for years against all official agencies, including academics, to try to establish the fact that their communities were being harmed. The most famous was a case where the authorities were eventually forced to admit that a collection of symptoms was the manifestation of a real condition, thereafter known by the town's name as Lyme Disease.

2.2 The Incommensurability Principle

The previous discussion can be summarized by using the philosophical concept of *weak comparability* (O'Neill, 1993; Martinez-Alier et al., 1998). From a philosophical perspective, it is possible to distinguish between the concepts of *strong comparability* (there exists a single comparative term by which all different actions can be ranked) implying *strong commensurability* (a common measure of the various consequences of an action based on a cardinal scale of measurement) or *weak commensurability* (a common measure based on an ordinal scale of measurement), and *weak comparability* (irreducible value conflict is unavoidable but compatible with rational choice employing, for example, multi-criteria evaluation).

In terms of formal logic, the difference between strong and weak comparability, and one defence of weak comparability, can be expressed in terms of Geach's distinction between *attributive and predicative adjectives* (Geach, 1967). An adjective *A* is predicative if it passes the two following logical tests:

(1) If *x* is *AY*, then *x* is *A* and *x* is *Y*;
(2) If *x* is *AY* and all *Y*'s are *Z*'s, then *x* is *AZ*.

Adjectives that fail such tests are attributive. Geach claims that "*good*" is an attributive adjective. In many of its uses it clearly fails (2): "*X is a good economist, all economists are persons, and therefore X is a good person*" is an invalid argument. The fact that a comparative holds in one range of objects does not entail that it holds in the wider range. Given a claim that "*X is better than Y*" a proper response is "*X is better what than Y?*" Similar points can be made about the adjective "*valuable*" and "*is more valuable than*". If evaluative adjectives like "good" and "valuable" are attributive in standard uses, it follows that their comparative forms have a limited range. That does not however preclude the possibility of rational choices between objects that do not fall into the range of a single comparative. Weak comparability is compatible with the existence of such limited ranges.

It is in terms of such descriptions that evaluation takes place. A location is not evaluated as good or bad as such, but rather, as good, bad, beautiful or ugly *in relation to different descriptions*. It can be at one and the same time a "*good W*" and a "*bad X*", a "*beautiful Y*" and an "*ugly Z*". The use of these value terms in such contexts is attributive, not predicative. Evaluation of objects relative to different descriptions invokes not just different practices and perspectives, but also the different criteria and standards for evaluation associated with these. It presupposes value-pluralism. An appeal to different standards often results in conflicting appraisal of an object: as noted above, an object can have considerable worth as a *U, V*, and *W*, but little as an *X, Y* and *Z*.

In conclusion, weak comparability implies *incommensurability* i.e. there is an irreducible value conflict when deciding what common comparative term should be used to rank alternative actions. Remembering that the presence of multiple-identities in complex systems can be explained in terms of the *epistemological plurality* and *ontological characteristics* of the observed system, I think that it is possible to further distinguish the concepts of social incommensurability and technical incommensurability (Munda, 2004). *Social incommensurability* can be derived from the concepts of reflexive complexity and Post Normal Science and refers to the existence of a multiplicity of legitimate values in society, that is, in one word, to democracy. *Technical incommensurability* comes from the multidimensional nature of complexity and refers to the issue of representation of multiple identities in descriptive models.

At this point, if we accept that real-world systems are multi-dimensional in nature, we also have to accept that the evaluation of public plans or projects has to be based on procedures that explicitly require the integration of a broad set of various and conflicting points of view. Consequently, multi-criteria evaluation is in principle an appropriate policy framework. It is interesting to note that analytic philosophy, theories of complexity, post-normal science, recent theories of rationality and modern public economics lead with different trajectories to the same conclusion, i.e. policy problems can be operationalized through a consistent multi-criteria framework.

The arguments developed in this section imply that there could be at least two different compromise solutions: a *social compromise solution* originating in value conflicts and a *technical compromise solution* for conflicting non-equivalent representations of the same policy options.

At this stage, the basic questions to be addressed are:

1. How is it possible to deal with technical incommensurability?
2. How can we deal with the issue of social incommensurability?
3. Which are the main consequences of technical and social incommensurability in a SMCE framework?

Answering these questions will be the subject of the rest of this book.

2.3 Reductionism "Must" Be Avoided

2.3.1 An Example of Ecological Reductionism: Carrying Capacity and the Ecological Footprint

Sustainable development clearly has a global dimension. However, the existence of interactions between local and global processes is also increasingly recognized. In particular, cities are open systems which impact on other physical areas and on the earth as a whole.

In the European context, for example, the reinforced focus on the city seems warranted as European countries face into a period of dramatic restructuring and transition (Nijkamp and Perrels, 1994; Cocossis and Nijkamp, 1995). The aim to make Europe economically more competitive has to be reconciled with the issue of environmental and cultural sustainability. At the institutional level for instance, EUROSTAT proposes a set of urban pressure indicators to deal with the urban sustainability issue (European Commission, 1996).[3]

It might be asked, for what reason so many different indicators should be introduced, when a unique physical *index* of human impact on the environment could be constructed simply from the concept of "carrying capacity". *Carrying capacity*, as defined in ecology, is the maximum population of a given species (frogs in a lake, for instance) that can be supported indefinitely in that given territory, without spoiling its resource base. Begon et al. (1996) clearly state that, even for animals, carrying capacity is "*an idealized concept not to be taken literally in practice*".

Authors with a background in biology and with an interest in population growth, such as Paul Ehrlich and his collaborators, have over the years become aware of the shortcomings of the idea of carrying capacity as applied to humans. This is why they proposed the formulation ***I = PAT, I*** standing for human impact on the environment, ***P*** for human population, ***A*** for affluence, and ***T*** for technology.

The definition of carrying capacity is irrelevant for humans for several reasons. First, the human ability to establish large differences in the *exosomatic use of energy* and materials points to a crucial question: at which level of consumption should the maximum population be established? Second, *human technologies* play a paramount role. To give an example transport is essential for determining urban carrying capacity because it influences the number of people which can enjoy a reasonable quality of urban life. Third, the *territories* occupied by humans are not given. We compete for them with other species as well as with other humans, whereby the concept of territory is socially and politically constructed. There is yet another reason why the notion of carrying capacity is not directly applicable to humans in any particular territory: *trade* may in fact be interpreted as the appropriation of the carrying capacity of other territories.

Urban growth rests on a trade-off between agglomeration economies (notably economies of scale and scope including higher wages) and diseconomies (e.g. population density and environmental decay). It is likely that environmental quality

---

[3] These indicators are: population density per area, land consumption, roads and parking areas, mono-functional areas, derelict areas, inhabitants per green area, accessibility of green areas, emissions of CO2, emissions of SO2 and Nox, emissions of VOC, emissions of PM10, emissions of lead, water consumption per capita, COD/BOD through (non-treated) waste water, non-treated waste water, non-treated waste water discharges to urban surface waters, soil contamination, municipal waste per capita, non-recycled municipal waste, household hazardous waste, energy consumption, share of private car transport, registered motor vehicles, traffic accidents with victims (injured and/or dead), mileage of commuters, people endangered by noise emissions, noise emissions of industry, noise levels of vehicle fleet.

problems may become more severe with urban size; however, factors such as land use, transportation system and spatial layout of a city are also critical factors in determining the "urban environmental carrying capacity".

Another indicator related to the idea of urban carrying capacity is the *ecological footprint index*. Ecological footprint overcomes some of the difficulties of traditional carrying capacity simply by inverting the usual carrying capacity ratio. In short, the ecological footprint measures land area required per person (or population), rather than population per unit area (Wackernagel and Rees, 1995; Folke et al., 1996).

The ecological footprint starts from the *assumption that every category of energy and material consumption and waste discharge requires the productive or absorptive capacity of a finite area of land or water*. If one sums up the land requirements for all categories of consumption and waste discharge of a defined population, the total area represents the ecological footprint of that population whether or not this area coincides with the population's home region.

More precisely, the ecological footprint of a specified population or economy can be defined as the area of ecologically productive land (and water) that would be required on a continuous basis:

(a) To provide all the energy/material resources consumed
(b) To absorb all the waste discharged, by a given population in a given area

From an operational point of view, the main categories of land use for the calculation of the ecological footprint would be the following:

1. Crop and grazing land required to produce the current diet (the sea area could also be included)
2. Land for wood plantations for timber and paper
3. Land occupied or built-over, as urban land
4. Land needed to absorb CO_2 emissions through photosynthesis, or alternatively land required to produce the ethanol equivalent to current fossil energy consumption

In Rees' hometown of Vancouver, the respective figures for these four items, per person, would be 1, 0.6, 0.2, and 2.3 ha (of middle-aged Northern temperate forest), i.e. over 4 hectares per person. One should note that only CO_2 is translated into a land requirement, and not other wastes, such as domestic waste, or other greenhouse gases, or radioactive waste; this is so because of difficulties of computation. The water catchment area and the waste water disposal area are not included either.

Of course, when considering the urban population, it is particularly important to acknowledge the existence of physical constraints on matter and energy flows, which are determined by the particular type of social structure. The *structure of a society* has huge relevance in determining the consequential ecological footprint for the same unit of human mass sustained, energy consumed or waste generated. Let us consider the case of food supply. A kilogram of grain consumed per person can have a cost of 2,000 kcal (in a poor society) or 35,000 kcal (in a rich society) according to the characteristics of the society. A rich society can be defined by the average need to produce food using only 5% of the available work force in agriculture (to produce grain at a throughput of 700 kg of grain per hour of labour). On the contrary, the situation of a subsistence society is far more "energy efficient". This is caused by a very low productivity of labour e.g. 10 kg of grain per hour of labour (the population is mainly composed of poor farmers). The same applies to the amount of land available (Giampietro, 1997).

What I want to emphasize here is the *problem of aggregation* (i.e. the somewhat mysterious convention that one needs to transform all the dimensions of ecological sustainability in a common measurement unit in terms of space connected to ecological footprint) and the necessary *reductionism* implied by the use of this index.

From a *land-use policy perspective*, the urban management suggestions arising from the computations of the ecological footprint can be misleading. Given that ecological footprint considers the land used to produce the current diet, this could imply an incentive towards intensive agricultural production systems. These systems will reduce the virtual space occupied by a city but at the same time will imply the use of much more energy and loss of biodiversity, due to the use of fertilizers and pesticides and the introduction of exotic species. It is true that to a certain extent these consequences will produce an increase in the land needed to absorb CO_2[4]. But what is the rate of compensability implied by these transformations? Are we sure that the decrease in the ecological footprint implied by a more energy intensive agriculture will correspond to an equal increase for the land needed to absorb CO_2?

Technically speaking, this will depend on the assumptions about the elasticity of substitution assumed between the different environmental pressures[5]. Unfortunately, in the computations of the ecological footprint index no specification of this elasticity is made and thus the compensation implied is completely unpredictable and non-transparent. Furthermore, even if the elasticity could be specified, what kind of biological productivity are we considering? What kind of soil? What kind of trees and of which age?

To give another simple land-use policy example, let us consider the issue of urban form. There is consensus that a *compact city* has less environmental impact than a *decentralized city* (see e.g. Frey, 1999). If there is high population pressure, taking only the environmental point of view into account, it would be better to have the people live in compact cities than spread out over the regional territory. But the ecological footprint index, will surely be very high for a compact city and on the contrary quite unpredictable in the case of a decentralized city. In this latter case the computations will depend crucially on what is defined as a homogeneous metropolitan area.

As we have previously seen, when dealing with complex systems operating on several hierarchical levels, the simultaneous existence of contrasting but "correct"

---

[4] This point was put to me by Joan Martinez-Alier.

[5] This is the same issue connected to the use of economic production function measured in monetary terms. In this case, the elasticities of substitution between different production factors are always clearly specified, e.g. a Cobb-Douglas type.

scientific assessments has to be accepted. Complex urban systems are entities which change their identity according to the particular hierarchical space scale at which they are described. A study of a block inside a city, of the administrative unit constituting a "Commune", or of the "metropolitan area" is likely to give completely different and contrasting views and policy suggestions. And so, if we take the example of the hierarchical level of the "Commune of Barcelona", the claim that the quality of life is steadily increasing seems to be correct (or at least this perception is shared by most of its inhabitants). If we look at the whole metropolitan area, the same statement appears less convincing in the light of one problem: the transfer of most of the polluting activities from the city centre to the periphery.

This is the reason why the ecological footprint is often computed for regions or countries. This leads us to another question: are political territories also relevant in ecological terms? And what about trade? The latter issue has been addressed intensively by van den Bergh and Verbruggen (1999). A discussion of the pros and cons of this index is also available at the "Forum on the ecological footprint" in Ecological Economics (2000).

In conclusion it should be emphasized that computing the inverse of the concept of carrying capacity will not help to eliminate its shortcomings. Indeed, by definition an inverse retains all the properties and limitations of the original concept, as this has become evident in the discussion above.

It is impossible to find scientifically sound conversion factors that can transform all ecological, economic and social dimensions in land as well as in energy, money and similar. The concepts of urban environmental carrying capacity and ecological footprint are examples of *ecological reductionism*, i.e. socio-economic and cultural aspects are completely neglected (e.g., transforming the Colosseum into a wooded area would theoretically improve the ecological footprint of Rome).

There have been various attempts to develop multi-dimensional systems of urban sustainability indicators (e.g., CEROI, ICLEI, and many others), without producing any consensus on the pros and cons of any specific system. However, I would like to address another issue here, relevant for the policy-making process and connected to the simultaneous use of various indicators: often some indicators improve while others deteriorate when computed for a specific city. How might such indicators be aggregated? As we know, multi-criteria evaluation is clearly relevant[6].

2.3.2 An Example of Economic Reductionism: The "Fetishism of Fictitious Commodities"

The starting point here is that it is impossible to deal with the concept of "economic value" (and connected economic policy instruments) as an *objective, value-free*

[6] See Munda and Nardo (2007) for a more formal analysis.

category. Economic development implies the creation of new assets in terms of physical, social and economic structures. In a process of “creative destruction” traditional environmental, social, and cultural assets owed to our common heritage may disappear.

Indeed, as discussed in Chap. 1, the key question is *value for what and for whom?* For example, if the objective is to reduce the tourist pressure on Venice, one might conceive of limiting the number of visitors by imposing an entry fee and using the money collected to maintain the city’s cultural heritage. However, one could argue that due to the “relative scarcity” of an economic good such as Venice, people will be quite willing to pay the price of an entry ticket. Thus, the economic instrument “entry ticket” will be useful for collecting money, but not for reducing the tourist pressure. As a consequence, the maximum number of visitors allowed per day should be clarified, and this can only be done on heuristic grounds since tourist carrying capacity can hardly be computed precisely.

Let us now move to a more fundamental question: people who do not visit Venice have an interest in its preservation? If the answer is yes, the concept of “total economic value” (see Box 2.1) immediately becomes relevant. To attribute monetary values to the historical heritage implies capturing *user* (actual, option and bequest) and *non-user* (existential, symbolic, etc.) *values*. Of course, to compute total economic values has nothing to do with the “true” or “correct” value. All monetary valuation attempts will suffer deep uncertainties such as: which monetary

Box 2.1 Total economic value

A fundamental concept in defining a value for non-market goods and services is the concept of *total economic value* (TEV). The TEV is the sum of four elements: the actual use value, the option value, the bequest value and the existence value. The *actual use value* is the value derived from the actual use of the good, e.g. an environmental space for recreation. The *option value*, instead, is the value derived from a possible use of it in the future by the current generation. Both of them have to do with individual preferences and the willingness to pay to conserve for example an ecosystem or a monument or the willingness to make use of it. Given the uncertainty about future uses, the option value is likely to be positive, indicating an interest in preserving the resource for the future. The *bequest value*, instead, implies an intergenerational dimension: it is a willingness to pay to preserve the good for the use of future generations (and not only for the future use of the current generation). Finally the *existence value* has to do with “the concern for, the sympathy with, and the respect for the rights or welfare of no- human beings” (Turner et al., 1994, p. 113).

valuation technique should be used[7]? Which time horizon should be considered? Which social discount rate?

Moreover, c*an we use money values as a social decision tool for sustainability policies?* If the answer is affirmative a measurement of social costs and benefits should be made on the basis of the so-called "*compensation principle*" (usually associated with the names of Hicks and Kaldor). According to this principle, the social cost of a given event is defined as the sum of money paid as compensation to those who have suffered injury. The level of utility[8] that the damaged people had before the event took place should determine the amount of compensation to be paid.

The general economic foundation of monetary compensation to the victims of an environmental or whatever else destruction is the concept of negative externalities. According to Baumol (1969) the relevant aspect of externalities is that *the activity of a subject negatively interferes in the utility function of another subject without an economic transaction between them*. We experience negative externalities in our everyday life: a barking dog at night, the smoke of a neighbour in a restaurant, the volume of teenagers' music, are a few examples. In many cases related with the environment this interference is in the utility function of a whole community, like for example the case of waste residuals or the pollution of a water source or costal zones.

Sustainability policies based on principles of compensation and substitution might sometimes be operable, but one should be very cautious *in applying such principles as a general rule*. The difficulties in substituting for the loss of environmental goods such as biodiversity (which is not even inventoried), or in compensating future generations for the uncertain, irreversible negative externalities we are causing today should be explicitly considered. These are allocations without any possibility of transactions in actual or fictitious markets. Who would be willing to accept compensation for the destruction of the "Sagrada Familia", the "Statue of Liberty" or the "Colosseum"? We could argue that, the presence of irreversibility and uncertainty urges us to transform the compensation principle into the *precautionary principle*[9] (it is more prudent a social conservationist attitude).

---

[7] Valuation techniques try to derive willingness to pay for a good or service by using the concept of consumer's surplus. Consumer's surplus is the difference between what an individual has to pay for a good (the market price) and what an individual would be willing to pay for each unit of the good rather than to go without it. Hence consumer's surplus is defined as the area below the demand function and above the price line. For environmental and artistic goods and services, which have no market price, the consumer's surplus is defined as the area below the demand function (and above the zero price line). A necessary condition for an effective calculation of the aggregate consumer surplus is knowledge of the demand curves for the elements of the targeted project.

[8] Historically, the word "utility" was used in economics to denote the subjective sensations which are derived from consumption. The economists of the late nineteenth century, who were concerned with constructing a theory of consumer choice, regarded utility as something which could be measured as an absolute quantity. They thought it possible to speak of the total quantity of utility derived from consuming a given bundle of goods, of subtracting such quantities from each other, and discussing how these differences changed as consumption varied (Gravelle and Rees, 1992, p. 74).

[9] The definition of the precautionary principle given at the Bergen Conference on Sustainable Development (1990) is the following: "*It is better to be roughly right in due time, bearing in mind the consequences of being wrong, than to be precisely right too late*".

Of course, this principle implies that the majority of society, mainly non-experts outside the economic system (i.e. outside market mechanisms), would decide the “amount” of cultural or natural capital desired. This becomes evident when talking about the “Sagrada Familia” (a church designed by Gaudí in the nineteenth century, still under construction in the city centre of Barcelona). Although some “experts” agree in that it should not be completed, society at large feels a strong commitment to and involvement in its construction. One reason might be found in its symbolic value for Catalan identification (Catalunya is a region in the north-east part of Spain characterized by a strong nationalistic feeling).

In this context, from an economic point of view the only instrument left is “*cost-effectiveness*”; that is given a certain “physical” target (e.g. the amount of cultural heritage to be preserved or the amount of contamination to be accepted), it is rational to try to achieve it by means of the lowest possible use of resources (i.e. at the minimum social cost). Obviously there are several possible targets. This is explicitly acknowledged in many instances of environmental management, such as water quality standards (Funtowicz et al., 1999).

In general two rankings are possible:

1. According to cost
2. According to physical target (e.g. the more monuments preserved, the better)

An inclusive debate might lead to the conclusion that the improvement of a physical target score would worth the extra economic cost, but equally the opposite judgement could be reached. In both cases we would have an ordinal ranking of alternatives and “cost-effectiveness” would “fall down” into “weak comparability” operationalized by multi-criteria evaluation, i.e. two criteria and two different rankings must be dealt with explicitly.

From the discussion above the following conclusion can be drawn: to attach prices to non-market assets (such as most of environmental and cultural ones), gives a positive signal to society and may contribute to a more rational use increasing the chances for a better conservation. When one wishes to preserve a monument or a natural area, a fundamental question is: is there any resource, which society is willing to assign to this objective? To answer this question the use of monetary techniques such as hedonic prices, travel costs or contingent valuation is desirable and useful[10].

[10] In the framework of the famous debate of the 1920s and 1930s on economic calculus in a socialist economy Hayek, replying to Neurath wrote (1925, p. 31): “Neurath was quite oblivious of the insuperable difficulties which the absence of value calculations would put in the way of any rational economic use of the resources…”. Or, as Von Mises had put it (Von Mises, 1920, in Hayek, ed. 1935, p. 111), “Where there is no free market, there is no pricing mechanism; without a pricing mechanism, there is no economic calculation”. Certainly, the market would sometimes fail to give economic value to environmental amenities, thus, the calculation of the profitability of a hydroelectric scheme would not include “the beauty of the waterfall which the scheme might impair”, except that attention could be paid “to the diminution of tourist traffic or similar changes, which may be valued in terms of money” (Von Mises, in Hayek, ed. 1935, p. 99). Through what is now called the “travel-cost method”, or similar methods, the market mechanism could be extended in a capitalist economy to positive or negative externalities.

However, one should remember that the market alone may be successful in efficient allocation of resources, but it does not give any guarantee at all for preservation of the cultural or natural heritage. Once something is on the market, it can be bought or sold, and so the willingness to accept and the compensation principle may easily cause the destruction of any asset. Monetary compensation is without a doubt the only possible tool when irreparable and irreversible damage occurs. This way, if an accident causing serious contamination occurs – as in the case of Seveso in Italy (1976), Bhopal in India (1984), the Exxon Valdez in Alaska (1989), or more recently the oil-tanker Prestige off the coasts of Galicia (2002) – it seems correct and appropriate to indemnify the victims.

It remains to be verified whether in the long run, compensation is an effective tool to prevent contamination, given that it does not guarantee the preservation of natural or artistic goods. Take the case of Catalan aquifers. In Catalonia ("*Catalunya*" in the Catalan language) there have recently been serious problems of contamination of the underground aquifers. The damage has been attributed to pigs which are imported from Holland, raised in Catalonia and subsequently sent back to Holland for slaughter. The Dutch like this solution because in their country there are serious problems of ground contamination and pig-breeding has become practically impossible. Who profits from this? Obviously, the Dutch and some Catalan families who owe their economic prosperity to this activity. Who pays the costs? Catalan society as a whole to which the Dutch have succeeded in transferring the environmental costs, and which is losing a resource of such vital importance as drinkable water. Private costs born by breeders are thus very different from those born by the whole Catalan community. Whenever monetary compensation to victims was paid, this did not necessarily bear any positive consequences for either the environment or, in the long run, the inhabitants of the region left without drinking water.

Economic value is different from environmental or artistic-cultural value. If we had to decide whether to save the Galapagos Islands or the inside sea in Holland, which value should one use? The economic would favour the inside sea, which, since totally eutrophized, represents an important economic service receiving all the nutrients coming from human activity. The ecological would obviously favours the Galapagos Islands. The choice of the values to be considered as socially predominant is a scientific or socio-political issue? Again, the issue of incommensurability of values becomes relevant.

The application of the precautionary principle introduces some elevated costs, but how much would its non-application cost? The burden could be enormous, as admitted by the European Environment Agency. Even the Economist (certainly a magazine well distant from being environmentalist) has recently suggested as a possible positive consequence of the accident of the Prestige (a ship which heavily contaminated the coasts of Galicia in northwest Spain) a stiffening of the European legislation on the subject of maritime transport (The Economist, November 23–29, 2002, p. 79). When we abstract from myopic logics, there is no doubt that for society it is ecologically and economically more convenient to apply the precautionary principle than to suffer a series of disastrous accidents. For this reason, when uncertainty and irreversibility are present, it may be prudent to exchange the compensation principle for the precautionary one.

Summarizing this discussion, we can say that to "*internalize*" externalities into the price system might in general have positive consequences from a sustainability perspective. But one should not forget the uncertainties and complexities which make it difficult to give physical and economic measures for externalities. Moreover, it is worth remembering how economic values depend on inter- and intra-generational inequalities in the distribution of the burdens of pollution and in access to natural resources. Thus externalities can be seen as "*cost-shifting*" or as "*ecological distribution conflicts*" (Martinez-Alier and O'Connor, 1996). In general, if the injured parties are poor (or even not yet born), the cost of the internalization of the externality will be low. This is why a lot of multinationals locate particularly dangerous production plants in developing countries where, in the case of accidents, they are generally forced to pay much lower monetary compensations than in western countries. The accident of the chemical plant of the Union Carbide in Bhopal, India, in 1984, is a sad example (Jasanoff, 1994; Rajan, 2002). Obviously, the institutional and juridical context is fundamental. In the case of the oil contamination caused by Texaco in Ecuador (with serious consequences for human health) the point in the trial was deciding whether the competent court should have been in the USA or in Ecuador. Texaco naturally insisted on having the trial held in Ecuador… (Martinez-Alier, 2002, pp. 102–107).

Accepting low values for a negative externality that has negative impacts on a poor community is a "political decision", that is far from being ethically neutral. Some years ago, an internal document of the World Bank, subsequently made public, suggested that toxic waste should be stored in Africa, since the cost of compensation would be extremely low and would therefore have to be considered as the most efficient solution (it is interesting to remember that the World Bank is supposed to be working on behalf of poorer countries).

Allen et al. (2003) summarize the basic sustainability issues in the following questions.

Sustainability of:

1. What?
2. For whom?
3. How long?
4. At what cost?

It is clear that economic instruments are designed to answer only the fourth question, they thus need to be complemented with other approaches if one wishes to deal with sustainability in a comprehensive manner. Monetary valuation methods are based on phenomena such as consumer's surpluses, market failures and demand curves which are only a partial point of view, since connected *with one institution only: markets*. From a sustainability point of view, issues connected with actions outside the markets and as well as behaviour of people different from the class of consumers should also be taken into account (Duchin and Lange, 1994). As noted by Funtowicz and Ravetz, (1994, p. 198), "the issue is not whether it is only the marketplace that can determine value, for economists have long debated other means of valuation; our concern is with the assumption that in any dialogue, all valuations or "numeraires" should be reducible to a single one-dimension standard". I would like to stress that incommensurability of values does not imply a hierarchy of values. "Intrinsic" value (or "end value" (Lockwood, 1997)) is sometimes considered "superior" to economic value, this is not the position I am defending. Beckerman and Pasek (1997, p. 65) rightly note that: "the frequent claim that the environment has some unique moral intrinsic value is unsustainable, its preservation often raises ethical and other motivations that are not commensurate with the values that people place on ordinary marketable goods".

I reiterate that I am not opposed to putting economic values on natural resources, to environmental sinks, to natural spaces or to cultural heritage. A location may be valuable for its biodiversity (measured in richness of species or genetic variety), and also as a landscape, yet also have economic value (measured by differential rent or by the travel cost method, or contingent valuation). These are different types of value. The point is that it is misleading to take sustainability decisions based on only one type of value. The "*fetishism of fictitious commodities"" and "energy myths*" must be avoided. Thus, instead of focusing on "missing markets" as causes of allocative disgraces, or trying to explain economic values by means of energy measures (clearly a non sense from an economic point of view) we should focus on the creative power of missing markets, because they push us away from economic commensurability, towards multi-criteria evaluation of evolving realities.[11]

The next chapter gives some guidelines for the real-world treatment of both technical and social incommensurability.

[11] "There is great pressure for research into techniques to make larger ranges of social value commensurable. Some of the effort should rather be devoted to learning – or learning again, perhaps – how to think intelligently about conflicts of value which are incommensurable" (Williams, 1972, p. 103). A call for dealing explicitly with incommensurability can also be found in Arrow (1997).

Chapter 3
Operationalizing Technical and Social Incommensurability in an SMCE Framework

3.1 The Orchestration of Sciences: Technical Incommensurability

One should note that the construction of a descriptive model of a real-world system depends on very strong assumptions about (1) the *purpose* of this construction, e.g. to evaluate the sustainability of a given city, (2) the *scale* of analysis, e.g. a block inside a city, the administrative unit constituting a commune or the whole metropolitan area and (3) the set of dimensions, objectives and criteria used for the evaluation process. A reductionistic approach to building a descriptive model could be defined as the use of just *one measurable indicator* (e.g. the monetary city product per person), *one dimension* (e.g. economic), *one scale of analysis* (e.g. the Commune), *one objective* (e.g. the maximization of economic efficiency) and *one time horizon*. If one wants to avoid reductionism, it will be necessary to take incommensurable dimensions into account and to use different scientific languages describing disparate but legitimate representations of the same system. This is what Neurath (1973) called the need for an "*orchestration of sciences*".

It is clear that a multi-criteria approach, being multi-dimensional in nature, seems an interesting framework in which operationalize Neurath's ideal. This virtue of multi-criteria approaches has been corroborated in a great number of real-world case studies employing a variety of methods (see e.g. Janssen, 1992; Maystre et al., 1994; Beinat, 1997; Stewart and Joubert, 1998; Moreno-Jiménez et al., 1999; Espelta et al., 2003). A real world case study involving the water supply system of the city of Palermo in western Sicily (Southern Italy) can help to clarify the point.

The case was part of a project commissioned by the region of Sicily and executed in the frame of the European Commission's DGXVI structural funds. This case study was developed during two years of interaction mainly between a multi-disciplinary team and the management body of the water supply system of the city of Palermo (plus some social actors involved in the final step of the study) (for more information on this case study see POP Sicily, full final report European Commission contract No.10122-94-03 TIPC ISP I or for a shorter version Munda et al., 1998).

Water resource management is characterized by the presence of strong competition among different categories of consumptive water uses and, as a consequence, among various interest groups. Such competition also exists between consumptive uses as a whole and "ecological uses" which aim to limit water diversion for off-stream uses in order to preserve the ecological equilibrium of ecosystems. This permanent state of competition may intensify into real conflict under drought conditions, i.e. when there is a temporary reduction of available water resources due to a long and severe decrease in rainfall (compared to mean or median natural values). The problem of water shortages due to drought is particularly relevant in southern Europe. In Sicily, the water distribution issue has deep historical roots. Indeed *mafia* started from the fighting over water control.

Water shortages not only depend on hydrological drought, which in turn follows from meteorological drought, but also on water supply system characteristics and demand levels, which are both affected by different drought mitigation measures. As a consequence, the purely technical hydrological solutions cannot be separated from their consequences for the socio-economic system. Although this was not evident in the beginning of the project, after a few meetings, hydrologists accepted that an economist could be of some help with this kind of problem. However, it was still very difficult to find a common language and to understand which contribution each could make towards a possible solution (or at least a better understanding) of such a complex problem.

The water system of Palermo provides water to municipal, agricultural and industrial users by using surface water and groundwater; a reservoir is also used for energy production.

It was agreed that alternative management options under drought conditions could be divided into two main groups:

- alternatives that try to satisfy 100% of water demands,
- alternatives that do not completely satisfy water demands.

To establish the alternatives, it was necessary to understand the structure of the Palermo water supply system and, given the technicalities involved (as one can see from its description in Fig. 3.1), it was immediately clear that this was the job of hydrologists. However, these alternatives had to be evaluated for the longest historic drought experienced in the water supply system (four years) according a set of criteria which included the economic dimension (e.g. associated financial costs and benefits for the company managing the water supply system, the energy production company, and so on), the social dimension (e.g. hygienic risk and social discomfort) and the environmental dimension (e.g. the in-stream flow requirement defined as the discharge which maintains a stream ecosystem or aquatic habitat). At this point the advantage of the multi-criteria structuring of the problem became evident. Each expert suddenly knew her/his comparative advantage.

From the experience of this case study, a first lesson could be learned: *a multi-criteria framework is a very efficient means of implementing a multi-inter-disciplinary approach*. The experts involved had diverse backgrounds (mainly in engineering, economics and mathematics). While the communication process was initially very difficult, once it had been decided to structure the problem in a multi-criterion fashion, it was astonishing to recognize that a common language had immediately been created.

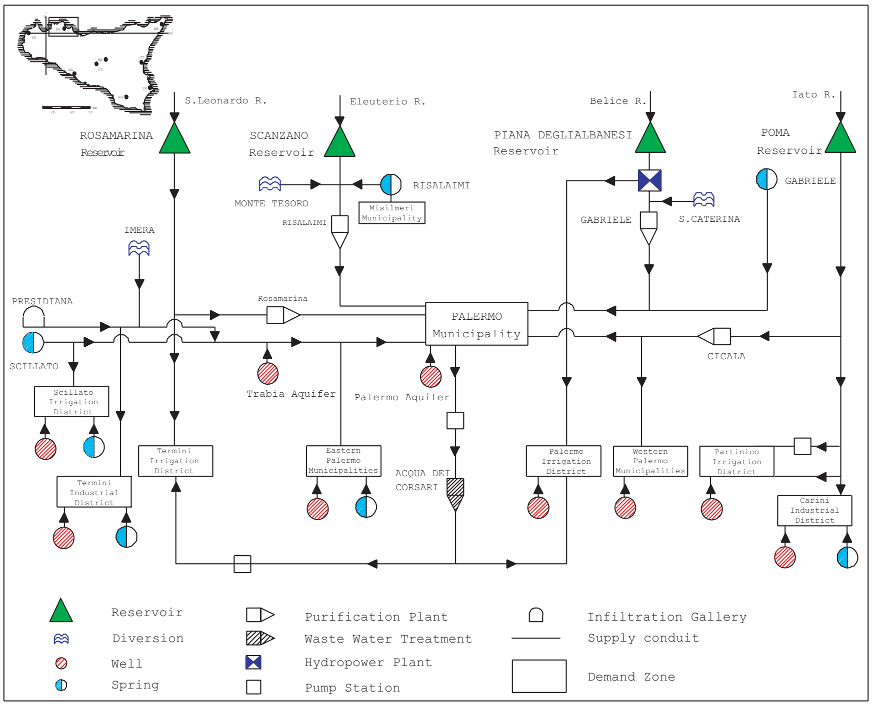


Fig. 3.1 Scheme of the Palermo Water Supply System (Source: Elaborations made by G. Rossi and his team at Catania University)

In terms of inter-disciplinarity, the issue is to find agreement on the set of criteria to be used; in terms of multi-disciplinarity, it is to propose and compute an appropriate criterion score. The efficiency of the interaction process may also increase greatly[1].

In the Palermo case study, it was also experienced that explicitly taking distribution issues into account increases the transparency of the study and facilitates an effective process of interaction between various social actors. This second lesson leads on to the issue of social incommensurability and public participation.

3.2 SMCE is about Democracy: Social Incommensurability

At this point in the discussion the question arises as to *who makes the decisions*? Some critics of multi-criteria evaluation say that *in principle*, in cost–benefit analysis, votes expressed on the market by the whole population can be taken into

[1] Here I refer to the idea of orchestration of sciences as a combination of multi/inter-disciplinarity. Multi-disciplinarity: each expert takes her/his part. Inter-disciplinarity: methodological choices are discussed across the disciplines (this definition has been discussed with R. Strand).

account (naturally with the condition that the distribution of income is accepted as a means of allocating votes)[2]. On the contrary, multi-criteria evaluation may be based on the priorities and preferences of only a few decision-makers. (We could say that the way these decision-makers reach their position is accepted as a way of allocating the right to express these priorities.) The criticism may be correct if a "technocratic approach" is taken, in which the analyst constructs the problem relying solely on expert input ("expert" meaning those who know the "technicalities" of a given problem).

For the formation of contemporary public policies, it is hard to imagine any viable alternative to *extended peer communities* (Funtowicz and Ravetz, 1991, 1994; Gowdy and O'Hara, 1996; Funtowicz et al., 1999; Corral-Quintana et al., 2001; De Marchi and Ravetz, 2001; Guimarães-Pereira et al., 2003, 2005; Kasemir et al., 2003). They are already being created, in increasing numbers, either when the authorities cannot see a way forward, or know that without a broad base of consensus, a policy will not succeed. They are called "citizens' juries'", "focus groups", or "consensus conferences", or any one of a great variety of names; and their forms and powers are correspondingly varied. But they all have one important element in common: they assess the quality of policy proposals, including the scientific and technical component. And their verdicts all have some degree of moral force and hence political influence. Here the quality is not merely in the verification, but also in the *creation*; as local people can imagine solutions and reformulate problems in ways that the accredited experts, with the best will in the world, do not find natural.

However, even a participatory policy process can be conditioned by strong value judgements such as, the relative importance of all the social actors (i.e. their weight); whether a socially desirable ranking should be obtained on the grounds of the majority principle; whether some veto power should be conceded to the minorities; whether income distribution effects are important, etc.

The management of a policy process involves many layers and kinds of decisions, and requires the construction of a *dialogue process* among many stakeholders, individual and collective, formal and informal, local and otherwise. This necessity has been winning ever more recognition in a multi-criteria decision-aid (MCDA) framework. Banville et al. (1998) offers a very well structured and convincing argumentation on the need to extend MCDA by incorporating the notion of the stakeholder. A social multi-criteria process must therefore be as *participative* and as *transparent* as possible; although in my opinion, participation is a *necessary* but not *sufficient* condition. This is the main reason I propose the concept of "Social Multi-criteria Evaluation" (SMCE) in place of "Participative Multi-criteria Evaluation" (PMCE) or "Stakeholder Multi-Criteria Decision Aid" (SMCDA) (Banville et al., 1998). To clarify this very important point, the experience of the so-called VALSE project (see VALSE full final report, Chap. 9, European Commission ENV4-CT96-0226, or for a synthesis De Marchi et al., 2000) is instructive.

---

[2] One should note that cost-benefit analysis can indeed be easily criticized both from the distributional and environmental points of view (see e.g. Munda, 1996; Spash and Hanley, 1995).

Troina is a small town (10,000 inhabitants) in north-east Sicily, Italy. On one hand, it seems there is a common assumption that there is a real water shortage which could be remedied by more effective use of existing resources. (Paradoxically, although water shortages are common in Sicily, Troina is an exception). On the other hand, there is a complex and heterogeneous collection of interests in the Troina water issue, which have had no effective dialogue. Hence, an effective structuring of the water problem at this early stage is an important task, so that eventual negotiations between social actors may have a better chance of a positive outcome. The steps of the overall evaluation process are schematized in Fig. 3.2.

One should note that policy evaluation is not a one-shot activity; on the contrary, it evolves as a *learning process*. It has to be realized that the evaluation process is usually highly dynamic, so that judgements regarding the political relevance of items, alternatives or impacts may display sudden changes, hence requiring a policy analysis to be flexible and adaptive in nature. This is the reason that evaluation processes have a *cyclical nature*. By this is meant the possible adaptation of elements of the evaluation process due to continuous feedback loops between the various steps and consultations among the actors involved (Nijkamp et al., 1990).

The first question to be answered is the following: *is "business as usual" a feasible option in the long run*? Business as usual, in this case, is a situation where power and water management are fragmented among the main actors and where infrastructure decisions are the only ones not requiring agreements. This can be considered the classic case of non-cooperative resource exploitation.

For example, the Municipality of Troina is trying to become self-sufficient in its drinking water needs by using its own spring water sources, even if this could be

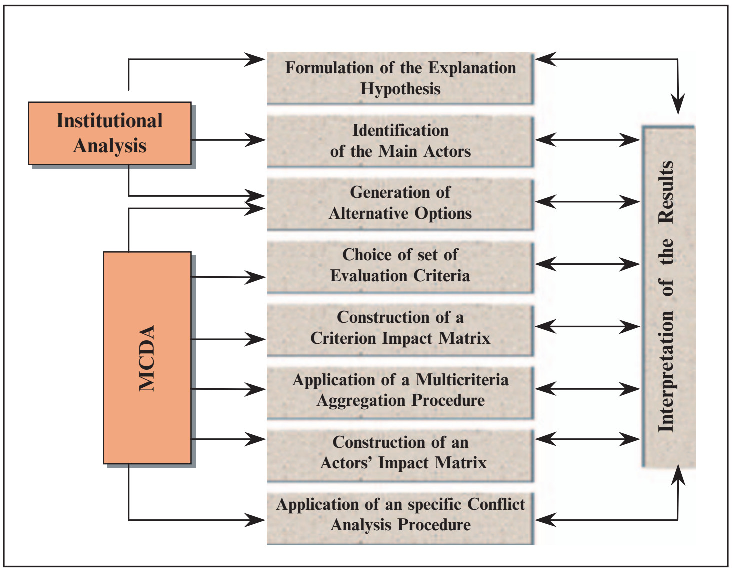


Fig. 3.2 Scheme of the evaluation process in the Troina case study

perceived by some as inefficient. To evaluate the business-as-usual option properly, it has to be compared to a set of different possible options on the basis of some evaluation criteria. At this point, an issue arises: alternatives and criteria for whom? This obliges us to acknowledge the preferences of the actors playing an important role in the dispute at hand.

Initially, only the actors playing an important role in the community of Troina (as a result of the institutional analysis) were taken into account. Later on, as a surprising result of the process of generating alternative options, it became clear to everybody that additional interest groups outside Troina also had to be taken into account. This learning process was very interesting particularly for the local administrators of Troina, who fully realized the importance of Troina water resources outside their own territory. As the Mayor acknowledged, such a process of structuring the problem was extremely useful for understanding the hierarchy of interests underlying the exploitation of local natural resources.

During the study, the top ranked position of the alternative proposing an *information campaign* was an unexpected surprise. The reaction to this result was the idea to stage within a very short time horizon, an exhibition on water management issues in the town of Troina. The Mayor and the municipal administration thought that the implementation cost of such a policy measure could be quite low and the potential positive impact on the community quite high. Of course, the political risks for the administration could also have been very high, since it was clear that a lot of powerful actors were working hard to maintain the status quo[3]. This point leads us back to the initial and principal question, is business as usual a defensible option?

One should note that business as usual was positioned almost at the bottom of the multi-criteria ranking. (In the NAIADE conflict analysis, it was in a low position for some actors and in a high to medium position for all the others.) Almost all the powerful social actors of the Troina community belong to this second group. We could say that the status quo is a compromise solution among the opposing internal interests. This may explain why few are willing to change the present situation (since it is very risky for the community at large). However, this situation seems much more of an *impasse* than a real equilibrium.

The study attempted to avoid the pitfalls of the technocratic approach, by applying different *methods of sociological research*. The "*institutional analysis*", performed mainly on historical, legislative and administrative documents, provided a map of the relevant social actors. Much insight was offered by "*participant observation*" as some contributors to the study were also members of the community and knowledgeable of its internal dynamics. The possible biases of this "insider perspective" were checked against the information obtained from some "*in-depth interviews*" with key local actors. Finally a "*survey*" using a questionnaire was performed on a random sample of the resident population, so as to explore their perception of the water issue in Troina (see also Box 3.1).

To elucidate the arguments to be developed in this section, I will refer to another case study, the DIAFANIS project (financed by the Spanish Ministry of Environment,

[3] In fact, I must say that the Mayor and his administration lost the next elections.

Box 3.1 Institutional Analysis

The social dimension of a problem can be explored using *institutional analysis*, a tool that can illuminate values, interests, roles, possible alliances and available resources of the social actors involved in a policy problem. The objective of institutional analysis is to gain insights into three aspects (Corral-Quintana, 2000).

1. **The problem at hand**. The actors' perceptions of the problem might diverge significantly, depending on their objectives and interests, their knowledge, resources and role. In order to understand the problem, as much information as possible must first be gathered on the *general context*, that is, the physical, economic, social and political characteristics of the territory where the conflict takes place. Secondly, the *legal framework* must be studied. Finally, a *chronology* of the events that led to the present situation must be pieced together.
2. **The social actors**. The social actors are those who can influence or whose interests are affected by the analysed policy options. For each group of social actors, roles, objectives, interests and resources are defined. Resources are the means that can be used in order to reach an objective, they can be economic (amount of money), political (capacity of influencing the decision making process), legal (advantages given by a law), cognitive (knowledge on the topic or on the decision process, or ability to understand other agents' behaviour).
3. **The interaction patterns**. The structure of the institutional network, the kind of interactions and the arena where interactions take place constitute indispensable information, as do influences and changes in the social actors' positions.

Different written and oral sources are to be used for carrying out an institutional analysis. In the first category we could mention local and national press, specialized magazines, official and informal documents produced by the social actors in order to explain their position, books, articles, and so on. In the second group we find individual interviews with key agents or with a random sample or focus groups. Normally, many complementary sources are used simultaneously.

see final report (in Spanish and Catalan) and Martí, 2001 (in Catalan)). This project was named "*diafanis*" to indicate that the emphasis of the approach was on the aspect of transparency.

The problem dealt with was the possible expansion of a skiing infrastructure in the Catalan Pyrenees (north-east Spain). It was very clear from the beginning that the choice of the geographical scale would determine the policy option considered desirable. In fact, local people living close to the area think that the expansion would bring more tourists and as a consequence more economic prosperity. This perception changes as soon as one leaves the immediate neighbourhood affected by the expansion project. Thus, for example, in Barcelona preservationists object to the project, since the area in question is close to a natural park and has even been declared by the autonomous government of Catalonia to be a possible natural area of European interest. What then is the appropriate scale? The local area, the entire Pyrenees, Catalonia or even Europe as a whole?

To understand if other possible courses of actions existed, an institutional analysis was carried out and consequently some participatory techniques were employed (see Fig. 3.3). By means of focus groups it was possible to gain an idea of people's preferences and then to develop a set of policy options. A limitation of the focus group technique, immediately evident, was that at the local level, some people were not willing to say publicly what they really thought, since they were afraid of the consequences for their everyday lives (social exclusion in small communities can be a calamity, or some people considered their jobs in danger, if for example, they were working for a hotel owner in favour of the skiing infrastructure). When far from the immediate vicinity of the affected area, this component of social control was almost absent. For this reason anonymous questionnaires and personal interviews are an essential part of the participatory process.

The selection of evaluation criteria was also based on what was learned through the participation process. However, at this stage a problem arose: should the evaluation

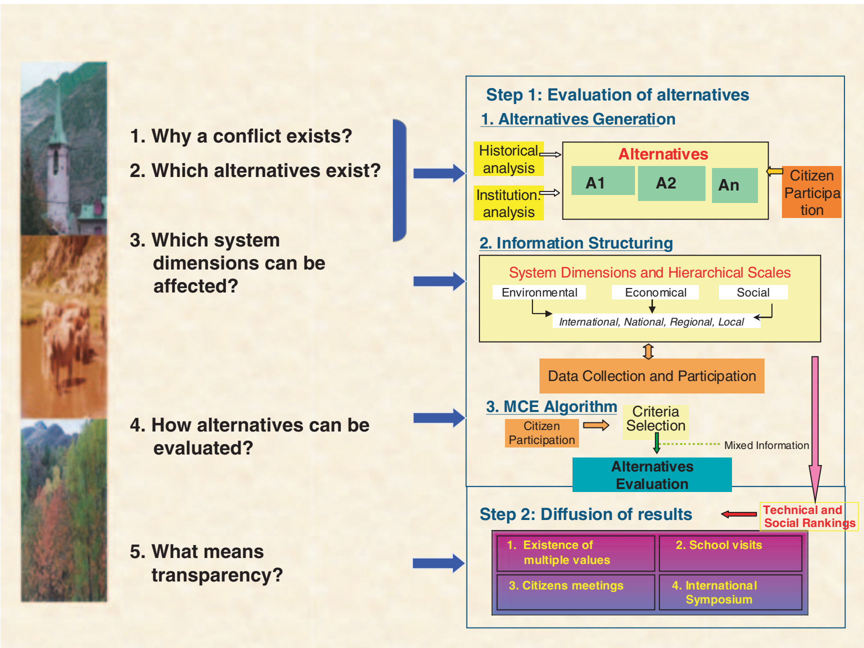


Fig. 3.3 Structure of the evaluation process in the DIAFANIS project

criteria come directly from the public participation process or should they be "translated" by the research team? It was soon understood that the raw material collected during interviews and focus groups could be used as a source of inspiration, but the technical formulation of criteria with properties such as "non-redundancy", "legibility" and so on (see Bouyssou, 1990) is a clear task for researchers. Of course, in this step, subjectivity is unavoidable (for example, there were many discussions within the team on how to contain the biases of certain members with strong ecologist convictions).

The same criticism of the subjective component of the research team can easily be made when summarizing the impacts of the various courses of action on the different social actors (e.g. to build the NAIADE conflict analysis procedure (see Chaps. 5 and 8)). This is obviously true, although the social scientists involved in the study greatly appreciated the possibility to work with an operational framework which allowed them to synthesize the large amount of non-formalized information collected during their field investigations.

Being conscious of the subjective and sometimes even arbitrary factors inherent in the study, a widespread information campaign was planned on the assumptions and conclusions of the study including local people, regional and national authorities, international scientists and even children at school.

Some interesting lessons emerged from these case studies:

1) One should not forget that the classical schematized relationship between decision-maker and analyst is indeed embedded in a social framework, which is of crucial importance in the case of public policy.
2) The combination of various participatory methods, which has proved powerful in sociological research, becomes even more so when integrated within a multi-criterion framework.
3) The use of a cyclical evaluation process allows for the incorporation of things learnt during the study. It is extraordinarily important that different participatory and interaction tools be used at different points in time. This allows for continuous testing of assumptions and unavoidable biases of the study team.
4) According to the geographical scale chosen, the relevant social actors with interests at stake can be found through institutional analysis. Institutional analysis is an essential step to identify possible "stakeholders" for a participative process. However, besides the unavoidable mistakes that may occur when carrying out an appropriate institutional analysis, there are even stronger reasons why a purely participatory study is undesirable.
5) The scientific team cannot accept uncritically all input of a participatory process, since:
    a) In a focus group, powerful stakeholders may influence all the others quite strongly.
    b) Some stakeholders might not desire or be able to participate, but for ethical reasons the scientific team should not ignore them.
    c) The notion of stakeholder[4] only recognizes relevant organized groups; this is why the term *"social actor"* is preferable.

[4] Banville *et al.*, 1998 discuss the notion of a stakeholder and its ambiguities deeply.

d) Focus groups are never meant to be a representative sample of the population. As a consequence, they can be a useful means of improving the researchers' knowledge of the institutional and social dimensions of the problem at hand, but they are never a way to derive consistent conclusions on social preferences.

These conclusions lead to the following personal (and thus arguable) *convictions:*

(1) Transparency is an essential feature to guarantee the quality of any study based on science for policy. In fact, all such studies should be *accountable* to the public at large for peer-review (accountability is a concept recently proposed by the European Commission in the White Paper on European Governance[5]).
(2) Multi-criteria methods are a powerful framework for policy analysis since this type of evaluation process can be very effective since it accomplishes the goals of being *inter-multi-disciplinary* (with respect to the research team), *participatory* (with respect to the local community) and *transparent* (since all criteria are presented in their original form without any transformations into money, energy or any other common measurement rod).
(3) Since decision-makers require legitimacy[6] for the decisions taken, it is extremely important that public participation and scientific studies do not become the justification for a lack of political responsibility. I strongly believe that the de-ontological principles of the scientific team and policy-makers are essential for assuring the quality of the evaluation process. Social participation does not imply that scientists and decision-makers have no *responsibility* for policy actions defended and eventually taken.
(4) As a consequence, ethics matter[7]. Let us imagine an extreme case in which a development project in the Amazon forest could affect an indigenous community with little contact with other civilizations. Would it be ethically more correct to invite them to a focus group… or to take into account the consequences of the project for their survival? And what about future generations and non-humans?
(5) A positive externality of participatory approaches is that the results obtained by the research team, i.e. data, findings, interpretations and insights, can sometimes also be returned to the community, which may then use them not just as a given, but rather as an input for deliberative democracy. In sum a participatory approach can also be an educational tool for the understanding of democracy.

---

[5] The White Paper on European Governance is inspired by the following five main principles: *openness, participation, accountability, effectiveness and coherence*.

[6] On the issue of legitimacy see also Roy and Damart (2002).

[7] The importance of the inclusion of ethical considerations in mathematical modelling and decision-making has also been discussed by Kleijnen (2001) and Rauschmayer (2000).

3.3 Implementing the Ideal SMCE Process

As already discussed, policy exercises are normally dynamic learning processes which must necessarily be adaptive in nature. This implies continuous feedback loops between the various steps and consultations among all the actors involved. Figure 3.4 can be considered an example of an ideal SMCE process, where scientific and social knowledge are perfectly combined. Of course, these steps are not rigid.

On the contrary, flexibility in real-world situations is one of the main advantages of social multi-criteria evaluation (see e.g. Vargas-Isaza, 2004 for an application of SMCE in Colombia, where there was an extreme situation involving social actors belonging to various informal armies (the so-called *actor armado*); Marti, 2005, who conducted a study with indigenous communities in Peru (see Box 3.2); or Sittaro, 2006, who applies SMCE in the context of indigenous communities in the Amazonian region of Ecuador).

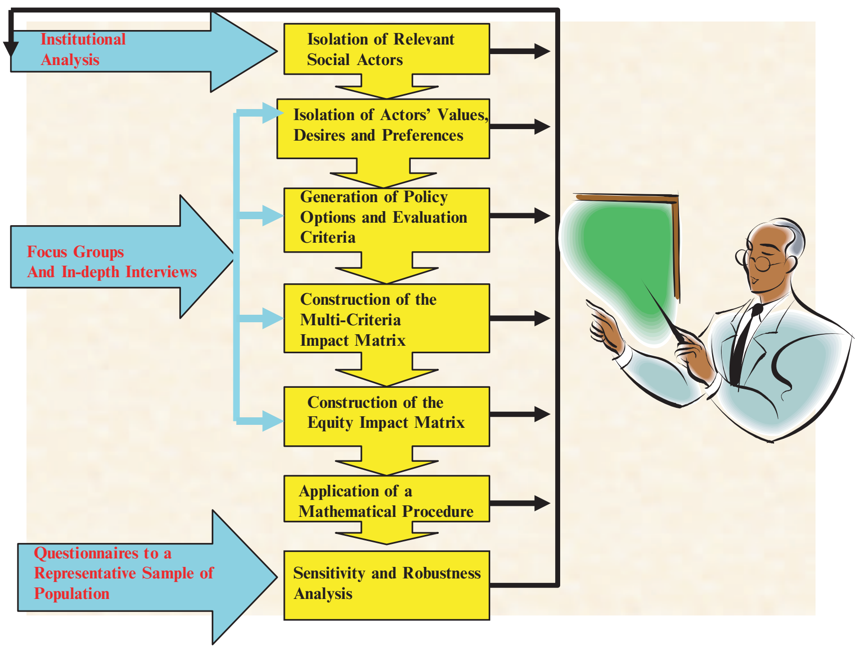


Fig. 3.4 The ideal problem structuring in SMCE

Box 3.2 SMCE with Indigenous Communities in Peru

In the framework of a research project on barter markets and biodiversity conservation in the Peruvian Andes, social multi-criteria evaluation was applied to the context of indigenous community development, in the region of "*Parque de la Papa*[8]. The issue tackled was an evaluation of the incommensurability of values associated to the biodiversity of medicinal plants (see Marti, 2005[9]). A multi-criterion evaluation matrix was developed by Quechua women without the use of written language. The matrix- building process started by taking into account the socio-cultural, political and ecological specificities of the region. The process included the following steps: (a) identification and inclusion in the process of the women in charge of preserving the knowledge of medicinal plants, (b) generation of evaluation criteria by means of collective deliberation, (c) experimentation and further improvements of the evaluation matrix, and (d) development of the evaluation by quantifying the criterion scores.

The women identified and then involved constituted the Group of Evaluation of Medicinal Plants (GEPM). For the generation of evaluation criteria and selection of medicinal plants, interviews with women in charge of the use of medicinal plants ("*curanderas*") were carried out, and advice was asked of experts in phitotherapy. Finally, the GEPM selected a total of 37 criteria and 43 medicinal plants to be included in the multi-criterion evaluation matrix. Later on, symbols were chosen to represent each one of the criteria. Materials and local objects such as kitchen utensils, dirt and wools were used among other things. The objects were chosen according to their association with the semantic meaning of the criteria. Objects and materials of vivid colours were chosen to keep the attention of the group as focused as possible, during the evaluation process. For each criterion, an ordinal scale of measurement was assigned between zero and five. The criterion score was represented by the number of beans. Once the evaluation system had been designed, the multi-criterion matrix was first experimented with by the GEPM in two community meetings. In these meetings the construction of the matrix was done on the floor with the help of wools.

Once the experimentation had been tested successfully, other meetings were organized. They lasted around five hours and the medicinal plants were evaluated in lots of twenty. When the sample of the plants had been definitely chosen, the women assumed different roles and also different positions in the ground space. The mediator located herself inside the matrix, one recorded the results in a lateral position and the group of consultants positioned themselves around the matrix. The mediator showed the plants to the consultants one by one and asked for its evaluation in relation to each criterion.

(continued)

[8] *Papa* in this context means potato.

[9] Research supported by the programme Sustaining Local Food Systems, Agricultural Biodiversity and Livelihoods of the International Institute of Environment and Development, and the "Asociación Kechua-Aymara para Comunidades Sostenible, ANDES".

Box 3.2 (continued)
An open discussion continued until the women arrived at a consensus on the number of beans to be assigned as a criterion score. In the case of persistent disagreements, the discussion included a comparison with the beans granted to the rest of species. Once the evaluation had been concluded, the final matrix was copied onto a paper document and the participants all signed the document to emphasize their role as knowledge keepers in relation to the incommensurability of values associated to the medicinal plants. The procedure described guaranteed that the evaluation had been adaptive, locally controlled and able to integrate different types of knowledge at different scales. The final result was the selection of 21 of the species evaluated to produce medicines for the indigenous community.

(Photo by Neus Marti)

In real-world applications, a very sensitive point is the synthesis of technical and social information to generate alternatives and evaluation criteria. Let us briefly review an application in the field of renewable energy (for more information see Munda et al., 2005[10] and Gamboa and Munda, 2007). In the last decade renewable energies and

[10] This research was carried out in the framework of the European Union project *"Development and Application of a Multi-Criteria Decision Analysis SoftwareTool for Renewable Energy Sources (MCDA-RES)"*, Contract NNE5-2001-273.

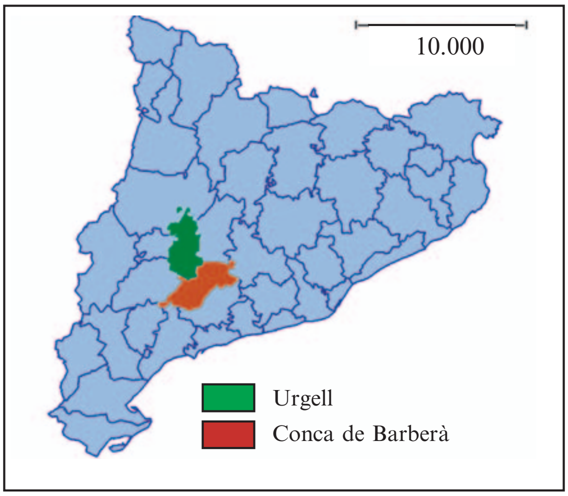


Fig. 3.5 L'Urgell and Conca de Barberà *Comarcas* in Catalonia

especially wind energy have experienced a strong boost from national and international authorities; typically presented as part of the strategy to deal with global warming and to accomplish the Kyoto Protocol. Although wind energy has a green image, it is difficult to find favourable positions for the installation of wind-farms. This opposition can depend on the extensive land use of wind-parks, their possible impacts on birds or their visual impact, as well as NIMBY (Never In My Back Yard) behaviour. The policy process for deciding the location of the wind turbines can itself be a source of conflict. For these reasons when designing, locating and evaluating alternative wind-park sites in Catalonia, a real-world social process was implemented including several social actors' visions. The impact zone is located in the western part of the Catalonian central depression (see Fig. 3.5) between the "*comarcas*" of Urgell and Conca de Barberà.

Two projects were proposed: one of 16 windmills of 850 kW, and a second of 66 windmills of 660 kW. In addition, there were two other projects planned to construct wind-farms of 75 and 15 windmills respectively, reaching 172 windmills in the area.

Early in this process, several positions emerged regarding the construction of those wind-parks. On one hand, some people started to raise objections to the wind-farms. First, they expressed the desire to participate in the design of the future of their *comarcas* and second, they saw territorial inequalities in the way Catalonia has been planning the energy production scheme.

On the other hand, some municipalities and citizens have agreed to the construction of the infrastructures. They see the wind-parks as an opportunity to generate income, to improve social services and to reverse the decline that characterizes these towns. By developing an institutional analysis study and applying various participatory techniques the social *"atmosphere"* could be captured (synthesized in Table 3.1).

Table 3.1 Socio-economic actors, scale of action and their position with respect to wind-parks

Social actor	Scale of action	Position regarding the wind-parks
Catalonian government	National	The Catalonian government has launched the Renewable Energy Plan for the year 2010. It projects that the construction of RES will grow, from 72.2 to 1.073 MW installed capacity. But recently they have made some declarations that they want to scale the installed capacity target to 3.000 MW
Town council of Vallbona de les Monges	Local – Province	The municipality wants the wind-parks to be installed. They see the economic benefits as a good opportunity to improve some social services, and/or to create others (like nursing for the aged)
Town council of Els Omells de Na Gaia	Local – Province	
Town council of Rocallaura	Local – Province	The three municipalities are negotiating together with the companies, trying to obtain equal and better retribution conditions from the promoters. However some of them say that if the economic profit is not enough to offset the actual social trends, then the wind-parks should not be constructed
Town council of Senan	Local – Province	The town council is fighting together with the inhabitants of Senan to oppose the wind-parks. They do not want to be surrounded by windmills, and they see the forest as a resource to develop tourism in the future
Consell comarcal de l'Urgell	Province - National	The president of the council has offered her mediation to help reach a compromise solution. But she shares the opinion of the mayors, in the sense that more financial income is needed to revitalize the towns and to offer more and better services
Politic representatives	Province	Representatives from different political parties have signed a motion asking for a moratorium on the wind-parks *Coma de Bertran* and *Serra del Tallat*, and defending the development of economic activities without interfering with local initiatives

(continued)

Table 3.1 (continued)

Social actor	Scale of action	Position regarding the wind-parks
Coordinadora por la defensa de a terra (Urgell, Conca de Barberà, Segarra, Garrigues)	Province	They think that it is not necessary to jeopardize the future of the towns to revitalize them. They are not against wind energy, but they do not approve the way the process has been carried out. They think that the solution has to be discussed by all the towns involved
Plataforma per Senan	Province	They see the projects as an undesirable gift from their neighbours. They do not share the way the process has bean carried out, and they say that to reach more equitable decisions. All the involved towns must take part in the discussion (See Town council of Senan above)
GEPEC	National	This is an environmental non-governmental organization, acting at the Catalonian level to redefine the Catalonian Energy Plan, with the participation of some social actors. They ask for a decentralized electricity production system next to the places of consumption. Regarding the location of wind-farms, they ask for special attention to the habitats of rare and endangered species, and to the biologic corridors. They ask also to apply the European Landscape Convention and for territorial equity
Enegía Hidroeléctrica de Navarra	National	The company is the promoter of one of the wind-parks. They are one of the main energy producers from RES in the Spanish territory, and one of their aims is to construct wind-parks as large as possible to provoke a change "*in the energy production culture*"
Gerrsa	National	The promoter of the *Coma Bertran* project. It was impossible to fix a meeting with them due to their reluctance to talk with people outside the government

One of the main features of the SMCE framework is that the alternatives can be constructed combining technical information and social actors' desires and preferences. In this case, sites additional to the ones originally proposed can be generated by using: (1) technical (and economic) feasibility, depending on wind availability, and (2) reduction of the visual impact for inhabitants.

Of course, there are several modes of generating alternatives; this is why continuous feed-back from the social and technical actors is needed. The evaluation criteria in general are a technical translation of social actors' preferences and desires. Their construction is a very delicate step. A summary of the whole process is presented in Table 3.2.

One should note that criteria and criterion scores are not determined directly by social actors. The impact matrix is a result of a technical translation operationalized by the scientific team. Even if the criteria are exactly those agreed with the social actors, the determination of the criterion scores is independent of their preferences. For example, an interest group can accept the use of a criterion measuring the effects of the various alternatives on employment, but the determination of the figure cannot

Table 3.2 Evaluation Criteria as a Translation of Social Actors' Desires and Preferences

Criteria	Social Desires and Preferences
Possible impact on other economic activities	Some people are worried about the consequences for the tourism sector in the long run and for residential value.
Land-owners' income	To generate additional income for farmers.
	To stabilize income.
	To improve the quality of life.
	There is a concern about who will profit-local or external owners.
	To avoid wealth concentration.
Distribution of income	To avoid concentration of revenue.
	To boost local development.
Municipalities' income	To increase the municipalities' budget
Visual impact	To offer more social services from the city council
	To avoid the industrialization of mountain areas
	To maintain rural population
	To protect tourism in the long term
Number of jobs	To attract and keep people in the region
	To preserve rural identity
	To revitalize the economic dynamics
	To avoid a decrease in land/property values of the region
Deforestation	To minimize the disturbance of ecosystems
	To avoid soil erosion
Noise annoyance	To protect human health
	To minimize the impact on wildlife
Avoided CO_2 emissions	To meet commitments to reduce emissions
Installed capacity	To promote a larger share of renewable energies into electricity production
	To guarantee economic viability

be (at least completely) controlled by them. This is the main reason why it is advisable to combine a social impact matrix with the technical impact matrix. While the latter is a good indication of the technical compromise solution, the former is a practical tool to search for social compromise; Chap. 8 is devoted to this issue. Recent real-world examples can be found in Gamboa (2003), who deals with the possible construction of one of the world's biggest smelter aluminium plants in Chilean Patagonia and the conflict surrounding it; and in Russi (2004), who combines social and technical solutions by constructing different impact matrixes for social actors in the framework of rural electrification by means of solar energy.

3.4 Social Multi-Criteria Evaluation as a Framework for Applying Public and Social Choice to Real-World Problems

The classical relationship between decision-maker and analyst often assumed by multi-criteria decision theory, and the related concept of "*decision aid*" as a learning process for the actors involved, seems most adequate in situations such as those defined as academic science and professional consultancy by Funtowicz and Ravetz (see Fig. 3.6). Since this process is best suited to the search for a technical compromise solution, I call it a "*technocratic approach*".

All the arguments and convictions discussed in Part A have led me to develop the concept of Social Multi-Criteria Evaluation (SMCE), the very essence of which is the recognition that (see Fig. 3.7):

- A multi-criteria framework is a very efficient means of applying *multi/inter-disciplinary* approach.
- Science for policy implies *responsibility* on the part of scientists towards society as a whole and not just towards a mythical decision-maker.
- *Public participation* is a necessary but not a sufficient component. Participation techniques are a tool for improving knowledge of the problem at hand, and not for eliciting input for uncritical use in the evaluation process. Social participation does not imply lack of responsibility.
- *Ethical judgements* are unavoidable elements of the evaluation exercise. These judgements always influence the results heavily. As a consequence, *transparency* in the assumptions used is essential.
- In this framework, mathematical aggregation conventions naturally play an important role, i.e. in assuring that the rankings obtained are *consistent* with the information and the assumptions used.[11]

---

[11] I insist on the importance of the algorithmic component in SMCE. Indeed, I have used the term *"non-algorithmic" multi-criteria evaluation* as an implementation tool for the incommensurability principle (Martinez-Alier *et al*., 1998). This term was intended to emphasize the importance of the decision-making process however I think it was an unfortunate choice, since it gives the impression that the algorithmic component is not useful at all.

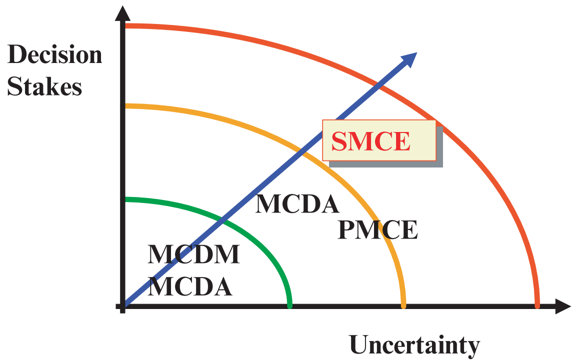


Fig. 3.6 Multi-Criteria Approaches in Relation to the Funtowicz–Ravetz Classification of Science for Policy

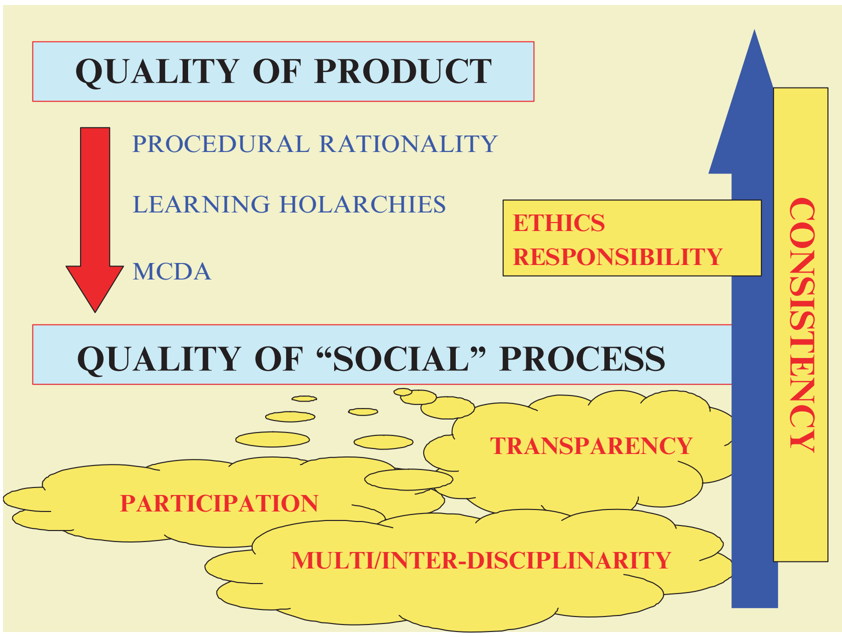


Fig. 3.7 Synthesis of a social multi-criteria evaluation process

This discussion leads on to the need to define the concept of *evaluation as the combination of representation, assessment* and *quality control connected with a given policy problem in relation to a given objective*.[12] This is why I use the term “*multi-criteria evaluation*” and not “*multi-criteria decision*” when a social context is implied.

[12] This definition has been developed following discussions with M. Giampietro.

As a conclusion, social multi-criteria evaluation can be considered an approach which is:

1. *Multi-inter-disciplinary*, to respect the plurality of scientific points of view.
2. *Participative*, to obtain as much input as possible from the general public.
3. *Transparent*, to make the assumptions adopted in the study clear.
4. *Consistent*, to assure that the results really proceed from the assumptions adopted.

Properties 1–3 are clearly dependent on the decision process; property 4 is more connected to the mathematical models used. In the next chapters in Part B I will try to isolate some properties that might be considered desirable for a discrete multi-criteria method to guarantee consistency in the framework of SMCE. Of course, in another framework, e.g. stock exchange investments, these properties might easily be irrelevant or even undesirable.

Part B
Consistency in Social Multi-Criteria Evaluation

"We should not expect ever to utilize in practice all the motive power of combustibles. The attempts made to attain this result would be far more harmful than useful if they caused other important considerations to be neglected. The economy of the combustible is only one of the conditions to be fulfilled in heat-engines. In many cases it is only secondary. It should often give precedence to safety, to strength, to the durability of the engine, to the small space which it must occupy, to small cost of installation, etc. To know how to appreciate in each case, at their true value, the considerations of convenience and economy which may present themselves; to know how to discern the more important of those which are only secondary; to balance them properly against each other; in order to attain the best results by the simplest means; such should be the leading characteristics of the man called to direct, to co-ordinate the labours of his fellow men, to make them co-operate towards a useful end, whatsoever it may be".

Sadi Carnot – *Thoughts on the motive power of fire, and on machines suitable for developing that power (original version: Réflexions sur la puissance motrice du feu sur les machines propres a développer cette puissance, Bachelier libraire, Paris, 1824), closing paragraph.*

Chapter 4
The Issue of Consistency: Basic Methodological Concepts

4.1 Dominance and Efficiency in Social Multi-Criteria Evaluation

In a discrete multi-criteria problem, there is a range of multi-criteria problem formulations which may take any of the following forms (Roy, 1985, 1996):

(α) The aim is to identify one and only one final alternative.
(β) The aim is the assignment of each alternative to an appropriate predefined category according to what one wants it to become afterwards (for instance, acceptance, rejection or delay for additional information).
(γ) The aim is to rank all feasible alternatives according to a total or partial preorder.
(δ) The aim is to describe relevant alternatives and their consequences.

Clearly the steps required in such a process demand a number of arbitrary, unavoidable subjective decisions. The degree of the subjective component may vary but it is always present. The consequence is that a complete axiomatization of a multi-criteria aggregation convention, i.e. a multi-criteria method, is quite difficult (Arrow and Raynaud, 1986). To deal with the problem, correctly identified by Arrow and Raynaud, two main approaches can be distinguished:

1. The attempt to find out under which specific circumstances each method might be more useful than others, i.e. for finding the right method for the right problem (see e.g. Guitouni and Martel, 1998).
2. The attempt to look for a complete set of formal axioms that could be attributed to a specific method (e.g. Arrow and Raynaud, 1986; Vincke, 1994).

I will try to isolate some properties that might be considered desirable for a discrete multi-criteria method in the framework of SMCE. Let us start with a principle that carries considerable epistemological implications: the principle that *dominated*

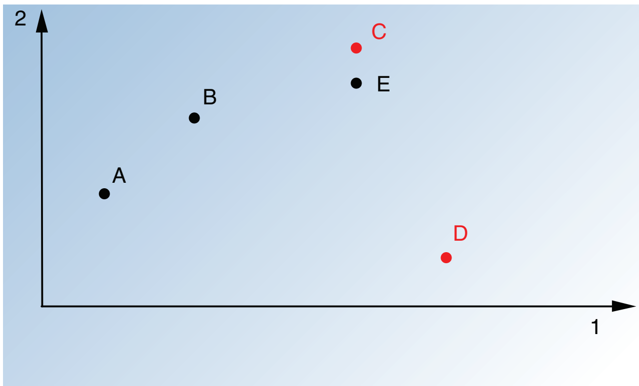


Fig. 4.1 Efficiency in a two-dimensional case

alternatives[1] can be ignored in an evaluation exercise and thus that only *efficient alternatives* have to be taken into account.

The concept of efficiency can easily be illustrated graphically (see Fig. 4.1 which refers to a two-criteria state space). Alternative ***C*** performs better than ***B*** in all respects and hence is preferred to ***B***. The same can be said for ***B*** compared with ***A***. Thus only ***C*** and ***D*** are efficient alternatives.

It has to be noted that efficiency does not imply that every efficient solution is necessarily to be preferred above every non-efficient solution; e.g. the non-efficient alternatives ***A*** and ***B*** are preferable to the efficient alternative ***D*** if the second criterion would receive a high priority compared to the first criterion (Nijkamp et al., 1990).

The principle that inefficient solutions may be ignored (often presented as a simple technical step) presupposes acceptance of the following assumptions:

(1) The assumption that all the relevant criteria have been identified. If relevant criteria are omitted, there are potential opportunity costs associated with assuming that it is safe to ignore dominated alternatives.
(2) The assumption that only the best alternative has to be identified (α problem formulation). Since the "second best" may have been eliminated during the technical screening, if more than one action has to be found, the elimination of the "inefficient" action may result in an opportunity loss (it should be noted that if the best action is removed from the set of feasible alternatives, then the second best becomes a member of the non-dominated set) (Bogetoft and

[1] ***Dominance***: an action ***a*** dominates an action ***b*** if ***a*** is at least as good as ***b*** for all the criteria taken into consideration, and much better than ***b*** for at least one criterion.

Efficient solution: an action ***a*** is efficient if there is no action ***b*** belonging to the set of alternatives taken into account and dominating ***a***.

Pruzan, 1991). If one is interested in the γ problem formulation, then dominated alternatives cannot be eliminated. It has to be remembered that in sustainability policies, it is often much more useful to have a ranking of policy options than to select just one. In fact sometimes the first alternative in the ranking can also be the most controversial from a social point of view. Thus in the context of sustainability dialectics, it may be more useful to implement the policy option that ranks second (and so technically is not "bad") but might reduce social conflicts.

(3) A third problem is connected with the question: how relevant are "irrelevant" alternatives? The axiom of "the independence of irrelevant alternatives" states that the choice made from a given set of alternatives *A* depends only on the ordering made with respect to the alternatives in that set. Alternatives outside *A* (irrelevant since the choice must be made within *A*) should not effect the choice inside *A*. The issue of the independence of irrelevant alternatives is particularly important and tricky when pair-wise comparisons are used[2].

To explain this point, let us imagine a football championship. To determine the winner all the teams have to compete pair-wise. Then we need to assess the performance of each team with respect to all the others, e.g. how many times a given team won or lost the match. By using this information, we can finally determine who won the championship. Let us now imagine that when the championship is about to end and team ***X*** is going to win (e.g. Barcelona), a new team ***Y*** is created (e.g. in Madrid). Would it be acceptable to allow this new team ***Y*** to play directly with ***X***? Would the supporters of team ***X*** accept that if ***Y*** wins, then ***Y*** will also win the championship? Of course not.

This example seems to give a clear answer to our problem, but let us now imagine that instead of ranking football teams, our problem is to buy a new car. If we are on the point of buying car ***A*** and a new car ***Z*** enters the market, can we simply compare ***A*** with ***Z*** or do we have to make all the pair-wise comparisons again? Now the answer is less clear cut. Moreover, let us imagine that the ranking at time *T* (without ***Z***) ranks car ***A*** better than ***B*** and that at time *T + 1* (when ***Z*** is considered in the pair-wise comparisons) ***B*** is ranked better than ***A*** just because ***Z*** is taken into consideration. Can this result be accepted? To answer this question in a definitive manner is very delicate matter. What is definite is that if pair-wise comparisons are used, then it has to be accepted that the alternative ***Z***, irrelevant for the evaluation between ***A*** and ***B***, can indeed change the relative evaluation of ***A*** and ***B***. This phenomenon is called "rank reversal".

From these simple examples we can draw some conclusions:

- When pair-wise comparisons are used, this information is not sufficient to derive a consistent ranking. It is necessary to exploit the relationships among all alternatives too. As a consequence no alternative is irrelevant.

---

[2] For a discussion of the definition of the axiom of independence of irrelevant alternatives see Chap. 6.

- If the set of alternatives is dynamic, i.e. new alternatives enter the evaluation process, all the pair-wise comparisons have to be made again. It is not possible to compare the new alternative only with the one that came first in the ranking.
- The principle that the final ranking depends on the relationship between the whole set of alternatives may cause the effect of rank reversal.

(4) Finally, a dominated action may be slightly worst than an efficient action, if indifference and/or preference thresholds are used, the two actions could then present an indifference relation (e.g. ***C*** and ***E***). This point will be further developed later on in the chapter.

As a conclusion we can state that in SMCE applications it is better to use aggregation procedures that do not a priori, exclude dominated alternatives.

4.2 Preference Modelling in SMCE

When considering two alternatives ***a*** and ***b*** according to each evaluation criterion, three situations are possible: (1) the score of ***a*** could be exactly the same as the score of ***b***, as a consequence ***a*** and ***b*** are indifferent. (2) The score of ***a*** could be better than the score of ***b***, thus ***a*** is preferred to ***b***, or (3) the reverse i.e. ***b*** might be better than ***a***. When the score of one alternative is better than that of another, a question arises: is the difference between ***a*** and ***b*** really significant? For example, in comparing the potential cost of two different routes for a new road, does a difference of a few hundred Euro really imply than the cheapest option has to be preferred? To avoid this kind of paradox, indifference and preference thresholds have been introduced, i.e. thresholds that indicate the degree of difference up to which two options have to be considered equivalent and from which degree of difference a preference relation exists.

More formally, given a set of evaluation criteria $G = \{g_m\}$, $m = 1, 2, \ldots, M$, and a finite set $A = \{a_n\}$, $n = 1, 2, \ldots, N$ of potential alternatives (actions), let us start with the simple assumption that the performance (i.e. the criterion score) of an alternative a_n with respect to a judgement criterion g_m is based on an *interval or ratio* scale of measurement. For simplicity of exposition, the assumption is made here that a higher value of a criterion is preferred to a lower one (the higher, the better).

The famous baldness paradox in Greek philosophy (i.e. how many hairs does one have to cut off to make a person with hair bald?), then Poincaré (1935, p. 69) and finally Luce (1956) all made the point that the transitivity of indifference relation is incompatible with the existence of a sensibility threshold below which an agent either does not perceive a difference between two elements or refuses to declare a preference for one or the other. Luce was the first to discuss this issue formally in the framework of preference modelling[3]. Mathematical characterizations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

By introducing a positive indifference threshold q the resulting preference model is the *threshold model:*

$$\left\{\begin{array}{lll} a_j P a_k & \Leftrightarrow & g_m(a_j) > g_m(a_k) + q \\ a_j I a_k & \Leftrightarrow & \left| g_m(a_j) - g_m(a_k) \right| \leq q \end{array}\right\} \tag{4.1}$$

where a_j and a_k belong to the set A of alternatives and g_m to the set G of evaluation criteria.

Real life experiments show that there is often an intermediary zone within which an agent hesitates between indifference and preference. This observation led to the so-called *double variable threshold model*, where indifference and preference thresholds vary according to the chosen scale, that is:

$$\left\{\begin{array}{lll} a_j P a_k & \Leftrightarrow & g_m(a_j) > g_m(a_k) + p\big(g_m(a_k)\big) \\ a_j Q a_k & \Leftrightarrow & g_m(a_k) + p\big(g_m(a_k)\big) \geq g_m(a_j) > g_m(a_k) + q\big(g_m(a_k)\big) \\ a_j I a_k & \Leftrightarrow & \left\{\begin{array}{l} g_m(a_k) + q\big(g_m(a_k)\big) \geq g_m(a_j) \\ g_m(a_j) + q\big(g_m(a_j)\big) \geq g_m(a_k) \end{array}\right. \end{array}\right\} \tag{4.2}$$

Relation Q has been called "weak preference" by Roy (1985, 1996). It translates the decision-maker's hesitation between indifference and preference and not "less strong" preference as its name might lead to believe. A criterion with both preference and indifference thresholds is called a pseudo-criterion (see Fig. 4.2).

A *pseudo-order structure* is a double threshold model upon which the following consistency condition is imposed:

$$g_m(a_j) > g_m(a_k) \Leftrightarrow \left\{\begin{array}{l} g_m(a_j) + q\big(g_m(a_k)\big) > g_m(a_k) + q\big(g_m(a_k)\big) \\ g_m(a_j) + p\big(g_m(a_k)\big) > g_m(a_k) + p\big(g_m(a_k)\big) \end{array}\right. \tag{4.3}$$

[3] The so-called "*Luce Paradox*" concerns cups of tea or coffee (according to the tastes of Anglo-Saxon or Latin people). Everybody would probably agree that in comparing various cups with different sugar content, there is always a maximum amount of sugar we would like in our cup. If a sufficiently large number of cups is considered and if no indifference threshold is used, due to the transitivity of the indifference relation property, the paradoxical result applies that *a cup with no sugar and a cup with only sugar are indifferent for the decision-maker*.

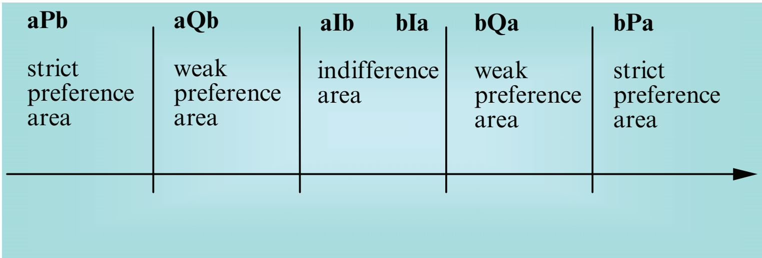


Fig. 4.2 Structure of a criterion with indifference and preference thresholds

In this framework, the PROMETHEE methods (Brans et al., 1984; 1986) are a very interesting attempt to introduce various parameters for indifference and preference thresholds with physical or economic interpretation for a single decision-maker. The preference modelling of the PROMETHEE methods is based on a function $F_j(a, b)$ which is a number between *0* and *1* which increases if $g_j(a) - g_j(b)$ is greater than zero. It equals zero if $g_j(a) \leq g_j(b)$. In order to estimate the various possible forms of $F_j(a, b)$, the decision-maker can choose among six different forms (see Chap. 5).

A problem inherent to all modelling procedures based on the notion of a pseudo-criterion is that they display a serious *lack of stability*. Such undesirable discontinuities make a sensitivity analysis (or robustness analysis) necessary; however, this important analytical step is very complex because of the combinatory nature of the various sets of data (Saltelli et al., 2004). One should combine variations of two thresholds (indifference and preference) and *k* possible scores of the *M* criteria (Perny and Roy, 1992).

A solution to this problem may be found in the concept of *valued preference relations and fuzzy sets*, i.e. a preference relation in which it is necessary to assign to each ordered pair of alternatives (a_j, a_k) a value $v(a_j,a_k)$ representing the "*strength*" or the "*degree of preference*" (Fishburn, 1970, 1973a; Roubens and Vincke, 1985) (see Sect. 4.3.2 and Chaps. 5 and 7).

4.3 Quantitative, Qualitative and Mixed Criterion Scores

4.3.1 Measurement Scales

The preference modelling we looked at in the previous section is based on the notion of "intensity of preference". This implies that the criterion scores considered are necessarily measured on an interval or ratio scale. However, in many real-world problems it is often necessary to use less ambitious measurements. The extension of multi-criteria modelling to qualitative information is the subject of this section.

Let us start by clarifying what a measurement scale is. The process of grouping individual observations into qualitative classes is measurement at its most primitive level. Sometimes this is called categorical or *nominal scaling*. The set of equivalence classes itself is called a nominal scale. The word *measurement* is usually reserved for the situation in which a number is assigned to each observation; this number reflects a magnitude of some quantitative property (how to assign this number constitutes the so-called *representation problem*).

There are at least three kinds of numerical measurement that can be distinguished: these are the *ordinal scale, the interval scale and the ratio scale* (Winkler and Hays, 1975; Roberts, 1979; Vansnick, 1990). Imagine a set of objects O, and suppose that there is some property that all objects in the set possess, such as value, weight, length, intelligence or motivation. Furthermore, let us suppose that each object o has a certain amount or degree of that property. In principle it is possible to assign a number $t(o)$, to any object $o \in O$, standing for the amount that o actually "has" of that characteristic. Ideally, to measure an object o, we would like to determine this number $t(o)$ directly. However because this is not always possible, it is necessary to find a procedure for pairing each object with another number, $m(o)$, that can be called its *numerical measurement*. The measurement procedure used constitutes a function rule $m: O \rightarrow R$, telling how to give an object o its $m(o)$ value in a systematic way. Measurement operations or procedures differ in the information that the numerical measurements themselves provide about the true magnitudes.

Let us suppose that there is a measurement procedure or rule for assigning a number $m(o)$ to each object $o \in O$, and suppose that the following statements are true for any pair of objects o_1 and $o_2 \in O$:

$$\begin{cases} m(o_1) \neq m(o_2) \ \textit{only if} \quad t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \ \textit{only if} \quad t(o_1) > t(o_2) \end{cases} \tag{4.4}$$

In other words, from this rule it is possible to say that if two measurements are unequal, and if one measurement is larger than another, then one magnitude exceeds another. Any measurement procedure for which (4.4) applies is an example of *ordinal scaling*, or measurement at the ordinal level.

A fundamental point in measurement theory is *uniqueness of scale*, by which is meant the admissible transformations of scale that allow the truth or falsity of the statement involving numerical scales to remain unchanged (*problem of meaningfulness*). In the case of an ordinal scale, it is unique up to a strictly monotonically increasing transformation (with infinite degrees of liberty). Other measurement procedures associate objects $o \in O$ with a real number $m(o)$ allowing much stronger statements to be made about the true magnitudes from the numerical measurements. Suppose that the statement of (4.5) is true:

$$\begin{cases} m(o_1) \neq m(o_2) \text{ only if } \quad t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \text{ only if } \quad t(o_1) > t(o_2) \\ t(o) = x \ \textit{iff} \ \ m(o) = ax + b, \text{where } \ a \in R^+ \end{cases} \tag{4.5}$$

That is, the numerical measurement *m(o)* is some *affine function* of the true magnitude *x*. When (4.5) applies, the measurement operation is called *interval scaling*, or measurement at the *interval-scale level*. An interval scale is unique up to a positive affine transformation (with two degrees of freedom).

When measurement is at the interval-scale level, any of the ordinary operations of arithmetic can be applied to the differences between numerical measurements, and the results can be interpreted as statements about *magnitudes* of the underlying property. The important part is the interpretation of a numerical result as a *quantitative statement* about the property shown by the objects. This is not possible for ordinal-scale numbers, but can be done for differences between interval-scale numbers. Interval scaling is about the best we can do in most scientific work, and even this level of measurement is all too rare in the social sciences. However, especially in the physical sciences, it is sometimes possible to find measurement operations making the statement of (4.6) true:

$$\begin{cases} m(o_1) \neq m(o_2) \ only\,if \quad t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \ only\,if \quad t(o_1) > t(o_2) \\ t(o) = x \ iff \ m(o) = ax, where \ a \in R^+ \end{cases} \tag{4.6}$$

When the measurement operation defines a function such as the statement contained in (4.6), then measurement is said to be at the *ratio-scale level*. For such scales, ratios of numerical measurements are unique and can be interpreted directly as ratios of magnitudes of objects. A ratio scale is unique up to a linear transformation; in this case, the ratio between differences is unique (with only one degree of liberty).

Of course, the fewer the admissible transformations of a scale, the more meaningful are the statements involving that scale. From this point of view, it is better to have a ratio scale than an interval scale, and it is better to have an interval scale than an ordinal scale.

4.3.2 Uncertainty in the Criterion Scores

Ideally the information available for a policy problem should be precise, certain, exhaustive and unequivocal. But in a real-world situation, it is often necessary to use information which does not have these characteristics and thus to deal with *uncertainty of a stochastic and/or fuzzy nature* in the data. Let us then introduce *a more realistic assumption*, i.e. that the set of evaluation criteria $G=\{g_m\}$, $m = 1, 2,\ldots, M$, on the set $A=\{a_n\}$, $n = 1, 2,\ldots, N$ of potential alternatives may include crisp (i.e. impacts measured on ordinal, interval or ratio scales), stochastic or fuzzy criterion scores.

If it is impossible to establish exactly the future state of the system under study, a stochastic uncertainty exists; this type of uncertainty is well known in decision theory and economics, where it is called "*decisions under risk*" (Knight, 1921;

Savage, 1954; Rothschild and Stiglitz, 1970, 1971; Arrow, 1971; Diamond and Stiglitz, 1974; Allais, 1979; Kahneman and Tversky, 1979; Kahneman et al., 1981; Machina, 1987; French, 1986; Markowitz, 1989; Moser, 1990). Applications of this concept in a multi-criteria framework can be found in D'Avignon and Vincke (1988), Martel and Zaras (1995) and Rietveld (1989), among others.

Another framing of uncertainty, called *fuzzy uncertainty*, focuses on the ambiguity of information in the sense that the uncertainty does not concern the occurrence of an event but the event itself, which cannot be described unambiguously. This situation is very common in human systems. These systems are *complex systems* characterized by subjectivity, incompleteness and imprecision. Zadeh (1965) writes: "as the complexity of a system increases, our ability to make a precise and yet significant statement about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics" (*incompatibility principle*). Fuzzy set theory is a mathematical theory for modelling situations in which traditional modelling languages that are dichotomous in character and unambiguous in their description cannot be used.

Fuzzy sets, as formulated by Zadeh (1965), are based on the simple idea of introducing a degree of membership of an element with respect to some sets. Fuzzy uncertainty considers all cases between 0 (non-membership) and 1 (complete membership), and it is represented by means of the membership functions.

Let us assume that the symbol U means the entire set (Universe of discourse). In classical set theory, given a subset Ω of U, each element $x \in U$ satisfies the condition: either x belongs to Ω, or x does not belong to Ω. The subset Ω is represented by a function $f_\Omega : \mathrm{U} \to [0,1]$. The function f_Ω is called a characteristic function of the set Ω. Fuzzy sets are then introduced by generalizing the characteristic function f_Ω. Let U again be a universe of discourse with $x \in U$. Then a fuzzy set $\tilde{\Omega}$ in U is a set of ordered pairs $\{\mathrm{x}, \mu_{\tilde{\Omega}}(x)\}$, $\forall x \in U$ where $\mu_{\tilde{\Omega}}$: $U \to \Theta$ is membership function which maps $x \in U$ into $\mu_{\tilde{\Omega}}(x)$ in a totally ordered set Θ (called the membership set) and $\mu_{\tilde{\Omega}}(x)$ indicates the grade of membership of x in $\tilde{\Omega}$. Generally, thez membership set is restricted to the closed interval [0, 1]. A fuzzy set is completely determined by its membership function.

A very useful concept for quantifying vagueness in criterion scores is that of a fuzzy number. A *fuzzy number* is simply a fuzzy set in the real line and is completely defined by its membership function such as $\mu_x : R \to [0,1]$. For computational purposes, this definition generally is restricted to those fuzzy numbers which are both normal and convex (Munda, 1995):

$$\textit{Normality:}\ \sup\{\mu(x)\}=1 \quad \text{with } x \in \mathrm{R}.$$

$$\textit{Convexity}: \begin{array}{l} \mu\{\lambda x_1 + (1-\lambda)\, x_2\} \geq \min\,\{\mu(x_1), \mu(x_2)\} \\ \textit{with } x \in R \textit{ and } \lambda \in [0,1] \end{array}$$

The requirement of convexity implies that the points of the real line with the highest membership values are clustered around a given interval (or point). This fact allows one easily to understand the semantics of a fuzzy number by looking at its distribution and to associate it with a properly descriptive syntactic label, e.g. "*approximately 135*" (see Fig. 4.3). The requirement of normality implies that, among the points of the real line with the highest membership value, there exists at least one which is completely compatible with the predicate associated with the fuzzy number. If a closed interval exists in which the membership function is equal to 1, it is called a flat fuzzy number.

A general type of fuzzy number is the so-called *L-R fuzzy number*; it is defined as follows:

$$\mu(x) = \begin{cases} \dfrac{F_L(x-m)}{\alpha} & \text{if } -\infty < x < m,\ \alpha \in R^+ \\ 1 & \text{if } x = m \\ \dfrac{F_R(x-m)}{\delta} & \text{if } m < x < +\infty,\ \delta \in R^+ \end{cases} \tag{4.7}$$

where m, α, δ, are the "middle" value, the left-hand and the right-hand variation, respectively. $F_L(x)$ is a monotonically increasing membership function and $F_R(x)$, not necessarily symmetrical to $F_L(x)$, is a monotonically decreasing function.

Fuzzy set theory also supplies a framework for representing "*qualitative information*" (measurements on an ordinal scale) by means of the concept of "*linguistic variable*". Human judgements, especially in linguistic form, appear to be plausible and natural representations of cognitive observations. We can explain this phenomenon by *cognitive distance*. A linguistic representation of an observation may require a less complicated transformation than a numerical representation, and therefore less distortion may be introduced in the former than in the latter.

Formally, a linguistic variable is represented by a quintuple (*X, T(x), U, G, M*) (Leung, 1988; Munda, 1995) where:

X is the name of the variable, e.g. "Evaluation Score";

T(x) is the term set of *X*, finite or infinite, such as good, very good and so on, in a universe of discourse ***U***. A *primary term* in *T(x)* is a term whose meaning must be defined a priori, and which serves as a basis for the computation of the meaning of the non-primary terms in *T(x)*;

G is a syntactic rule by which the non-primary terms in the term set are generated. It is possible to use a context free grammar or a regular grammar. In *G* it is possible

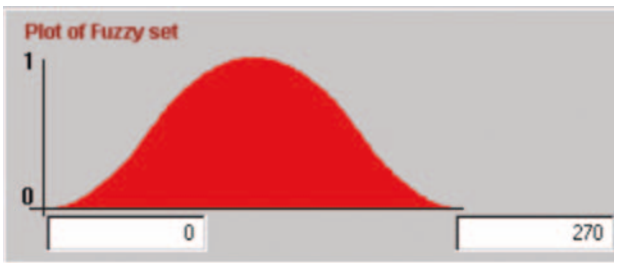


Fig. 4.3 A symmetric fuzzy number "approximately 135"

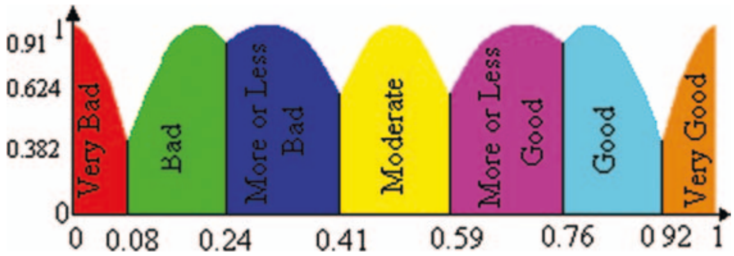


Fig. 4.4 Evaluation scores represented by means of linguistic variables

to find *primary terms, hedges* (not, very, more or less, etc.), *relations* (better than, worse than, etc.), *conjunctions* (and, etc.), and *disjunctions* (or, etc.);
M is a semantic rule which associates each term with its meaning (a fuzzy subset in *U*). Through *M*, a compatibility (membership) function $\boldsymbol{\mu}$:$U \rightarrow [0, 1]$ is constructed (e.g. $\boldsymbol{\mu}_{good}$ shows the degree to which a numerical score is compatible with the concept of *good* and equivalently $\boldsymbol{\mu}_{good}$ may be viewed as the membership function of the fuzzy set *good*, see Fig. 4.4 for an example). The basic assumption behind the idea of linguistic variables is that a kind of "*intensity of preference*" exists for the representation of qualitative impact scores, i.e. the measurement scale is not purely ordinal.

Type 1

If linguistic variables whose meaning can be translated into a measure on an interval or ratio scale are present in a decision model, generally it is because of a lack of information or of the right instrument of measurement. Therefore, we have a qualitative evaluation of a variable that in theory could be measured on an interval or ratio scale. So it is reasonable to suppose that it is possible to transform the qualitative information into quantitative with a certain degree of precision. The parameters, necessary in this case, may easily be established, because this is a case of the so-called ""*informational fuzziness*" depending mostly on the subjective culture of the person in charge of the evaluation. For example, the proposition "that man is tall" may have different meanings for different people, but everybody can easily indicate the "tolerance interval" of his own evaluation. Such a representation takes into account some "labels" of the term set, e.g. young, young and/or very young, etc., and the problem is to find which values of the base variable are compatible with these terms.

Type 2

In the case of linguistic variables with no meaning on an interval or ratio scale, the qualitative information does not depend on lack of information, but on the nature of information that is essentially fuzzy (*intrinsic fuzziness*). Therefore, whereas in the other cases the stochastic and the fuzzy representation may be competitive, in this case the fuzzy representation is the only one possible. For example, if linguistic propositions (like "beautiful flower" or "quality of life") clearly have no quantitative base variable, how can we represent them? It seems that there is a set of hidden and fuzzy standards in one's mind with a justification for this type of concepts, but they are more than a human being can rationally handle simultaneously (Zimmermann and Zysno, 1983).

A first approach to this problem may be to try *to decompose the concept* that one wants to represent into a series of quantitative measurable variables. This approach presents two main problems, viz. the explication of the quantitative variables, and the aggregation procedure to be used.

A second approach is to define an *artificial quantitative base variable*, assuming that the real space is one-dimensional. The interval of the real space is chosen from [–1, 1], [0, 1], or [0, 10], etc., as desired and can be subdivided into a series of fuzzy sets representing linguistic values (e.g. very negative, moderately positive, very positive, etc.). Then, a link or mapping between quantitative (numerical) and qualitative (linguistic) values is established. This "*direct estimation*" approach has been criticized because of its lack of theoretical foundation. Recently, some psychologists (Norwich and Turksen, 1982, 1984; Wallsten et al., 1986) have developed a graded pair comparison procedure, which allows for simultaneous testing of the necessary axioms, scaling of the responses in order to obtain memberships and tests of goodness of fit. Subsequently, empirical experiments have demonstrated a high level of similarity between membership values determined through graded pair comparison and direct magnitude estimation. Thus it seems that a theoretical justification can be established for the quantification of the vague meanings of inexact linguistic terms by means of direct estimation.

A third interesting approach follows from the notion of the *type 2 fuzzy set*. A type 2 fuzzy set is a fuzzy set whose membership values are fuzzy sets on [0, 1]. This corresponds to the case in which the decision-maker is not able (or not willing) to characterize the grade of membership by an exact number, but gives an evaluation such as "the grade of membership is high, medium", etc. It is always possible to define a fuzzy set of type $n = 2, 3, \ldots$, if its membership function is a mapping from U to a set of fuzzy subsets of type $n - 1$; therefore, it is possible that in order to reduce the fuzziness, many transformations will be required, thus diminishing drastically the computational efficiency of the algorithm.

Another possible approach has been developed primarily in the field of psychological research (Hersh et al., 1979), called the "*yes-no paradigm*". In this approach, an element x of the universe of discourse U, is presented to the subject and he has to decide whether the element is a member of A, A being a fuzzy subset of U. The fraction of positive responses across replications (within or across subjects) is considered a measure of $\mu_A(x)$. It has been noted that the main problem of this approach is that it confounds fuzziness with response variability and that it can be interpreted as an indication that words have various, but nevertheless precise meanings for different people and/or at different times.

4.3.3 Dealing with Mixed Information on the Criterion Scores

Let us now try to find other desirable properties for SMCE. It has been argued that the presence of qualitative information in evaluation problems concerning socio-economic and physical planning is a rule, rather than an exception (Nijkamp et al., 1990). The idea of *technical incommensurability* implies that there is a clear need for methods that are able to take into account information of a "mixed" type (both qualitative and quantitative criterion scores plus various possible sources of uncertainty).

In multi-criteria decision theory, a clear distinction is made between quantitative and qualitative methods. Essentially, there are two approaches for dealing with qualitative information: a direct and indirect (Nijkamp et al., 1990; Munda et al., 1994a). In the *direct approach*, qualitative information is used directly in a qualitative evaluation method; in the *indirect approach*, qualitative information is first transformed into cardinal (i.e. interval or ratio scale of measurement), then, one of the existing quantitative methods is employed.

Cardinalization is especially attractive in the case of available information of a "mixed type". In this case, the application of a direct method would usually imply that only the qualitative content of all available (quantitative and qualitative) information be used, which would give rise to an inefficient use of this information. In the indirect approach, this loss of information is avoided; the question is, however, whether there is a sufficient basis for the application of a certain cardinalization scheme. Two examples of cardinalization of a qualitative evaluation matrix are the *expected value method* (Rietveld, 1989) and *multi-dimensional scaling techniques* (Kruskal, 1964; Nijkamp, 1979; Keller and Wansbeek, 1983).

(1) In the expected value method ordinal criterion scores are replaced by quantitative scores by using a transformation procedure aiming at deriving the centroid of a convex polyhedral set S consistent with the underlying ordinal information. It is proved that the expected value is identical to the centroid of the polyhedron S. One has to note that the statistical distribution, normally used as a transformation procedure, gives rise to a linear cardinalization curve. If one wants to derive a concave structure, it can be proved that the probability density function has to be rewritten in a more complex way. However, by means of integration, Rietveld (1989) proves that a series of precise expected values can be obtained. These results can again be interpreted in terms of the centroid of a certain polyhedron. Weighted summation can be used to rank the alternatives.
(2) A completely different approach to cardinalization is the use of multidimensional scaling techniques. Such techniques aim to transform qualitative data input into a cardinal output of lower dimensionality. In a sense, a scaling technique may be regarded as a kind of qualitative principal component analysis. It is clear that several concepts from multidimensional scaling analysis may also be applicable to ordinal multiple criteria problems. For instance, one may use a scaling technique in order to transform a qualitative evaluation matrix into a cardinal matrix with lower dimensionality. Then the cardinal configuration of the initial

qualitative matrix provides a metric picture of the Euclidean distances both between the alternatives and between the effects. This is a normal standard operation. A limitation of this elegant but complex evaluation approach is that it requires a sufficient number of degrees of freedom to allow a multi-dimensional scaling. This implies that unless a sufficiently large number of evaluation criteria is used, no consistent scaling results can be obtained (Nijkamp et al., 1990).

An approach to tackling mixed information is the EVAMIX method (Voogd, 1983). EVAMIX involves the construction of two measures: one dealing only with the ordinal criterion scores and the other with the quantitative criterion scores. By making various assumptions about standardization and aggregation, several methods can be defined by which an appraisal score for each alternative can be calculated. The *most important assumptions* behind the EVAMIX approach concern the definition of the various standardization functions (at least three different techniques can be distinguished). Other assumptions concern the weights for the ordinal and cardinal criteria, and finally the additive relationship of the overall dominance measure. The global structure of the EVAMIX method is summarized in Fig. 4.5.

A problem connected to *all multi-criteria methods* that try to take mixed information into account, but which is particularly evident in the EVAMIX approach, is

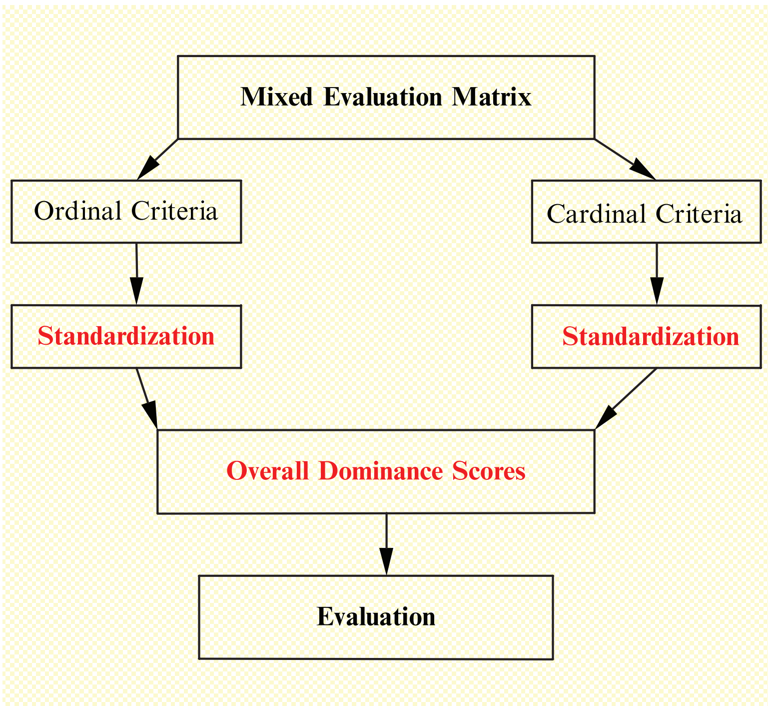


Fig. 4.5 Scheme of EVAMIX

the problem of equivalence of the procedures used in order to standardize the various criterion scores.

4.4 Compensability and the Meaning of Weights

Another sensitive point concerns compensability and the corresponding meaning of weights in multi-criteria methods. In short, the information contained in the impact matrix which is useful for solving the so-called discrete multi-criterion problem is as follows:

- *Intensity of preference* (when quantitative criterion scores are present)
- *Number* of criteria in favour of a given alternative
- *Weight* attached to each criterion.
- *Relationship* of each alternative to all the other alternatives

Combinations of this information generate different aggregation conventions, i.e. manipulation rules for the available information so as to arrive at a preference structure (see Chap. 5). In this framework, it is crucial to highlight that the aggregation of several criteria implies taking a position on the fundamental issue of compensability. *Compensability* refers to the existence of trade-offs (in the economic literature this concept is more often called "*marginal rate of substitution*"), i.e. the possibility of offsetting a disadvantage on some criteria by a sufficiently large advantage on another criterion, whereas smaller advantages would not do the same. Thus a preference relation is non-compensatory if no trade-off occurs and compensatory otherwise. The use of weights with intensity of preference is the basis of compensatory multi-criteria methods and gives the weights the meaning of *trade-offs*. Conversely, the use of weights with ordinal criterion scores is the basis of non-compensatory aggregation procedures and gives the weights the meaning of *importance coefficients*_(Keeney and Raiffa, 1976; Roberts, 1979; Bouyssou, 1986; Bouyssou and Vansnick, 1986; Vansnick, 1986; Podinovskii, 1994).

To apply the compensatory approach, one has to determine, for each criterion, a mapping $\phi_m : g_m \rightarrow \boldsymbol{R}$ which provides an interval scale of measurement. One must also assess the scaling constants in order to specify how the compensability should be achieved, given the scales ϕ_m between the different criteria (Roberts, 1979). Note that the scaling constants which appear in the compensatory approach depend on the scales ϕ_m, thus they do not characterize the intrinsic relative importance of criteria. In Box 4.1, this result is proved by using as an example, the case of the so-called "*composite indicators*" (Munda and Nardo, 2003, 2005). In Box 4.2, it is shown the empirical relevance of this methodological discussion on weights in the framework of cost–benefit analysis (Munda, 1998).

The concept of importance used here can be classified as *symmetrical importance*, that is "if we have two non-equal numbers to construct a vector in R^2, then it is preferable to place the greatest number in the position corresponding to the most important criterion" (Podinovskii, 1994, p. 241).

Box 4.1 Weights as trade-offs in composite indicators

A typical composite indicator, I, is built as follows (OECD, 2003, p. 5):

$$I = \sum_{i=1}^{N} w_i x_i$$

where x_i is a normalized variable and w_i a weight attached to x_i, with $\sum_{i=1}^{N} w_i = 1$ w_i=1 and $0 \leq w_i \leq 1$, $i = 1,2,\ldots,N$. It is clear that from a mathematical point of view a composite indicator entails a weighted linear aggregation rule applied to a set of variables. A variable x_i expresses different levels of some underlying dimensions, e.g. GDP per capita or number of publications in international journals. We assume a total strict order P_i on x_i (e.g. GDP of European countries ordered from highest to lowest). The conclusion that weights are not importance coefficients can easily be shown as follows, when differentiability is allowed (adapted from Vincke, 1992, pp. 36–37). Suppose that country a is evaluated according to some variables $(x_1(a),\ldots,x_n(a))$, then the *substitution rate at a*, of the variable j with respect to the variable r (taken as a reference variable) is the amount $S_{jr}(a)$ such that, country b whose evaluations are: $x_l(a)=x_l(b)$, $\forall l \neq j,r$; $x_j(b)=x_j(a)-1$; and $x_r(b)=x_r(a) + S_{jr}(a)$ is indifferent to country a. Therefore, $S_{jr}(a)$ is the amount which must be added to the reference variable in order to compensate for the loss of one unit from variable j. Consider now a composite indicator $I(x_1, x_2,\ldots,x_n)$ aggregating all variables and suppose that the indicator is the same for the two countries. Let $z(a)=(x_1(a), x_2(a),\ldots, x_n(a))$ and $z(b)=(x_1(b), x_2(b),\ldots,x_n(b))$. Then as a first approximation one has:

$$0 = I(z_b) - I(z_a) = \sum_{i=1}^{n} \left(\frac{\partial I}{\partial x_i}\right)_{z_a} \left(x_i(b) - x_i(a)\right)$$

$$= -\left(\frac{\partial I}{\partial x_j}\right)_{z_a} + S_{jr}(a)\left(\frac{\partial I}{\partial x_r}\right)_{z_a}$$

and manipulating

$$S_{jr}(a) = \frac{\left(\partial I / \partial x_j\right)(z_a)}{\left(\partial I / \partial x_r\right)(z_a)} \tag{4.8}$$

When the function I is a weighted sum of all the normalized variables, i.e.

$$I(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} w_i x_i \tag{4.9}$$

(continued)

Box 4.1 (continued)

from expression (4.8) one obtains:

$$S_{jr}(a) = \frac{w_j}{w_r} = \text{constant.} \tag{4.10}$$

This means that in the weighted summation case, the substitution rates are always equal to the weights of the variables up to a multiplicative coefficient. As a consequence, the estimation of weights is equivalent to that of substitution rates. But precisely this implies a compensatory logic. Thus, the relevant question to ask is in terms of a "*gain with respect to one variable allowing to compensate loss with respect to another*" and NOT in terms of "*symmetrical importance*" of variables (placing the highest number beside the most important criterion). This means that the linear aggregation rule entails weights whose meaning is necessarily compensatory and never of "*symmetrical importance*".

However common practice in constructing composite indicators, as summarized in a recent OECD report, follows the rule that: "Greater weight should be given to components which are considered to be *more* significant in the context of the particular composite indicator" (OECD, 2003, p. 10). It thus seems that current practice is fundamentally flawed.

To illustrate the issue, consider the hypothetical example presented in Table 4.1.

Table 4.1 Illustrative example with three countries and three variables

	GDP (Millions of Euro)	Populations (Number of Inhabitants)	Percentage of Protected Species
A	32,000	1,000,000	60
B	80,000	3,000,000	70
C	100,000	5,000,000	40

Consider first the measurement scale. Suppose that in the construction of a sustainability composite indicator, the trade-off between protected species and GDP is set such that a decrease of one point in the percentage of protected species can be compensated for by an increase of 100,000,000 Euro of GDP. This trade-off can be expressed as $\frac{w_{species}}{w_{GDP}} = 100,000,000$. If the measurement scale of GDP is changed and is now measured per capita, the trade-off indicated above would now be modified e.g. in "1% less of protected species can be compensated for by 100 Euro more of GDP per capita". Thus in this case one has $\frac{w_{species}}{w_{GDP}} = 100$. Since the measurement scale of the variable *protected species* has not changed, the only weight that must change value is the one attached to GDP, which in the second formulation has to

(continued)

Box 4.1 (continued)

increase considerably (since the numerator remains constant and the value of the ratio decreases).

One obvious observation might be that in a composite indicator variables are normalized and thus effects due to measurement scales should disappear. This however is not true. Consider for example, the normalization technique *distance from the group leader*, which assigns 100 to the leading alternative and other alternatives are ranked in percentage points away from the leader (see Box 5.2), that is $100\left(\frac{\text{actual value}}{\text{maximum value}}\right)$. By applying this normalization technique while keeping the original trade-off "1% decrease in species versus 100 million Euro GDP" one has to standardize the value 100 mill. Euros according to the new scale. This is equivalent to dividing this value by the score of the country with the highest GDP: $\frac{100}{100{,}000} = 10^{-3}$. If income was expressed as GDP per capita, then the trade-off would be a "1% decrease in species versus $\frac{100}{32{,}000} = 3.125 \bullet 10^{-3}$ increase in GDP per capita".[4] Again trade-offs and corresponding weights must change according to the range of variation of the measurement scale considered. One may easily check that this kind of consequence applies independently of the normalization technique chosen. The conclusion is that in the case of a linear aggregation rule, trade-offs depend on the scales of measurement, and since weights are connected to the values of trade-offs they also depend on the scales of measurement.

Clearly trade-offs can be evaluated only if one knows the quantitative scores of the variables involved without any uncertainty. On the other hand, the concept of importance is connected to the variable itself and NOT with its quantification. Whether protected species are considered *more or less important* than GDP is a quality of the variables which is independent from any measurement scale employed.

As we will see in Chap. 5, an aggregation framework allowing complete compensability and as a consequence using weights as trade-offs, is the Multi-Attribute Value Functions approach. ELECTRE II (Roy and Bertier, 1973) is an example of a method trying to minimize compensability and using weights as importance coefficients (although the existence of a veto considered as intensity of preference may give rise to some perplexity over the meaning of weights). A method that requires weights to be importance coefficients is the REGIME method when only ordinal

[4] A is the country with the highest GDP per capita with €32,000, followed by B with €26,667 and C with €20,000.

Box 4.2 Distributional weights in cost–benefit analysis

The concept of individual preference relevant to cost–benefit analysis (CBA) is the preference recorded on the market place (or which would be recorded if there were a market), and not the preference recorded by a simple political vote. This kind of "economic voting" is preferred to classical political voting procedures for various reasons. One is that "the use of money values permits some expression of the intensity of preference in the vote: it enables the individual to say how deeply he wants or does not want the project or good in question" (Pearce and Nash, 1981, p. 7).

The assumption underlying the net present value rule used in CBA is that of an additive social welfare function, such as $SW = \sum_h U_h$, where the subscript h denotes the individual to whom the utility function applies. Under the assumption that the marginal utility of money income (λ) is *identical for all individuals*, the variation of this social welfare function indicating the social worth of a project is:

$$\begin{aligned}\Delta SW &= \sum_h \sum_i \frac{\partial U_{ih}}{\partial Y_{ih}} \bullet \Delta Y_{ih} \\ &= \lambda \sum_h \sum_i P_i \Delta Y_{ih} = \lambda \sum_i P_i \Delta Y_i\end{aligned} \tag{4.11}$$

Where the h subscript denotes the individual to whom the utility function and quantity of the good Y_i apply. The translation into monetary terms is accomplished by the equation $\lambda \frac{\partial U_i}{\partial Y_i} = P_i$ where P_i is the (relative) price of good i.

The fact that intensity of preference is taken into account inside a linear aggregation rule, has the consequences, proved in Box 4.1 that weights must be considered as trade-offs. At this point a question arises: in their standard use, are the so-called *distributional weights* used as importance coefficients or as trade-offs? Let us try then to answer this question.

The function used in cost–benefit analysis can be considered an additive social welfare function, where the assumption of the constancy of the marginal utility of income across individuals needs to be accepted. Given that society is unlikely to be indifferent about various distributions of income, some ways of integrating the distributional aspects into the analysis have to be found. The most popular methodology is to introduce distributional weights explicitly; by using different weights for different social groups.

It has to be noted that the failure to use any weighting system implies making the value judgement either that the existing distribution of income is optimal and/or the change in income distribution is negligible. If, and only if, one is happy with such a value judgement, it is reasonable to use un-weighted

(continued)

Box 4.2 (continued)

market valuations to measure costs and benefits. *Therefore, there is no escape from value judgements*.

The methods used to obtain weights are based on a variety of philosophical and methodological principles. A review of the already classical scientific debate on these issues can be found in Dasgupta and Pearce (1972), Dasgupta et al. (1972), Pearce and Nash (1981) and Ray (1984). To give some examples, a possible approach for obtaining distributional weights is to observe the weights implicit in past government decisions (Weisbrod, 1968). Another possibility is an analysis of the progressivity of the income tax schedule (Krutila and Eckstein, 1958); by means of this approach the marginal rate of a tax can be converted into a surrogate for the marginal utility of income, the relevant weights are the inverses of the marginal tax rates. As a consequence the gains (or losses) of lower-income groups are weighted more heavily than the gains of high-income groups. Another approach is to scale down higher incomes and scale up lower incomes to equalize their influence on the cost–benefit outcome. Foster (1960) suggests that gains and losses should be weighted by the ratio of the average national personal income to the individual,s income. Lastly, different sets of weights can be obtained by trying to assess the likely shape and elasticity of a marginal utility of income function.

From this brief overview we can deduce that all the proposed methods for obtaining weights in CBA are based on the concept of weights as *coefficient of importance*. This result can be summarized by quoting the following sentence: "if the decision-maker considers individual 2 more "deserving" than individual 1 he will weight 2's losses more heavily than 1's gains i.e. $\lambda_2 > \lambda_1$" (Dasgupta and Pearce, 1972, p. 65). Unfortunately, since CBA is based on a completely compensatory mathematical model, as discussed in this section, weights can only have the meaning of a trade-off ratio, as a consequence *a theoretical inconsistency exists*.

criterion scores are used (Hinloopen et al., 1983). However, when mixed information is considered (Hinloopen and Nijkamp, 1990), weights are more likely to be considered as trade-offs and not as importance coefficients. In Chap. 7, I will develop an aggregation convention with the property that the manipulation rules are consistent with the notion of weights as importance coefficients.

In the framework of SMCE, one should note that compensability also has consequences for the so-called weak/strong sustainability debate (Faucheux and O'Connor, 1998). The so called "*weak sustainability*" concept states that an economy can be considered sustainable if it saves more than the combined depreciation of natural and man-made capital. "We can pass on less environment so long as we offset this loss by increasing the stock of roads and machinery, or other man-made (physical) capital. Alternatively, we can have fewer roads and factories so long as we compensate by having more wetlands or mixed woodlands or more education" (Turner et al., 1994, p. 56). The concept of "*strong sustainability*" is based on the assumption that certain sorts of natural capital are deemed critical and not readily substitutable by man-made capital.

The purpose of "*green accounting*" is to provide information on the sustainability of the economy but there is no settled doctrine on how to combine different and sometimes contradictory indicators and indexes in a way immediately useful for policy (in the sense that GDP or other macroeconomic statistics have been useful for policy) (Funtowicz et al., 1999, 2002). The expression "*Taking nature into account*" (much used both in the UN system and in the European Union) hides the tension between money valuation, and appraisal through physical indicators and indexes (which themselves might show contradictory trends). So far, the elementary question of whether the European economy is moving towards or away from sustainability cannot be answered with consensus on the indicators and the integrative framework to be used.

A point of scientific controversy in the contemporary debate concerns the use of monetary or physical indexes. Examples of monetary indexes are Daly and Cobb (1989) ISEW (Index of Sustainable Economic Welfare), Pearce and Atkinson (1993) Weak Sustainability Index, the so-called El Serafy approach (Yusuf et al., 1989). Examples of physical indexes are HANPP (Human Appropriation of Net Primary Production (Vitousek et al., 1986), the Ecological Footprint (Wackernagel and Rees, 1995), MIPS (Material Input Per unit of Service) (Schmidt-Bleek, 1994).

Although these approaches may look different, they all have at least two common characteristics:

1. The subcomponents needed for the building the aggregate index are *ad hoc*. No clear justification is given why e.g. diet enters in the computation of the ecological footprint and the generation of waste does not.
2. All the indexes are based on the assumption that a *common measurement rod* needs to be established for aggregation purposes (money, energy, space, and so on). This creates the need of making very strong assumptions on conversion coefficients to be used and on compensability allowed (i.e. until which point improved economic performance results in environmental destruction or social exclusion).

It is clear, at this stage, that since indicators reveal contradictory trends (see the urban sustainability example of Chap. 1), multi-criteria aggregation conventions can be an adequate mathematical framework avoiding the arbitrariness of conversion coefficients (for an example of application of a multi-criteria framework to the sustainability of the Ecuadorian economy see Falconi, 2002). In particular, non-compensatory and partial compensatory multi-criteria aggregation conventions are the only ones that may allow an implementation of the strong sustainability concept at both micro and macro levels of analysis.

4.5 The Use of Weights in an SMCE Framework

4.5.1 Weights and Technical Incommensurability

In the Multi-Criteria Decision Aid (MCDA) literature there are several procedures designed to elicit decision-maker's priorities in the form of weights. In short, it is possible to distinguish two main approaches (Nijkamp et al., 1990):

- *Direct estimation of weights* (trade-off method, rating method, ranking method, verbal statements, pair-wise comparisons)
- *Indirect estimation of weights* (weights based on previous choices, weights based on a ranking of alternatives, interactive estimation of weights)

All these methods present different features in terms of time needed, complexity, transparency, etc., and therefore their performance depends on the specific problem faced, but in general, the weighting of criteria is open to criticism for the following reasons:

(1) in the case of single-person decisions, an abstract hypothesis it is accepted that the decision-maker has a clear idea in her/his mind of her/his own scale of preferences, and that she/he is capable of expressing these clearly without contradiction, while concretely, the logic of the "*choice*" assumes a reduction of the confusion inevitably present in the mind of the decision-maker when he is about to face a problem (Munda, 1993). This initial state of confusion has been proved empirically; in fact, in experimental research, it has been noted that, regarding the psychological climate surrounding the decision, there are substantially three phases in the decision process:

(1) Initial disorientation
(2) Re-orientation process
(3) Solution

As a consequence, it may happen that the decision-maker, even after weighting the different criteria with precision, is supplied by a method with results which do not satisfy her/him because the solution obtained is no longer in accordance with her/his scale of real preferences. Many scientists maintain that the decision-maker is not satisfied because she/he is inconsistent in her/his preferences, thereby forgetting that an inconsistent decision-maker represents the reality to which the mathematical model must be adapted; therefore, unless we admit explicitly that the model is a purely formal one, the irrationality is in the model, not in the decision-maker!

(2) Another important point concerns the interpretation of the *meaning* of the weights supplied by the decision-maker. As we have seen in the previous section, weights are used with the meaning of "trade-off" or "coefficients of importance". In the case of non-compensatory methods, the inter-criteria information required is a relation of relative importance between coalitions of

criteria. Such a concept of relative importance is often translated into numbers called weights[5]. In the case of compensatory methods (e.g. weighted sum) the weights have to be considered as scaling factors and then their meaning is that of a trade-off ratio.

The above considerations imply that there should be consistency between a given aggregation procedure and the questions asked to the decision-maker in order to elicit a set of weights. Otherwise one runs the risk of combining weighting techniques with aggregation models which are not theoretically compatible (examples of these inconsistencies have been shown in Boxes 4.1 and 4.2).

Even if weights are elicited in a well-defined and consistent procedure, it may happen that such weights are not precisely determined (this situation is very common when various actors are involved in the decision process). In such cases two solutions are normally attempted:

- Sensitivity analysis aiming at verifying (by means of different vectors of weights) the robustness and stability of the results obtained with the initial vector of weights. But as it has been noted "one has to recognize that this procedure does not directly and specifically deal with imprecise weights, being only, a way of bypassing the problem" (Bana e Costa, 1990).
- Procedures aiming to deal directly with situations of poor weighting information (e.g., Outweigh Analysis, Bana e Costa, 1990; Regime Analysis, Nijkamp et al., 1990). A common problem of this kind of procedure is that many assumptions need to be made for their correct axiomatization.

(3) In cases of *social decisions*, it is often an impossible task to establish a weighting of the different criteria which satisfies all the decision-makers. For this reason, after the experiment regarding the building of the Paris underground network, in ELECTRE IV, Roy decided to eliminate the weighting of criteria. This practical solution was also adopted by me in the NAIADE method, but probably it is not the right solution. Let us look at this issue in more depth.

To begin we have to think about the relationship between criteria and social actors. In general in the formulation of a set of evaluation criteria, two main tendencies can be distinguished in MCDA literature. On one hand, one may wish to build a decision model as close as possible to the real-world problem; this may increase the number of evaluation criteria to a level such that its application becomes almost impossible. On the other hand, one may wish to use a small number of criteria so that the model is simpler and faster to use; this may lead to an oversimplification.

According to Bouyssou (1990), a family of criteria should have two important qualities:

---

[5] In ELECTRE methods, the concept of importance of a criterion is taken into account by means of
- Its weight (importance coefficient)
- Its veto threshold

(1) "*Legibility*", i.e. the family should contain a sufficiently small number of criteria so as to be a discussion basis allowing the analyst to assess inter-criteria information necessary for the implementation of an aggregation procedure.
(2) "*Operationality*", i.e. the family should be considered by all actors as a sound basis for the continuation of the decision-aid study.

In addition, a family of criteria must also satisfy a number of technical properties, leading to the concept of a *consistent family of criteria* (Roy, 1985). However, all these advices one can find are quite "*technocratic*" in nature, i.e. they assume that finally the choice of criteria is ultimately a business of the *analyst*. On the contrary, I think that the analyst in an SMCE framework, must be a social facilitator and a "*translator*" of social concerns into technical criteria, but clearly the source of these criteria is society.

In Spain, for example, during the Franco dictatorship, there was an important policy criterion: *safety of the northern frontier bordering France*. Nowadays nobody even remembers the existence of this (Franco's) attitude towards frontiers. This highlights the fact that policy criteria are contingent on the social and political framework of a given historical period. To give another example, today the environmental dimension is becoming increasingly important in evaluation projects, although it was almost irrelevant until the seventies.

These reflections can be synthesized graphically in Fig. 4.6. As we know there are different legitimate *values and points of view* in society. This creates social pressure to take various *policy dimensions*, e.g. economic, social and environmental, into

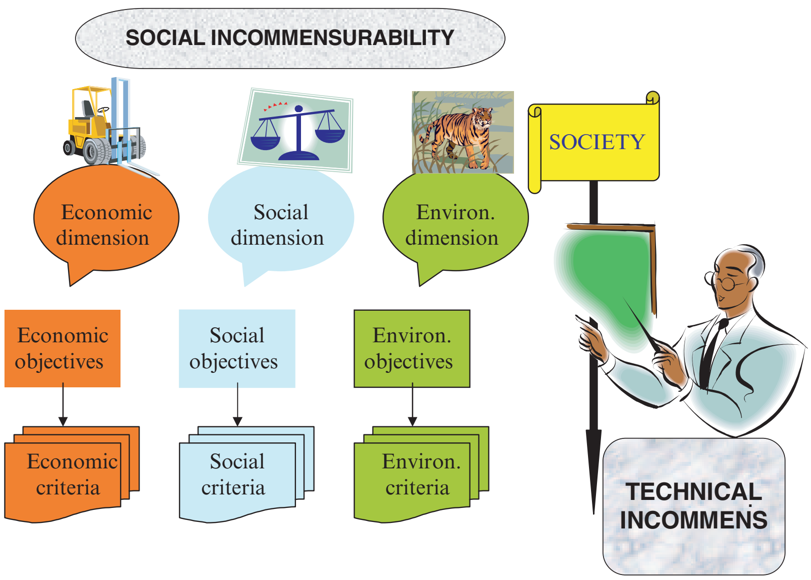


Fig. 4.6 A schematized vision of the hierarchical structure of a policy problem

account. These dimensions are then *translated by analysts* into *objectives and criteria*[6]. This gives rise to a question: who should attach criterion weights and how?

Let us start with the pragmatic solution of *no criterion weighting* (solution adopted in the ELECTRE IV and NAIADE methods). This approach normally reduces conflicts in the problem structuring step, but the question here is: is it *normatively* correct? Indeed the fact that all the criteria have the same weight does not guarantee at all that objectives, dimensions at above all social groups have the same weight.

This would be guaranteed only under the condition that all the dimensions have the same number of criteria. This of course is quite unnatural and artificial and even dangerous. Analysts could be tempted to choose the same number of criteria for each dimension even if these criteria were completely *redundant*. A better solution would be to give the same weight to each dimension and to split each weight among the objectives and criteria of any dimension proportionally.

We arrive then at the conclusion that by giving the same weight to all the criteria the different social dimensions have different weights (since any dimension will be weighted according to its number of criteria). On the contrary *different criterion weights* can guarantee that all the *dimensions are considered equal*!

To demonstrate this point, let us look at the sustainability assessment exercise presented in Chap. 1. The indicators presented in Table 1.2 indeed belong to three dimensions, i.e. economic, social and environmental, considered essential in any sustainability assessment. Let us then try to understand to which dimension each single indicator belongs. The following rough classification may be made:

Economic dimension

1. City product per person

Environmental dimension

2. Use of private car
3. Solid waste generated per capita

Social dimension

4. Houses owned
5. Residential density
6. Mean travel time to work
7. Income disparity
8. Households below poverty line
9. Crime rate

Clearly the social dimension is implicitly receiving a much greater weight than any other dimension (considering that six indicators over nine belong to this dimension). A reasonable decision might be to consider the three dimensions equally important. This would imply giving the same weight to each dimension and then

---

[6] This hierarchical structure of a multi-criterion problem is similar to some extent to the one proposed by Saaty (1980). However, I use the basic idea of hierarchy but no technical proposals of AHP.

splitting this weight among the indicators. That is, each dimension would have a weight of 0.333; then the economic indicator would have a weight of 0.333, the two environmental indicators a weight of 0.1666 each, and each of the six social indicators a weight equal to 0.0555.

As one can see, if dimensions are considered, weighting indicators by means of importance coefficients is crucial. From a technical point of view, the assumption of equal weighting of dimensions can be justified from both theoretical and empirical perspectives. Theoretically since *a priori* there is no reason why one of the typical sustainability dimensions (economic, social and environmental) should receive a higher priority than the others. Pragmatically, because in a multi-inter-disciplinary framework, the equal weighting assumption is equivalent to an equal weighting of different disciplines and thus internal team conflicts can sometimes be reduced.

But at this point one should consider another question: why, from a social point of view, dimensions should be weighted equally. We have to remember in fact that one basic assumption of SMCE is that different dimensions are associated with different groups in society (see Fig. 4.6). Can we find any normative justification for weighting social groups?

4.5.2 Weights and Social Incommensurability

There are at least four ethical principles which are susceptible of being transformed into social weights:

- Political democracy
- Economic democracy
- Sustainability
- Precautionary principle

Let us discuss them briefly.

(1) ***Political democracy*****.** The basic idea is that the majority of the population has the right to make decisions. In this case, the derivation of social weights is relatively easy: each dimension should have a weight equal to the proportion of population supporting the values represented by this dimension. This approach presents *two main problems* (Moulin, 1988). One is pragmatic in nature, i.e. finding the exact percentage of population can be a difficult and expensive process (for instance holding a kind of referendum for every complex decision that affects a community). The other more theoretical, i.e. the minority always loses without any compensation. For these reasons, economists have proposed the concept of economic democracy.

(2) ***Economic democracy*****.** The main idea of *economic democracy* (based on cost–benefit analysis) is that it is possible to avoid any type of referendum simply because it is possible to derive people's preferences by looking at their behaviour on the market. Moreover, in this way it is also possible to find the intensity of preference of the economic agents (while this is not possible in the political democracy since the principle is one person, one vote) by looking at their consumer surplus and to compensate minorities by means of the Kaldor-Hicks principle. *Problems* connected with this approach are at least twofold (see also Sect. 2.3.2 and Box 4.2):

(1) Distributional issues. Willingness to pay very often depends on *ability to pay*, and not on preferences. Income distribution is then the basic principle one should accept to arrive at social decisions.
(2) There are goods and services for which markets do not exist, e.g. environmental goods. For this kind of goods and services one should then invent artificial markets or derive implicit markets to obtain a so-called shadow price. The most general valuation method is contingent valuation.

(3) ***Sustainability*****.** probably implies that the economic, social and environmental dimensions should be taken into consideration *equally*. From a societal point of view, this implies that an environmental pressure group should receive the same weight as an economic pressure group even if, for example, only 5% of society has some environmental concerns. One way to explain this apparent paradox is that one should accept that this 5% of the population is also speaking for animals and future generations who are not taken into account by either political or economic democracy principles. Another way to defend this position is that humans should behave according to a *Kantian principle of universality* such that any negative impact of their actions on the planet should be avoided. Thus even if no social group is willing to support environmental friendly policies, there is an ethical imperative to take them into account.
(4) ***The precautionary principle*****.** This principle maintains that in all situations where there is a simultaneous presence of uncertainty and irreversibility, prudence is the correct driving principle of social actions. For example, in the case of global warming, due to the fact that there is no clear scientific evidence of what may happen with the present trend of greenhouse gas emissions, it is better to be prudent and to reduce those emissions now, even if the economic cost might be very high. A consequence of the adoption of this principle is that environmental pressure groups might even receive a higher weight than other social actors.

This discussion of weights is a clear example of the difference between a participatory multi-criteria study and a social multi-criteria one. The following conclusions on the use of weights can be drawn:

1. *A plurality of ethical principles* seems the only consistent basis for deriving weights in an SMCE framework.
2. In social decision processes, weights cannot be derived as inputs from participatory techniques. This is *technically* very difficult (e.g. which elicitation method should be used? Which statistical index is a good synthesis of the results obtained? Do average values of weights have any meaning at all?), *pragmatically* undesirable (since deep conflict among the various social actors is likely) and even *ethically* unacceptable (if one accepts that future generations and non-humans must also be taken into account).

3. Weights in the framework of SMCE are clearly meaningful only as *importance coefficients* and not as trade-offs (since different ethical positions can only be translated into weights as importance coefficients). Veto power for minorities is a very important characteristic.
4. *Sensitivity and robustness analysis* have a completely different meaning with respect to the case of single-person and technical decisions.[7] In fact in the case of social decisions weights derive only from a few clear cut ethical positions. This means that sensitivity or robustness analyses have to check the consequences for the final ranking of only these positions *and not of all the possible combinations of weights*. Sensitivity and robustness analysis are thus a way to improve transparency[8] (for more technical details on sensitivity analysis see Saltelli et al., 2004).

4.6 Conclusion

Summarizing the discussion developed in this chapter, we could conclude that the following properties are desirable in SMCE for sustainability public choice.

- To have a ranking of all the alternatives is more useful than to select one alternative only; this implies that dominated alternatives cannot be excluded a priori.
- Indifference and preference thresholds should be explicitly taken into account.
- Mixed information of the widest range should be addressed in a consistent way.
- Complete compensability should be avoided to allow for a strong sustainability approach.
- Weights in this framework are clearly meaningful only as importance coefficients and not as trade-offs.
- Sensitivity and robustness analyses have to check the consequences for the final ranking of only some clear ethical positions and not of all possible combinations of weights.

---

[7] I have discussed this point with Serafin Corral-Quintana. I think that he correctly maintains that in a policy framework, sensitivity analysis should consider the willingness of social actors to implement a given course of action more than combinations of weights (Corral-Quintana, 2001). I disagree on the fact that actors should be considered a function of the power they have to support or fight a policy action. I think this has a descriptive content but not a normative one. This is why I insist on the ethical dimension of a normative exercise such as SMCE.

[8] On this point I disagree with Kleijnen (2001), who claims that "modellers should try to develop robust models", in the sense that models should not be very sensitive to modellers' assumptions. Some ethical positions might be very disparate and thus lead to different rankings of the policy options. What is essential in a social framework is then the transparency in the assumptions.

Chapter 5
The Issue of Consistency: Basic Discrete Multi-Criteria "Methods"

5.1 The Total Comparability Axiom: Multi-Attribute Value Functions

The clearest presentation of the theoretical framework of Multi-Attribute Value and Utility Theory[1] can be found in Keeney and Raiffa (1976). Keeney (1992) focuses mainly on operational aspects while Beinat (1997) develops a series of real-world case studies in which value functions are used for environmental management problems. The main characteristic of this approach is that all its formal properties are completely known.

From a theoretical point of view, it is a very elegant and attractive solution to the discrete multi-criterion problem. From the operational point of view, it is the most important theory behind *MCDM*, which assumes a concrete decision-maker who always "believes that in a specified decision context there is a particular preference structure that is appropriate for him" (Keeney and Raiffa, 1976, p. 80).

Multi-attribute value theory is based on the following hypothesis: in any decision problem there exists a *real valued function V* (called *value function*) defined on the set *A* of feasible actions, which a well identified decision-maker wishes, consciously or not, to examine (see Box 5.1). This function aggregates the different criteria (in this context generally referred as *attributes*) taken into consideration, so that the problem can be formulated as

$$max\ V(gi(a)) : a \in A \tag{5.1}$$

where $V(gi(a))$ is a value function aggregating the m criteria; the role of the analyst is to determine this function.

The main underlying assumption of the multi-attribute value functions approach is the identification of human rationality with consistency. More analytically, it is assumed that the preference relations of the decision-maker are a complete pre-order, i.e.:

[1] In this context, the terminology "*value function*" is used when preferences are assumed to be under certainty and "*utility functions*" when probability distributions are present. Here I discuss value functions only.

Box 5.1 Ordinal Utility Theory in Economics

Simon (1983) notes that humans have at their disposal neither the facts nor the consistent structure of values nor the reasoning power needed to apply the principles of utility theory. For instance, the assumption that human behaviour is transitive, can give rise to the famous Luce's paradox (see Chapter 4). In microeconomics, where the assumption that an economic agent is always a utility maximizer is a fundamental one, it is generally admitted that this behavioural assumption has a predictive and not a descriptive meaning. Modern theory of consumer behaviour starts from the notion of "consumer preference". A common set of assumptions (or axioms) underlies preference relations ("completeness", "transitivity", "continuity", "monotonicity", "convexity", and "non-satiation"; see any standard microeconomics textbook for explanation). This implies *ordinal utility theory*, in which utility does not require measurable units. The descriptive realism of the consumer's behaviour as a utility maximizer has been under attack since its inception. In one of the most forceful defences of the maximization hypothesis, Friedman (1953) states that "truly important and significant hypotheses will be found to have 'assumptions' that are widely inaccurate, descriptive representations of reality, and in general, the more significant the theory, the more unrealistic the assumptions (in this sense)". In essence, Friedman believes that the validity of a theory (i.e. its non-falsity) depends purely on its consequences, in other words, that only predictability matters and that the empirical non-realism of assumptions is irrelevant to developing good economic theory models. One should note that the principle underlying multi-attribute value functions is exactly the opposite: *it is possible to describe and aid real-world decisions by using ordinal utility theory*.

- All the states are comparable (between two alternatives **a** and **b** only a preference or an indifference relation can exist)
- Transitivity of the preference relation
- Transitivity of the indifference relation

Therefore, in the "*complete transitive comparability axiom*", preferences can be modelled by means of two binary relations ***I*** and ***P*** having the following properties:

- ***I*** (indifference): reflexive, symmetrical and transitive,
- ***P*** (strict preference): irreflexive, asymmetrical and transitive.

One of the most important consequences of using multi-attribute value functions is that *complete compensability* is always assumed. As stated clearly by Keeney and

Raiffa (1976, p. 66) "our problem is one of value tradeoffs". Since there exists a function *V* by which criteria (attributes) $g_1, g_2, \ldots, g_n$, can be aggregated, there must also exist functions w_{ij} (called trade-offs between the *i*-th and the *j*-th criteria) measuring the amount that a decision-maker is willing to accept on the *j*-th criterion to compensate the loss of a unit on the *i*-th criterion (an amount which may vary according to the point considered in the criteria space).

It is often practically difficult to determine such tradeoffs in precise terms, especially when the criterion scores are expressed in units which appear to lack a common measure. Generally, the trade-offs are the basis of discussions between the analyst and the decision-maker towards constructing the function *V*.

More analytically, a trade-off can be described as following (see also Box 4.1): w_{ij} such that

$$(g_1, \ldots, g_i - 1, \ldots, g_j + w_{ij}, \ldots, g_n)\, I\, (g_1, \ldots, g_i, \ldots, g_j, \ldots, g_n). \tag{5.2}$$

When differentiability is possible, it is:

$$w_{ij} = \frac{\partial V_{(g)} / \partial_{g_i}}{\partial V_{(g)} / \partial_{gj}} \tag{5.3}$$

The most simple (and most commonly used) analytical form is the linear aggregation rule. Let us examine carefully the axioms behind this aggregation rule. Given a set of evaluation criteria $G=\{g_m\}$, $m=1, 2, \ldots, M$, and a finite set $A=\{a_n\}$, $n=1, 2, \ldots, N$ of potential alternatives (actions), the problem is to find $max\ V(g_m(a)) : a \in A$, where $V(g_m(a))$ is an additive value function. If we also assume the existence of inter-criteria information in the form of weights $W=\{w_m\}$, $m = 1,2, \ldots, M$, with $\sum_{m=1}^{M} w_m = 1$ derived as trade-offs[2], the maximization problem can be write as shown in (5.4).

$$max\ V(g_m(a)) = max \sum_{m=1}^{M} w_m v_m(g_m(a)) : a \in A, \tag{5.4}$$

[2] The weights in the additive model are scaling constants which allow marginal value functions to take values on the same interval between zero and one. As a consequence, in this framework weights depend on the measurement scale of value functions. As clearly shown by Anderson and Zalinski (1988), when weights depend on the range of attribute scores, such as in the context of a linear aggregation rule, the interpretation of weights as a measurement of the psychological concept of importance is completely inappropriate.

where any single value function v_m is scaled from zero (worst $v_m(g_m(a))$) to one (best $v_m(g_m(a))$). One must note that this scaling process has nothing to do with a normalization exercise. The main difference is the fact that a normalization technique is an objective change in the measurement unit or origin consistent with measurement scale rules (see Box 5.2 for a review of the main normalization techniques), while

Box 5.2 Main Normalization Techniques

Several techniques can be used to normalize criterion scores (or variables in the context of composite indicators) (OECD, 2003; Saisana and Tarantola, 2002):

Standard deviation from the mean, which imposes a standard normal distribution (i.e. a mean of 0 and a standard deviation of 1). Thus, positive (negative) values for a given alternative or country indicate above (below)-average performance:

$$\left(\frac{\text{actual value - mean value}}{\text{standard deviation}}\right)$$

Distance from the group leader, which assigns 100 to the leading alternative while other alternatives are ranked as percentage points from the leader:

$$100\left(\frac{\text{actual value}}{\text{maximum value}}\right)$$

Distance from the mean, where the (weighted or un-weighted) mean value is given 100, and alternatives receive scores depending on their distance from the mean. Values higher than 100 indicate above-average performance:

$$100\left(\frac{\text{actual value}}{\text{mean value}}\right)$$

Distance from the best and worst performers, where positioning is in relation to the global maximum and minimum and the index takes values between 0 (laggard) and 100 (leader):

$$100\left(\frac{\text{actual value - minimum value}}{\text{maximum value - minimum value}}\right)$$

Categorical scale, where each criterion is assigned a score (either numerical, such as between [1…k], k>1, or qualitative-high, medium, low) depending on whether its value is above or below a given threshold.

"value functions are mathematical representations of human judgements. They offer an analytical description of the value system of the individuals involved in the decision and aim at capturing the parts of human judgements involved in the evaluation of alternatives." (Beinat, 1997, p. 8).

Each criterion is linked to a specific value function which states the attractiveness of each alternative in comparison with all others (by means of attractiveness scores measured on an *interval scale*, independently of the fact that the original criteria may have been either qualitative or quantitative in nature).

The construction of value functions, as well as trade-offs may require a rather lengthy process of interaction between decision scientists and concrete decision-makers (or experts). For example, value functions may increase with the value of the criterion score (e.g. salary levels in evaluating job offers or employment creation in publicly provided goods), decrease (e.g. environmental impact in an infrastructure location problem or a price factor), or peak (e.g. in the evaluation of the number of parking places in a city centre, since a low number of parking places has adverse effects on commercial activities, while a high number increases the risk of congestion and decreases the quality of the urban area (Beinat, 1997, p. 9)). See Fig. 5.1 for graphical examples.

An important point to consider is that each criterion can have its own value function because of the existence of preference independence. This property is a necessary condition for the existence of a linear aggregation rule.

The following theorem holds:

Theorem 5.1 (Debreu, 1960; Krantz et al., 1971; Keeney and Raiffa, 1976)
Given the family of criteria $g_1, g_2, \ldots, g_m$, *an additive aggregation function exists if and only if these criteria are mutually preferentially independent.*

Given the set of evaluation criteria G, a subset of criteria Y is *preferentially independent* of $Y^c=Q$ (the complement of Y) only if any conditional preference among elements of Y, holding all elements of Q fixed, remain the same, regardless of the levels at which Q are held. The *criteria* $g_1, g_2, \ldots, g_m$ are *mutually preferentially independent* if every subset Y of these criteria is preferentially independent of its complementary set of criteria.

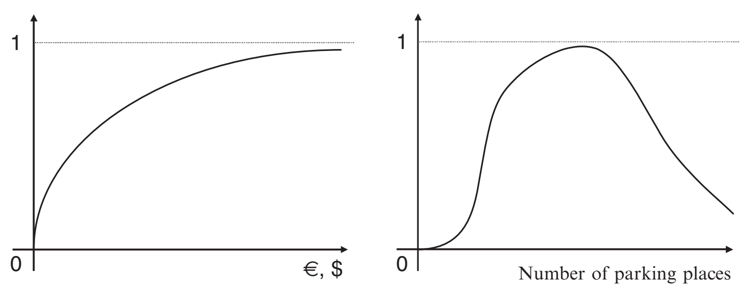


Fig. 5.1 Examples of possible shapes of value functions

Preferential independence is a very strong condition from both the epistemological and operational points of view. It implies that the trade-off ratio between two criteria $w_{x,y}$ is independent of the values of the *M-2* other criteria, i.e.:

$$\frac{\partial w_{x,y}}{\partial q} = 0 \ \forall x, y \in Y, \forall q \in Q \text{(Ting, 1971)} \tag{5.5}$$

From an *operational point of view* this means that an additive aggregation function permits the assessment of the marginal contribution of each criterion separately (as a consequence of the preferential independence condition). The marginal contribution of each criterion can then be added together to yield a total value. If, for example, environmental dimensions are involved, the use of a linear aggregation procedure implies that among the different aspects of an ecosystem there are no phenomena of synergy or conflict. This appears to be quite an unrealistic assumption (Funtowicz et al., 1990). For example, "laboratory experiments made clear that the combined impact of the acidifying substances SO_2, NO_X, NH_3 and O_3 on plant growth is substantially more severe than the (linear) addition of the impacts of each of these substances alone would be." (Dietz and van der Straaten, 1992).[3]
Let us consider the following example:

A and *B* are two different social scenarios, where
M represents an amount of monetary income
S represents a measure of social degradation
E represents an indicator of environmental degradation

Formally, the preference independence condition in this case is:

$$\begin{aligned} &\forall [M(A)0, M(B)0], \\ &[M(A)1, S(A)0, E(A)0]\ \mathbf{P}\ [M(B)1, SB)0, E(B)0] \Rightarrow \\ &[M(A)1, S(A)1, E(A)1]\ \mathbf{P}\ [M(B)1, S(B)1, E(B)1] \\ &S(A)_i = S(B)_i \ \text{ and } \ E(A)_i = E(B)_i \ \ \forall\, i \end{aligned} \tag{5.6}$$

One should note in this context the need to accept a strong implicit assumption, i.e. that *society is indifferent to complete different cases of risk* (described by *S* and

[3] From an *epistemological point of view*, preferential independence implies the separability of values. This is quite an important issue in philosophy; for example, in ethics the thesis of the unity of the virtues is defended to different degrees by Plato, Aristotle and Aquinas. "…it is true that a virtue often cannot be treated apart from the company it keeps. Courage is not an excellence when it appears amid vices-for example as a disposition of the dedicated Nazi.… For example, one's appreciation of the value of freedom, "positive" or "negative", in a particular society, cannot be simply treated as separable from what individuals realise with that freedom. Different political and ethical values when applied in a particular context cannot be applied in isolation from one another" (O'Neill, 1993, p. 114).

E in (5.6)). The possibility that some levels of environmental and social degradation are unacceptable cannot be admitted.

To summarize, we can conclude that the assumption of preference independence is essential for the existence of a linear aggregation rule (and single value functions). Unfortunately, it is usually never tested whether preference independence applies in a given policy problem,[4] although this assumption has very significant and often undesirable consequences.

It has to be noted that any single value function v_m is measured on an *interval scale* while the overall value function V has an *ordinal meaning* only. This implies that indifference and preference thresholds cannot be used in the framework of multi-attribute value functions.

The additive model is completely compensatory. As a consequence weights are never importance coefficients, *they are always trade-offs*. This step is quite difficult on practical grounds[5] (since trade-offs are very difficult to obtain and some consistency requirements are always necessary); while on theoretical grounds a weak sustainability philosophy is always implied (given that a good economic value function assessment can always compensate for a bad environmental assessment). In any case, if one wishes to implement an MCDM framework to elicit decision-makers'

[4] In practice the test of mutual preferential independence is often neglected because it is considered extremely time-consuming. For M criteria, there are $M(M-1)/2$ pairs that must be independent, thus the number of conditions to verify becomes astronomically large as M increases even modestly. However, some findings of Leontieff (1947a, b) can save much of the potential work (see Keeney and Raiffa, 1976, p. 112–114 for more detail).

[5] There is unanimity in the literature that the only method in which weights are computed as scaling constants and in which there is no ambiguous interpretation is the so-called *trade-off method*, starting with revealed preferences. No weight importance judgement is required in this method. The trade-off method can be briefly described as follows. Let us continue with the example of composite indicators seen in Box 4.1, and consider two countries A and B, differing only in the scores of variables x_k and x_t. The problem is then to adjust the score of e.g. x_k for B, in such a way that A and B become indifferent. Formally, it is:

$$I(A) = I(B) \Leftrightarrow I(x_1,...,x_k^{'},...,x_t^{'},...,x_n) = I(x_1,...,x_k^{''},...,x_t^{''},...,x_n)$$

$$\Rightarrow \sum_{\substack{i=1 \\ i\neq k,t}}^{N} w_i x_i + w_k x_k^{'} + w_t x_t^{'} = \sum_{\substack{i=1 \\ i\neq k,t}}^{N} w_i x_i + w_k x_k^{''} + w_t x_t^{''}$$

$$\Rightarrow w_k x_k^{'} + w_t x_t^{'} = w_k x_k^{''} + w_t x_t^{''}$$

This last equation is an equation in the unknown w_k and w_t. To compute the N weights as trade-offs, it is necessary to assess $N–1$ equivalence relations, which together with the usual normalization constraint $w_i + \ldots + w_n = 1$ determine a linear system of N equations in the N unknown weights. Of course if some uncertainty in the variable scores exists, this method cannot be applied. As one can easily understand to assess weights as trade-offs, as should always be done when using a linear aggregation rule, is a much harder job than to use weights as importance coefficients.

preferences, this approach is no doubt the most appropriate one. Efforts to implement a more user friendly additive value model have been made in the so-called MACBETH (**M**easuring **A**ttractiveness by a **C**ategorical **B**ased **E**valuation **T**ec**H**nique) method (Bana e Costa et al., 2005), where only qualitative judgements about differences of value are required.

From a mathematical point of view, an advantage of the value function approach is that the axiomatic system is known and complete and that the independence of irrelevant alternative axiom is always respected (since alternatives are evaluated one-by-one, no preference reversal phenomenon can occur).

5.2 The Partial Comparability Axiom: Outranking Methods

According to the "*fundamental partial comparability axiom*" (Roy, 1985, 1996), preferences can be modelled by means of four binary relations ***I*** (indifference), ***P*** (strict preference), ***Q*** (large preference), and ***R*** (incomparability).

By means of the large preference relation, all the other relations can be obtained:

- *a**P**b* ⇔ *a**Q**b and not b**Q**a*
- *a**I**b* ⇔ *a**Q**b and b**Q**a*
- *a**R**b* ⇔ *not a**Q**b and not b**Q**a*

In order to avoid giving a discriminating role to differences that are barely significant, indifference and preference threshold are introduced (see Chap. 4).

The concept of partial comparability is the basis of the so-called "*outranking methods*". An action ***a*** outranks an action ***b*** only if ***a*** is at least as "good" as ***b*** according to all the criteria considered. Outranking methods can make some incomparable actions comparable because realistic information exists; other actions remain, nevertheless, incomparable.

Briefly, these models entail aggregating the criteria into a partial binary relation *a**S**b* (an outranking relation) based on the concordance and discordance indexes, and then "exploiting" this relation; each of these two steps may be treated in a number of ways according to the problem formulation and the particular case considered.

To illustrate let us think of a Parliament: the concordant coalition can be considered as the sum of the votes of the members in favour of a given option; according to the majority rule of democracies, this option will be approved if it obtains at least more than 50% of the votes. According to the normative tradition in political philosophy, all coalitions, however small, should be given some fraction of the decision power. One measure of this power is the ability to veto certain subsets of outcomes. This explains the use of the condition of non-discordance.

One should also note that in the concordant coalition the "*number*" of criteria are counted, i.e. only ordinal information is used. In the computation of the discordance index, the "*intensity*" of such discordance is considered. One has to note that in this way all the information contained in the impact matrix is used (both the number of criteria in favour of a given alternative and the intensity of preference of any single criterion), but no compensability is implied and weights can be used as importance coefficients.

This is probably the most ingenious idea in the ELECTRE (**EL**imination **E**t **C**hoix **T**raduisant la **RÉ**alité) methods. In fact in the concordance index, since no intensity is considered weights should not be trade-offs. On the other hand in the discordance index, the intensity is used, but for limiting compensability than to allow it. Of course, to use all the criteria in terms of number and intensity, it is necessary to compare alternative ***a*** with ***b*** and also ***b*** with ***a*** (the criteria in the concordance coalition of ***a*** with ***b*** are in the discordance coalition of ***b*** with ***a***, thus any single criterion is used both in number and intensity) thus greatly increasing the number of pair-wise comparisons.

As an example, we shall apply ELECTRE 1 to a family of criteria without discrimination thresholds. In ELECTRE 1, the proposition *aSb* is accepted if the concordant coalition $C(aSb) = \{g_i \in G : g_i(a) \geq g_i(b)\}$ is sufficiently important (condition of concordance) and if for the other criteria the difference $g_i(b) - g_i(a)$ is not too large (condition of non-discordance). The importance of a coalition is represented by the sum of the weights (w_i) of the criteria belonging to that coalition. Thus the index $c(a, b)$ defined by:

$$c(a,b) = \frac{\sum_{i \in C(aSb)} w_i}{\sum_{i=1}^{M} w_i} \tag{5.7}$$

represents the relative importance of $C(aSb)$ among the set of all criteria. Whether or not $C(aSb)$ is sufficiently important is then judged by comparing $c(a,b)$ to a threshold $s > 1/2$ called the concordance threshold.

In order to determine which differences in the discordant criteria are judged too large, a veto threshold v_i (which may vary with g_i) is defined on each criterion in such a way that the existence of a discordant criterion such as $g_i(b) - g_i(a) \geq \Diamond v_i$ prohibits acceptance of *aSb* whatever the value $c(a, b)$. Thus in ELECTRE 1:

$$aSb \Leftrightarrow \left[c(a,b) \geq s \ and \ g_i(b) - g_i(a) < v_i \quad \forall i \notin C(aSb)\right] \tag{5.8}$$

The calculations of concordance and discordance indexes as well as ranking procedures vary according to the ELECTRE method considered. The greatest level of sophistication perhaps, being reached in ELECTRE III (Roy, 1978).

The ELECTRE III method starts by comparing each action ***a*** with each other action ***b***, with the objective of assessing the credibility of the assertion: *action **a** outranks action **b***.

The calculation of this credibility is constructed in a way that it:

- characterizes a group of criteria considered to be in concordance with the affirmation being studied and assess as the relative importance of this group of criteria compared with the remaining ones;
- characterizes among the criteria not in concordance with the affirmation studied, those whose opposition is strong enough to reduce the credibility which would result from taking into account just the concordance, and calculates the possible reduction in credibility that would thereby result.

Given any pair of actions, a "*fuzzy outranking relation*" is defined. This relation summarizes the results of the comparisons of all possible pairs of actions. Values of the credibility of the outranking close to 0 or 1 represent a solidly established relation, whereas values close to 0.5 indicate a weak affirmation.

The exploitation of the fuzzy outranking relation is achieved by constructing two pre-orders Z_1 and Z_2 using a descending and ascending distillation process (respectively) and by then trying to combine these to produce a partial pre-order $Z = Z_1 \cap Z_2$.

A common property of all outranking methods is that they are partially non-compensatory; thus a weak sustainability philosophy is generally avoided and weights can be used as "coefficients of importance". However, this property, very important in our context, is disputed by some. It is argued that weights also depend on intensity, since the complete formalization of importance depends on both the weight and the veto threshold corresponding to a given criterion. This has two main implications: (1) when changing the weight of a criterion, its veto threshold must also change, (2) the fact that weight and veto threshold are connected means that criterion weight also depends on intensity, how can we then be sure that it is not a trade-off? (Suggestions for deriving weights in the context of ELECTRE methods can be found in Figueira and Roy, 2002).

Issues connected with the outranking methods, as well as with other approaches based on pair-wise comparisons are:

- The axiom of independence of irrelevant alternatives is not respected. Thus the phenomenon of rank reversal may appear.
- The so-called Condorcet Paradox may appear (see Chap. 6), i.e. alternative ***a*** may be ranked better than ***b, b*** better than ***c*** and ***c*** better than ***a***.

A problem, specifically connected with the outranking approach, is that, it is necessary to establish a large number of *ad hoc* parameters, i.e. indifference and preference thresholds, concordance threshold, discordance thresholds and weights. This may cause a loss of transparency and consistency in the model.

Roy's defence is that these parameters allow the decision-maker to express her/his preferences and to learn more about them, thereby operationalizing the concept of decision aid (MCDA). This argument is valid in the case of single-person decisions when the aim is to create a learning process.

5.3 The PROMETHEE Methods

As in ELECTRE III, PROMETHEE (**P**reference **R**anking **O**rganization **METH**od for **E**nrichment **E**valuations) (Brans et al., 1986) involves in building a valued outranking relation, by using concepts and parameters that have some physical (or economic) interpretation for the decision-maker, an MCDA philosophy is then adopted.

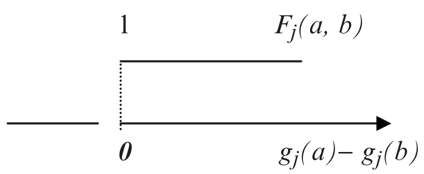
Criterion weights w_j, according to the authors *increasing with the importance of the criterion*, are used. Given *M* criteria and *N* alternatives, the multi-criteria preference index $\Pi(a, b)$ of each ordered pair of actions *(a, b)* is computed by considering the weighted average of the preference functions *Fj(a, b)*. It is:

$$\Pi(a,b)=\frac{\sum_{j=1}^{M} w_j F_j(a,b)}{\sum_{j=1}^{M} w_j} \tag{5.9}$$

Fj(a, b) is a number between 0 and 1 which increases if $g_j(a) - g_j(b)$ is large, and equals zero if $g_j(a) \leq g_j(b)$. According to the way the decision-maker's preferences increase with the difference $g_j(a) - g_j(b)$ for each criterion, the form F_j is chosen from among the following options:

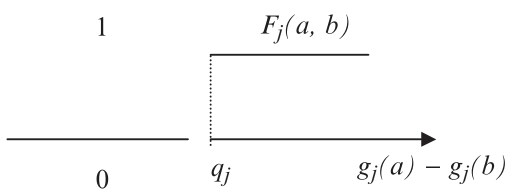
First form

Immediate strict preference; no parameter to be determined.



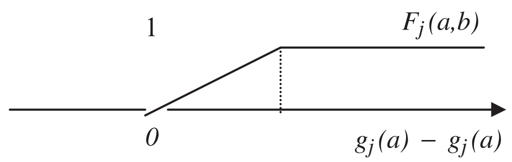
Second form

An indifference threshold exists which must be fixed.



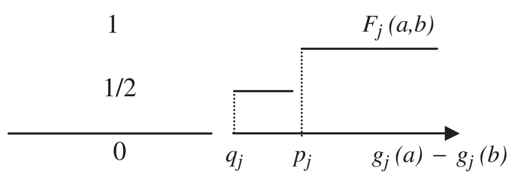
Third form

Preference increases up to a preference threshold, to be determined.



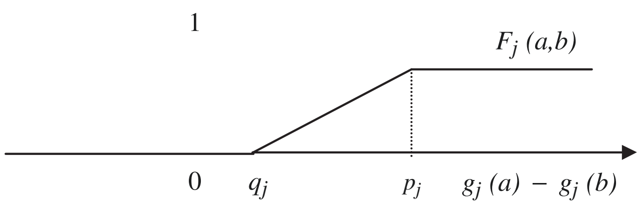
Fourth form

An indifference and a preference threshold exist which must be fixed; between the two, preference is average.



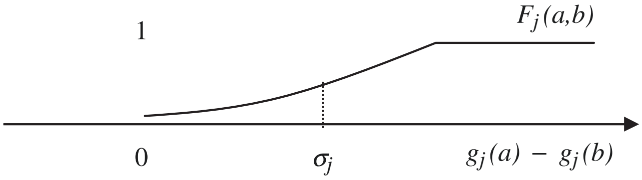
Fifth form

An indifference and a preference threshold exist which must be fixed; between the two, preference increases.



Sixth form

Preference increases following a normal distribution, the standard deviation of which must be fixed.



From a mathematical point of view, forms 1, 2 and 3 are obviously particular cases of form 5; it is nevertheless easier to present them separately to the decision-maker since each one corresponds to a very specific attitude. According to the authors, in general, these six forms are sufficient to cover a large number of possible attitudes, and $F_j(a, b)$ along with the corresponding multi-criteria preference index represent a type of *preference intensity* (Brans et al., 1986, p. 232).

The multi-criteria preference index of the PROMETHEE method is quite similar to the concordance index of ELECTRE III, with the difference that no discordance concept is introduced in PROMETHEE. The multi-criteria preference index $\Pi(a, b)$ determines a valued outranking relation on the set of N alternatives. This relation can be represented as a valued outranking graph, the nodes of which are the N alternatives. Between two nodes (alternatives) ***a*** and ***b***, there are two arcs having values $\Pi(a, b)$ and $\Pi(b, a)$ (Brans et al., 1986, p. 232).

As in ELECTRE III, two complete pre-orders are built; one according to a *leaving flow* index and another one according to an *entering flow* index. Both are computed for each individual alternative. The leaving flow is the sum of the *values* of the arcs leaving node ***a*** and therefore provides a measurement of the strength of ***a***. symmetrically, the entering flow measures the weakness of ***a***. Thus the higher the leaving flow and the lower the entering flow, the better the action. In PROMETHEE 1 the intersection between the two pre-orders is considered; as a consequence the final pre-order may present some incomparability relations between alternatives. In PROMETHEE 2, a complete pre-order is computed (with no incomparability relation) by characterizing each alternative by its *net flow*, i.e. the difference between the leaving and the entering flow.

The PROMETHEE approach is based on a simple mathematical structure; this makes it easy to use. Given that models of preference intensity are used, the degree of compensability allowed is high. Moreover, weights cannot be considered as importance coefficients, they should be derived as trade-offs. The possibility of rank reversals is high.

5.4 The REGIME Methods

Originally this method was designed for dealing with ordinal criterion scores and ordinal weights (Hinloopen et al., 1983). Later it was adapted for dealing with mixed criterion scores (Hinloopen and Nijkamp, 1990).

As we have seen in Sect. 5.2, in ELECTRE methods the concordance index takes into account the number of criteria in favour of a given alternative and NOT its intensity. The concordance index can thus be computed when only ordinal criterion scores are present. This is the main starting point of the ordinal REGIME method. The added sophistication in comparison with the ELECTRE family is that weights are also considered ordinal, thus probability concepts need to be introduced into the algorithm.

Ordinal weights are interpreted as originating from unknown quantitative weights. A set *S* is defined, containing the whole set of quantitative weights that conform to the qualitative priority information. In some cases the sign will be the same for the whole set *S*, and the alternatives can be ranked accordingly. In other cases the sign of the pair-wise comparison cannot be determined unambiguously. In other words, let us define C_{ab} as the sum of weights that are related to the criteria for which ***a*** is better than ***b***. Then the concordance index for the same alternatives is calculated, but this time by considering the criteria for which ***b*** is better than ***a***, i.e. C_{ba}. After having calculated these two sums, the net concordance index $\mu_{ab} = C_{ab} - C_{ba}$ is calculated. If the sign is positive, this indicates that alternative ***a*** is more attractive than alternative ***b***; otherwise the opposite holds. Due to the ordinal nature of the information on weights, in the index μ_{ab} no information exists regarding the size of the difference between the alternatives; it is only the sign of the indicator that matters. However there can sometimes be ambiguity in the sign of the index μ_{ab}. These cases are defined as *critical regimes*. Let us examine the numerical example shown in Table 5.1 with three alternatives and four criteria (Hinloopen et al., 1983, pp. 85–93).

In this impact matrix, rank number 3 indicates the best outcome while rank 1 is assigned to the worst outcome per criterion. If one compares alternative ***a*** with alternative ***b***, the following regime is obtained:

[−1 +1 +1 −1], where −1 means that ***b*** is better than ***a*** and +1 that ***a*** is better than ***b***. Let us assume the following ordinal ranking of criterion importance:

$$w_2 \geq w_4 \geq w_3 \geq w_1 \geq 0, \ \ where \sum_{i=1}^{4} w_i = 1 \tag{5.10}$$

According to (5.10) it is $\mu_{ab} = -w_1+w_2+w_3-w_4$ which is non-negative in all cases (since w_2 is greater than both w_1 and w_4 and w_3 is greater than w_1), thus ***a*** is definitely to be preferred to ***b.*** Now instead of (5.10), let us assume the ordinal ranking of criterion importance shown in (5.11).

$$w_1 \geq w_2 \geq w_3 \geq w_4 \geq 0, \ \ where \sum_{i=1}^{4} w_i = 1 \tag{5.11}$$

Table 5.1 Example of an ordinal impact matrix

Alternatives	a	b	c
Criteria			
g_1	1	2	3
g_2	3	2	1
g_3	3	1	2
g_4	2	3	1

By using (5.11), the sign of μ_{ab} can be both positive and negative. Let us in fact assume two vectors of weights both consistent with (5.11):

$$W_1 = \begin{bmatrix} 0.40 & 0.35 & 0.20 & 0.05 \end{bmatrix}$$

$$W_2 \begin{bmatrix} 0.45 & 0.30 & 0.15 & 0.10 \end{bmatrix}$$

It is easy to verify that by using W_1 μ_{ab} is positive whereas by using W_2 μ_{ab} is negative.

This difficulty in REGIME is circumvented by partitioning the set of feasible weights so that for each subset of weights a definite conclusion can be drawn about the sign of the pair-wise comparison. The distribution of the weights within *S* is assumed to be uniform and therefore the relative sizes of the subsets of *S* can be interpreted as the probability that alternative ***a*** is preferred to alternative ***b***. Probabilities are then aggregated to produce an overall rating of the alternatives, based on a success index or success score. A frequency matrix approach is used to isolate the final alternative (see Chap. 6).

The REGIME method presents the advantage of taking into account *ordinal information*. It should be remembered that weights can also be simply ordinal in nature thus being clearly only importance coefficients. From this point of view REGIME can be considered *entirely consistent in its use of weights as importance coefficients*. The only assumption required is that of a uniform distribution that is not that restrictive (this assumption is based upon the Laplace criterion in the case of decision-making under uncertainty).

However, when *mixed information on criterion scores* is present, the aggregation procedure becomes cumbersome. To deal with both qualitative and quantitative information, it is assumed that qualitative information is the representation of unknown quantitative information. Building on this assumption, a cardinalization scheme is applied and then all information is dealt with as quantitative.

Besides considerations on the way the ordinal information is transformed into quantitative, there is here a basic problem with weights. In fact, if weights are connected to quantitative information and the intensity of preference concept is used, weights *can only be trade-offs and not importance coefficients*. Given that when mixed information is present, all the information is considered quantitative, REGIME loses its desirable characteristic of dealing with weights as importance coefficients.

5.5 The Analytic Hierarchy Process

The Analytic hierarchy process (AHP) has been developed by Saaty (1980), it structures the decision problem into levels which correspond to the decision-maker's understanding of the situation: objectives, criteria, sub-criteria, and alternatives.

By breaking the problem up into levels, the decision-maker can focus on smaller sets of decisions. The AHP is based on four main axioms:

1. Given any two alternatives (or sub-criteria), the decision-maker is able to provide a pair-wise comparison of these alternatives for any criterion on a *ratio scale* which is reciprocal.
2. When comparing any two alternatives, the decision-maker never judges one to be infinitely better than another for any criterion.
3. One can formulate the decision problem as a hierarchy.
4. All criteria and alternatives which impact on a decision problem are represented in the hierarchy.

The axioms above describe the two basic tasks in the AHP (Harker, 1989): structuring and solving the problem as a hierarchy and eliciting judgements in the form of pair-wise comparisons. In order to help a decision-maker to assess these pair-wise comparisons, Saaty (1980) created a nine-point intensity scale between two elements. The numbers suggested to express degrees of preference between two elements A and B are shown in the Table 5.2:

Numbers in between (2, 4, 6, 8) can be used to represent compromises between categories.

The elicitation of priorities for a given set of options for a given criterion involves the completion of an $N \times N$ matrix, where N is the number of options under consideration. However, since the comparisons are assumed to be reciprocal, one needs to answer only $N(N-1)/2$ of the comparisons.

Saaty proposed an *eigenvector approach* for the estimation of the weights from a matrix of pair-wise comparisons. The eigenvector approach is a theoretically and practically proven method for estimating the weights. The eigenvector also has an intuitive interpretation since it is an averaging of all possible ways of thinking about a given set of alternatives.

After estimating the weights, the decision-maker is also provided with a measure of the inconsistency of the given pair-wise comparisons. It is important to note that the AHP does not require decision-makers to be consistent, but rather provides a measure of inconsistency as well as a method to reduce this measure if it is deemed to be too high. After generating a set of weights for each alternative for any criterion, the overall priority of the alternatives is computed by means of a linear, additive function.

Table 5.2 Saaty's nine-point scale

If A is..… as (than) B	Then the Preference Number to Assign is
Equally as important as	1
Moderately more important than	3
Much more important than	5
Very much more important than	7
Exceedingly more important than	9

The AHP is a very widespread approach in many real-world applications; moreover *it is the only method that explicitly deals with the issue of hierarchy in decision problems*. However, it is thought a good decision tool when the decision-maker is clearly identifiable and expresses her/his preferences and takes responsibility for the decision taken. This may be the case in entrepreneurial decisions but is hardly ever so in social decisions.

The basic assumptions (1) and (4) are indeed very strong. Assumption (1) seems disputable because although it is possible that there is mixed information in the impact matrix, it is supposed that it is always possible to transform even ordinal information into a ratio scale, which is the most precise scale of measurement! Assumption (4) is subject to all the arguments developed in this book on the limits of descriptive models in general (see Chap. 2) and of the concept of efficiency in particular (see Chap. 4).

One should note that AHP can be considered a particular case of the multi-attribute value functions approach, thus being completely compensatory in nature. This implies that the weights derived in an AHP framework *are always in the form of trade-offs and never of importance coefficients*. Finally, contrary to multi-attribute value functions, AHP may display rank reversals quite often.

5.6 The NAIADE Method

NAIADE (**N**ovel **A**pproach to **I**mprecise **A**ssessment and **D**ecision **E**nvironments) (Munda, 1995) is a discrete multi-criteria method whose impact (or evaluation) matrix may include crisp, stochastic or fuzzy measurements of the performance of an alternative with respect to an evaluation criterion, thus it is very flexible for real-world applications.

A peculiarity of NAIADE is the use of conflict analysis procedures to be integrated with the multi-criteria results. NAIADE can give the following information:

- Ranking of the alternatives according to the set of evaluation criteria (i.e. technical compromise solution/s)
- Indications of the distance of the positions of the various interest groups (i.e. possibilities of convergence of interests or coalition formations)
- Ranking of the alternatives according to actors' impacts or preferences (i.e. social compromise solution/s)

From a mathematical point of view, two main issues are faced:

1. The problem of equivalence of the procedures used in order to standardize the mixed criterion scores
2. The problem of comparison of fuzzy numbers typical of all fuzzy multi-criteria methods

These two issues are dealt with a new semantic distance that is useful in the case of continuous, convex membership functions also allowing a definite integration (see Chap. 7).

The whole NAIADE procedure can be divided into four main steps:

1. Pair wise comparison of alternatives according to each criterion
2. Aggregation of all criteria
3. Ranking of alternatives
4. Social conflict analysis

The comparison between the criterion scores of each pair of actions is carried out by means of the semantic distance. Another important advantage of the preference modelling introduced in NAIADE is for sensitivity analysis. The modelling procedure based on the notion of a pseudo-criterion, may present a serious lack of stability. Such undesirable discontinuities make a sensitivity analysis (or robustness analysis) necessary. However, one should combine variations of two thresholds and k scores of the M criteria; fuzzy sets can help in overcoming this combinatorial problem. In fact by using membership functions, small variations of input data (scores, thresholds) will modify in a continuous way.

A first possible approach is to associate a fuzzy outranking relation with a pseudo-criterion as in ELECTRE III or to use a generalized criterion as in PROMETHEE. It has to be noted that both methods require that some parameters have to be fixed. A second approach could be the use of fuzzy criterion scores; however, in this case the problem of comparison of fuzzy numbers arises. This is why in NAIADE an approach based on both the underlying philosophy of a "fuzzy pseudo-criterion" and the use of fuzzy criterion scores is used.

The value 2, in the example shown in Fig. 5.2 indicates that below this value the difference between the two alternatives is not sufficient to state that one is better than the other. In Fig. 5.3, the value 3.5 indicates that starting from this value, the difference between the options is indeed sufficient to declare that one is better than the other. A problem inherent in the use of precise indifference and preference thresholds is that they can create the strange situation that e.g. in our case up to the

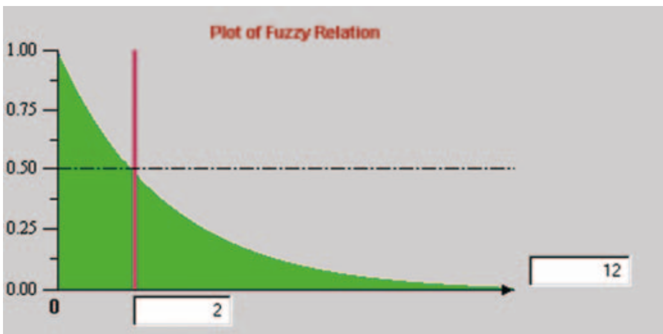


Fig. 5.2 Example of credibility degrees of a fuzzy indifference relation

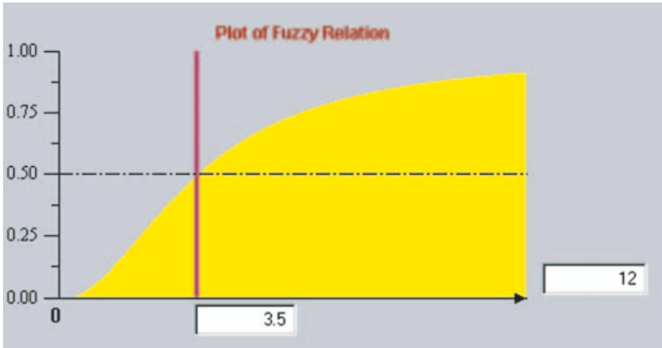


Fig. 5.3 Example of credibility degrees of a fuzzy preference relation

value 1.9999 one would conclude that the two options are indifferent and starting from 2.0001 one would definitely state that the preference relation seems plausible.

For this reason, credibility degrees of the preference and indifference relations are introduced in NAIADE. Credibility degrees are measured on the *y-axis* (while in the *x-axis* the difference intensities for two options and one single criterion are represented), in the case of indifference they indicate that zero difference intensity makes the credibility equal to 1, and then the greater the difference intensity the smaller the credibility of an indifference relation. This credibility is greater than 0.5 up to the value of the indifference thresholds and smaller than 0.5 starting from the indifference thresholds. The credibility of an indifference relation then necessarily must be a monotonically decreasing function like the one shown in Fig. 5.2. In the case of preference the reverse holds. At zero difference the credibility of preference is zero, then the greater the intensity the more credible the preference relation. This credibility is greater than 0.5 when the preference threshold is overtaken. As a consequence, the credibility degree of a preference relation can only be a monotonically increasing function like the one shown in Fig. 5.3.

As one can see thanks to the preference modelling based on the use of credibility degrees, the issue of significance of difference intensities is dealt with properly, and no abrupt transition from indifference to preference is allowed.

Given the information on the pair-wise performance of the alternatives according to each single criterion, it is necessary to aggregate these evaluations in order to take all criteria into account simultaneously. This is done by using a kind of concordance index aggregating the various credibility degrees obtained according to the criteria used. This is done by introducing parameters that *allow us to establish the degree of compensability* desired in the problem at hand. The final ranking of the alternatives in a complete or partial preorder (γ problem formulation), is obtained by means of the basic idea of positive (leaving) and negative (entering) flows of the PROMETHEE methods.

Finally, in the light of an equity analysis, it is possible to evaluate the impact of various actions on different income/interest groups (see Chap. 8). One should note that there might sometimes be a serious divergence between the multi-criteria ranking and the equity ranking. This is mainly because the information provided by these rankings is different in nature (they would otherwise be redundant).

The NAIADE method presents three main advantages:

- The possibility to take into account various forms of mixed information in an equivalent way.
- The possibility to determine the degree of compensability allowed in the aggregation procedure.
- The explicit use of a conflict analysis procedure, so distinguishing clearly the technical and social compromise solutions.

However this method suffers also from some serious limitations:

- The impossibility of using weights explicitly. Moreover if they are used, *they could be only trade-offs* and never importance coefficients.
- The large number of *ad hoc* parameters needed for the elaboration of the multi-criteria impact matrix.
- The fact that qualitative information can be used only in the form of linguistic variables and can never be measured on a purely ordinal scale.
- Rank reversals may easily appear.

5.7 Ideal and Reference Point Approaches

Ackoff (1978) writes: "An ultimately desired outcome is called an "ideal". If one formulates a problem in terms of approaching an ideal solution, one minimizes the changes of overlooking relevant consequences in decision making. Seeking the ideal is the best way to open and stimulate the mind to creative activity".

Briefly, the philosophy underlying the multi-criteria methods based on ideal point concepts can be summarized as follows (Yu, 1973, 1985; Zeleny, 1974, 1982). Multi-criteria problems are characterized by conflict because of perceived absence of an obvious "best" option; therefore, the only way to resolve the conflict is to find or invent an ideal point. The only way to decrease the intensity of the conflict is to find or generate alternatives which are as close as possible to the ideal point.

Coombs (1958) assumes that there is an ideal level of attributes for objects of choice and that the decision-maker's utility function decrease monotonically on both sides of this ideal point. He shows that probabilities of choice depend on whether compared alternatives lie on the same or opposite sides of the ideal. The ideal point procedures are characterized by the following axiom of choice: *alternatives that are closer to the ideal are preferred to those that are farther away. To be as close as possible to the perceived ideal is the rationale of human choice*.

A concept similar to the ideal alternative, its mirror image, the anti-ideal which can be defined on any properly bounded set of feasible alternatives. The question is, do humans strive to be as close as possible to this ideal or as far away as possible from the anti-ideal? Zeleny (1982) writes: "our answer-both. As a matter of fact we propose that humans are capable of switching between the two regimes according to the given circumstances of the decision process…Naturally, the compromise set based on the ideal is not identical with the compromise set based on the anti-ideal. This fact can be used in further reducing the set of available solutions by considering the intersection of the two compromises". The issue of how losses and gains are normally evaluated by human beings has been studied in depth by Kahneman and Tversky (1979). Applications of their prospect theory inside multi-criteria decision theory have been done by Gomes and Lima (1992).

One of the traditional ideal point approaches is to compute the "distance" of each action from the ideal point and then rank them in terms of their proximity to the ideal. One issue related to this approach is that each action is considered completely independent of the set of all the other actions; the assumption is that humans compare each action with the ideal rather than with each other. Clearly this philosophy is close to that of multi-attribute value functions. Indeed, the distance functions used in this approach can be considered a type of value function. As we already know, an advantage is that no preference reversal is possible.

Aspiration levels (or goals) express the decision-maker's ideas about the desired outcomes of the decision in terms of a certain level to be aimed at for each objective. There is a close link between the concept of aspiration level and the theory of satisficing behavior (Simon, 1983). The usual way in which aspiration levels are treated is by means of goal programming (Spronk, 1981). An advantage of goal programming is that it always provides a solution, even if none of the goals are realizable, provided that the feasible region is non-empty. This is possible by using deviational variables, which show whether the goals have been attained or not. In the latter case, they measure the distance between the realized and aspired levels. An approach that can be viewed as a generalization of goal programming and ideal point techniques is the "*achievement scalarizing functions*" method (Wierzbicki, 1982). The basic idea is constructing a mathematical basis for satisficing decision-making by introducing the wishes of the decision-maker as basic *a priori* information in the form of aspiration levels (reference points). Achievement scalarizing functions can be considered as a modification of traditional value functions, thus issues such as *compensability* and *weights as trade-offs* here apply again. Interesting properties of this approach are its methodological foundation (since satisficing behaviour seems to be quite *descriptive* for human beings) and the fact that the number of parameters needed is very low, just the specification of goals.

In general these approaches are more suitable for describing human behaviour than for social decisions. However an interesting application in a sustainability framework is the possibility of using reference point approaches for benchmarking sustainability indicators (Munda, 2005b).

To give an example, let us consider the urban sustainability example presented in Chap. 1. A reasonable question to ask is: even if we have a very reliable ranking, what is the use of knowing that Moscow is better overall than Amsterdam or vice versa? Let us try to shed some light on this issue. First of all, one should note that for the majority of indicators used in assessment exercises, no clear reference point is available. For instance, when GDP is used nobody knows the ideal value of a country GDP, thus it is quite common to compare with other countries GDP, e.g. the USA one.

In order to obtain a set of reference values, an "ideal point" can be defined by choosing the best values reached for any single indicator. In our example, the vector defining the ideal value (called "ideal city") is presented in Table 5.3.

A first, very simple procedure could be the application of a normalization rule known as "*distance from the group leader*", which assigns 100 to the leading country and ranks other countries as percentage points away from this leader country (see Box 5.2). By applying this normalization rule, which can be considered a simple ideal point method, to the indicator scores of Amsterdam and New York (taking care that when the objective is minimization the leader is the city with the lowest indicator score – the leader is of course the ideal city) the results presented in Table 5.4 are obtained. As shown in Figs. 5.4–5.6, these results can also be presented graphically for making their interpretability easier. The numerical results are synthesized by using the so-called radar diagrams, in which the ideal city reaches the score 100 on any indicator.

According to these computations, Amsterdam reaches the ideal values on mean travel time to work and income disparity therefore these issues can be considered resolved. Households below the poverty line and city product per person are not a great concern. All the issues connected to the other indicators signal on the contrary, a problem in this city. Their policy relevance is high and their resolution should be urgent (above all, the crime rate).

Table 5.3 Multidimensional representation of the "ideal city"

Criteria \ Alternatives	Ideal City
Houses owned (%)	50.5
Residential density (pers./hectare)	72
Use of private car (%)	10
Mean travel time to work (minutes)	22
Solid waste generated per capita (t./year)	0.2
City product per person (US$/year)	30952
Income disparity (Q5/Q!)	5.25
Households below poverty line (%)	15
Crime rate per 1000 (theft)	4.3

Table 5.4 Benchmarking exercise using the distance-from-the-leader method

	Amsterdam	New York
Economic dimension		
City product per person	91.27	100
Environmental dimension		
Use of private car	16.6	30.7
Solid waste generated per capita	50	32.78
Social dimension		
Houses owned	4.35	20.39
Residential density	47.33	100
Mean travel time to work	100	60.27
Income disparity	100	37.47
Households below poverty line	73.17	92.02
Crime rate	2.98	7.58

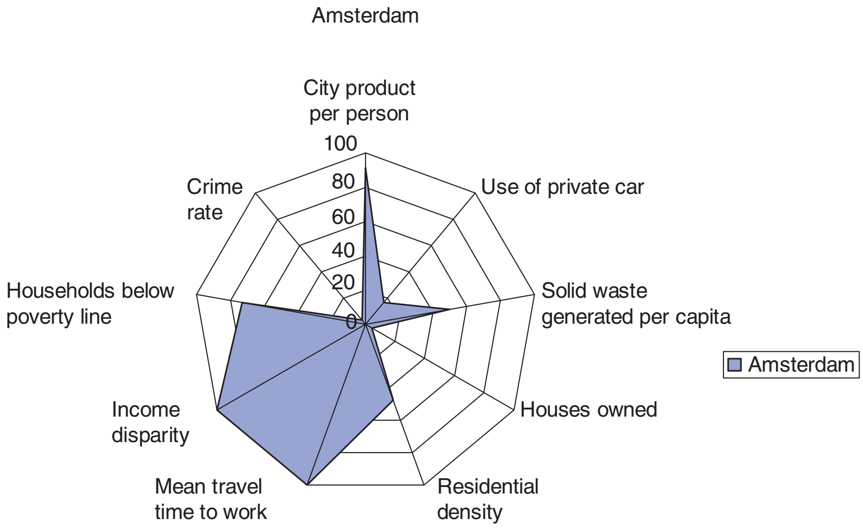


Fig. 5.4 Radar diagram for Amsterdam's sustainability benchmarking

New York is doing well on city product per person and residential density (where it meets perfectly the ideal values), rather well on households below the poverty line and mean travel time to work (where it is not so far from the ideal values) and rather badly on the other values where it is definitely worse than the ideal values used, and as a consequence, in our hypothetical situation, the policy issues connected with these indicators should be considered important priorities. In reality,

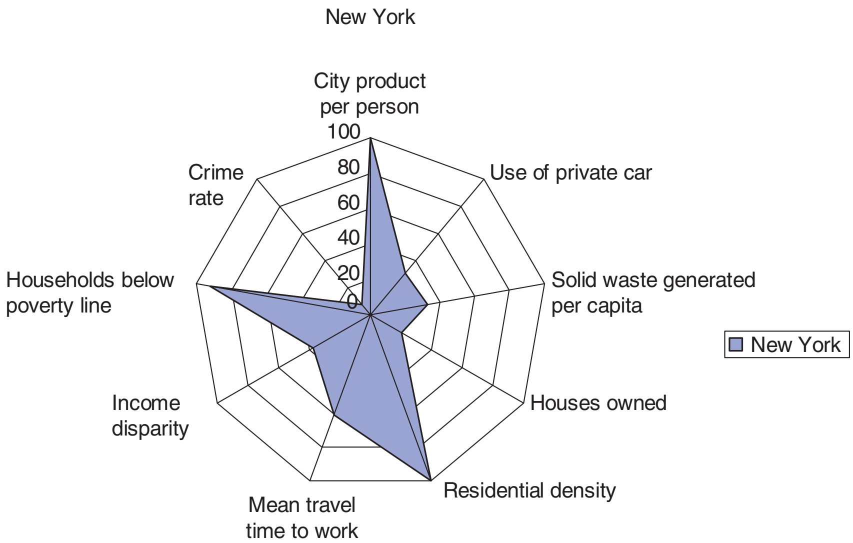


Fig. 5.5 Radar diagram for New York's sustainability benchmarking

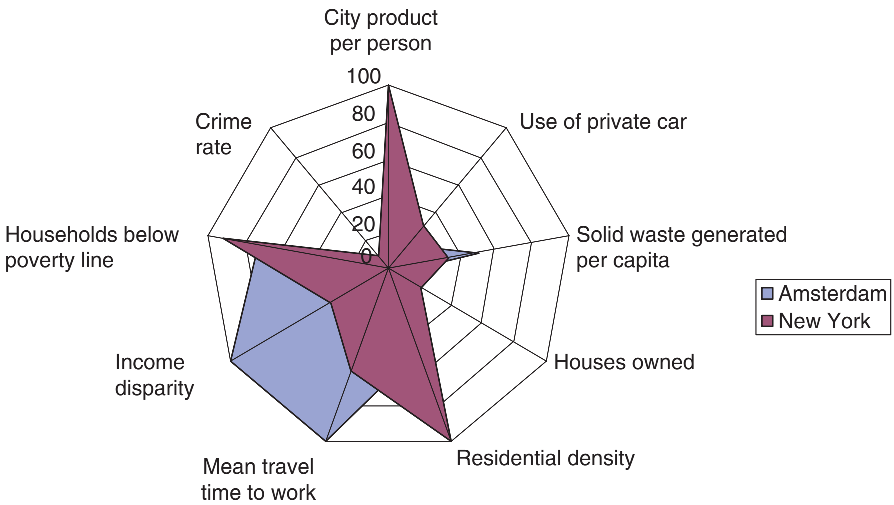


Fig. 5.6 Radar diagram comparing Amsterdam and New York

one should note that these ideal values depend on the cities we are comparing. In this case, the four cities are so heterogeneous that probably their comparison is meaningless. Moreover the issue of information quality applies: who really believes that residential density is not a problem in New York?

5.8 Conclusion

To conclude this chapter, it is possible to state quite safely that a problem of multi-criteria decision theory is that many different mathematical aggregation conventions (or methods) exist. For the case of SMCE, the following considerations may be useful when attempting to select appropriate methods.

The idea of *social incommensurability* makes the following properties desirable in a social multi-criteria method:

- Weights in this framework are clearly meaningful only as importance coefficients and not as trade-offs. As a consequence, complete compensability cannot be implemented.
- Conflict analysis procedures explicitly looking for social compromises should integrate an SMCE exercise.
- In a policy framework, to have a ranking of all the alternatives is more useful than to select one alternative only.

The idea of *technical incommensurability* makes the following properties desirable in a social multi-criteria method:

- Partial or complete non-compensability is an essential consistency requirement.
- Indifference and preference thresholds should be explicitly taken into account.
- Mixed information of the widest range should be addressed in a consistent way.
- Simplicity, i.e. the use of as few parameters as possible, is a very desirable property.
- The hierarchical dimension of a policy problem should be explicitly considered.

In Table 5.5, the methods presented in this chapter are confronted with the desirable properties proposed for SMCE. An evaluation has been made of whether certain properties are present or not and sometimes also of the degree (low, medium, high) (the latter by means of a subjective evaluation). For mixed information, it is held that a method can deal with different criterion scores if no transformation is required (such as value functions or cardinalization schemes).

Table 5.5 Evaluation of some multi-criteria methods according to desirable properties for SMCE

	Comp.	Imp. Coeff.	Mix. Inf.	Simpl.	Hier.	γ	Thresh	Conf. Anal.
Value Funct.	High	No	No	High	No	No	No	No
ELEC.	Low	Not Clear	Partly	Low	No	Yes	Yes	No
REG. (1983)	Low	Yes	No	High	No	No	No	No
REG. (1990)	High	No	Yes	Low	No	No	No	No
NAIA.	Medium	No	Yes	Low	No	Yes	Yes	Yes
AHP	High	No	No	Low	Yes	No	No	No
PROM.	High	No	No	Medium	No	Yes	Yes	No
Ideal. Point	High	No	No	Medium	No	No	No	No

However, even if quite rough, the information contained in Table 5.5 leads to an important conclusion: no one of these methods displays all the properties considered desirable in an SMCE framework. For this reason in Chap. 6 the theoretical foundations of ranking procedures are further studied by using concepts from social choice theory. Chap. 7 presents a new multi-criteria method which aims to offer all the desirable properties for social multi-criteria evaluation in the search for technical compromise solutions. Chap. 8 then deals with the search for social compromise solutions.

Chapter 6
The Issue of Consistency: Lessons Learned from Social Choice Literature

6.1 Social Choice and Multicriterion Decision-Making

Vansnick (1990) showed that the two main approaches in multi-criteria decision theory i.e. compensatory and non-compensatory can be directly derived from the seminal work of Borda (1784) and Condorcet (1785). Indeed from the social choice literature, it can be seen that various ranking procedures used in multi-criteria methods have their origins in social choice. To give a few examples: the weakness-strength approach, typical of outranking and PROMETHEE methods (Brans et al., 1986; Roy, 1996), has a clear derivation from two Condorcet consistent rules, i.e. the Copeland (1951) and Simpson rules (1969); Arrow–Raynaud proposed a sequential procedure (building on Köhler, 1978) which is also based on some principles of the Condorcet rule; the so-called frequency matrix approach (Hinloopen et al., 1983; Matarazzo, 1988) comes directly from Borda algorithm, or the permutation method (Paelinck, 1978), has a strong connection with the so-called Kemeny's method, and so on.

Surprisingly this quite obvious relation between multi-criteria decision theory and social choice was completely ignored in the development of the multi-criteria decision literature we have referred to in the previous chapter. In this chapter I critically review some basic concepts of social choice that are also of fundamental importance in the multi-criteria evaluation framework.[1]

In 1986 Kenneth Arrow and Hervé Raynaud published a very influential book entitled "*Social choice and multicriterion decision-making*", in which the formal analogies between the discrete multi-criterion problem and the social choice problem are analysed in depth. This book was based on the assumption that, in the case that all criteria have ordinal impact scores, if one considers the evaluation criteria as voters, a multi-criteria impact matrix and a voting matrix are identical. As a consequence all results of social choice also fully apply to multi-criteria decision theory (at least when no intensity of preference and no indifference/preference thresholds are involved).

[1] For reasons of clarity I always use the term criterion/criteria instead of voter/voters. All proofs are omitted; these can easily be found in the references given.

Formally, the discrete multi-criterion problem with ordinal criterion scores can be defined by the following three axioms (adapted from Arrow and Raynaud, 1986, pp. 81–82).

Axiom 1: *Diversity*. Each criterion defines a total order on the finite set *A* of alternatives to be ranked.

Axiom 2: *Symmetry*. The only preference information, provided by criteria, is the ordinal pair-wise preferences.

Axiom 3: *Positive Responsiveness*. The degree of preference between two alternatives ***a*** and ***b*** is a strictly increasing function of the number of criteria (or weights) that rank ***a*** before ***b***.

A good place to start is Arrow's impossibility theorem (Arrow, 1963). This theorem shows that if one formally determines the properties which should define the concept of democracy, a very sad conclusion can be proved: the only political system respecting all the properties is a dictatorship! A legitimate question thus arises: does this paradoxical result also apply to the discrete multi-criterion problem? Arrow and Raynaud (1986, pp.17–23) answer this question. Let us assume that to arrive at a total ranking of all alternatives a multi-criterion aggregation convention needs at least to satisfy three axioms.[2]

Axiom 1: *Unrestricted Domain*. The values that can be taken by the criteria are unrestricted and the mathematical aggregation convention must respect unanimity.

Axiom 2: *Independence of irrelevant alternatives*. The ranking of the alternatives in *A* depends only on the alternatives belonging to *A*. "This means that it is of no importance for the decision if you have forgotten in the application of the method some (poorly ranked) alternatives:.… The complete set of alternatives is always very large and only a relatively small subset can be identified. It is thus essential that the result of the method on a small set of alternatives does not vary if forgotten alternatives are taken into consideration" (Arrow and Raynaud, 1986, p. 19).

Axiom 3: *Positive Responsiveness*. That, as explained in the previous page, requires that preference is a monotonically increasing function of the number of criteria.

[2] The Arrow's original impossibility theorem (Arrow, 1963) is slightly different, above all with respect to the independence of the irrelevant alternatives axiom. In the social choice literature formulation, it is called the *axiom of binary independence*, i.e. the social ranking of each pair of alternatives depends only on the preferences of each *voter* on that specific *pair* of alternatives. The ranking of any other alternative is irrelevant for this social ranking. Indeed, in the version proposed by Arrow and Raynaud (1986) the axiom of independence of irrelevant alternatives is closer to the definition given by Chernoff (1954), which is derived from Nash's bargaining theory. For a deeper discussion of the independence of irrelevant alternatives axiom and its various definitions see e.g. Ray (1973) and Bordes and Tideman (1991).

The following theorem then applies: *the only ranking respecting all these axioms must coincide with the ranking supplied by one of the criteria taken into consideration*. In other words, the correct solution of a multi-criterion problem comes from a mono-criterion optimization. A consequence of this theorem is that no perfect multi-criterion aggregation convention can exist. "Reasonable" ranking procedures must then be found. In the framework of social multi-criteria evaluation, this gives rise to two questions: is it possible to find a ranking algorithm consistent with the desirable *properties*[3] of social multi-criteria evaluation? And the reverse, is it possible to ensure that no essential property will be lost?

In social choice, the reaction to Arrow's theorem has been the search for less ambitious voting structures, making it necessary to retain a few basic requirements only. These basic requirements are generally threefold:

1. *Anonymity*: all criteria must be treated equally.
2. *Neutrality*: all alternatives must be treated equally.
3. *Monotonicity*: more support for an alternative cannot jeopardize its success.

One should note that, while anonymity is clearly essential in the case of voters, it is not the case in the multi-criterion problem since criterion weights can normally be introduced. Life may therefore be easier in the multi-criterion case. We will realize the consequences of abandoning anonymity later in this chapter. Let us now examine some ranking procedures directly coming from the social choice tradition. Emphasis will be put on Arrow's result, in the sense that the limitations of these procedures will be made clear.

Let us start with the numerical example shown in Table 6.1, where 21 criteria rank four alternatives (*a, b, c, d*):

We can assume that the objective is to isolate one alternative (α-decision problem formulation). A first possibility is to apply the so-called *plurality rule*, meaning that the alternative which is most often ranked in first place is the winner. Thus, in our case, alternative ***a*** would be chosen since eight criteria put it in first position. However, if one looks carefully at Table 6.1, it can be seen that alternative ***a*** also has the strongest opposition, since 13 criteria put it in last position.

It is interesting to note that this paradox was the starting point for Borda's and Condorcet's research at the end of the 18th century, but the plurality rule is still the most common electoral system in the 21st Century. This is a clear example of what Arrow's impossibility theorem signifies in the real-world exercise of democracy.

---

[3] Often this search for clear properties characterizing an algorithm is indicated as the axiomatic approach. However, one should note that properties or assumptions are NOT axioms. As summarized perfectly by Saari (2006, p. 110) "Many, if not most, results in this area are merely properties that happen to uniquely identify a particular procedure. But unless these properties can be used to construct, or be identified with *all* properties of the procedure (such as in the development of utility functions in the individual choice literature), they are not building blocks and they most surely are *not* axioms: they are properties that just happen to identify but not characterize a procedure. As an example, the two properties (1) Finnish-American heritage (2) a particular DNA structure, uniquely identify me, but they most surely do not characterize me".

Table 6.1 Numerical example with 21 criteria and 4 alternatives

Number of Criteria	3	5	7	6
	a	a	b	c
	b	c	d	b
	c	b	c	d
	d	d	a	a

Source: Moulin, 1988, p. 228

From the plurality rule paradox two main lessons can be learned:

1. Good ranking procedures should respect the entire ranking of alternatives and not the first position only.
2. It is important to consider not only what a majority of criteria prefer, but also what they reject.

The second lesson leads to the use of veto indexes such as for example proposed by Moulin (1981). In the framework of multi-criteria decision theory the first attempts to use veto indexes were with the ELECTRE methods (Roy, 1985, 1996). However, one should note that in the ELECTRE methods veto indexes refer to pair-wise comparisons of alternatives (by looking at the intensity of preference of quantitative criterion scores belonging to the discordance coalition) and not to the whole overall ranking. Other applications of veto concepts will be discussed in Chaps. 7 and 8. Here we will focus more on the first lesson above.

The Borda solution to the plurality rule paradox is the following scoring rule: given *N* alternatives, if an alternative is ranked last, it receives no points; it receives 1 point if it is ranked next to the last. The scoring process continues like this up to $N-1$ points awarded to the alternative ranked first[4]. The Borda winner is the option with the highest total score. Note that a crucial implicit assumption underlies a Borda rule: the arbitrary transition from an ordinal scale of measurement (individual ranking of feasible alternatives) to an interval or ratio scale one (according to the scoring rule adopted).

Let us then apply the Borda rule to the example shown in Table 6.1. As a first step we can present the information contained in Table 6.1 in a *frequency matrix* fashion. This is done in Table 6.2. This Table shows how many criteria put each of the alternatives into each of the four positions in the ranking. The sum of each row or column is always a constant equal to the number of evaluation criteria (21 in our example). By applying the Borda scoring rule, the following results are obtained:

$a = 8 \times 3 = 24$
$b = 5 + 9 \times 2 + 7 \times 3 = 44$
$c = 10 + 5 \times 2 + 6 \times 3 = 38$
$d = 6 + 7 \times 2 = 20$

[4] The plurality rule can indeed be seen as a particular case of a Borda count where a single point is assigned to a voter's top ranked alternative and zero to all others.

Table 6.2 A frequency matrix for the application of the Borda rule

Alternatives	a	b	c	d	
Ranking					Points
1st	8	7	6	0	3
2nd	0	9	5	7	2
3rd	0	5	10	6	1
4th	13	0	0	8	0

Table 6.3 Outranking matrix derived from the condorcet approach

	a	b	c	d
a	0	8	8	8
b	13	0	10	21
c	13	11	0	14
d	13	0	7	0

It can be seen that the preferred option is no longer ***a,*** but rather ***b***. The plurality rule paradox has thus been solved.

Let us now look at the Condorcet approach. The Condorcet rule is based on a pair-wise comparison between all alternatives considered. For each pair, a *concordance index* is computed by counting how many criteria are in favour of each alternative. In this way an outranking matrix whose elements hold the "constant sum property", is built. The pairs whose concordance index is higher than 50% of criteria are selected. Given the transitivity property, a final ranking is determined.

To make the procedure clearer, let us apply it to the example in Table 6.1. The outranking matrix is shown in Table 6.3: in this case, the constant sum of $e_{ij}+e_{ji}=21 \quad \forall i \neq j$.

The majority threshold is composed of 11 criteria. The pairs with a concordance index higher than 11 are the following:

$b\mathbf{P}a=13$, $b\mathbf{P}d=21$, $c\mathbf{P}a=13$, $c\mathbf{P}b=11$, $c\mathbf{P}d=14$, $d\mathbf{P}a=13$. Clearly, alternative ***c*** is the Condorcet winner since it is always preferred to any other alternative. Alternative ***b*** is preferred to both ***a*** and ***d***. Between ***a*** and ***d***, ***d*** is preferred to ***a***. Thus the final ranking is the following: $c \rightarrow b \rightarrow d \rightarrow a$.

As one can see, the derivation of a Condorcet ranking may sometimes be a long and complex computation process. To solve this problem, so-called Condorcet consistent voting rules have been developed. These rules are practical, quick-searching algorithms for the Condorcet winner or for reasonable substitutes for it, when a Condorcet winner does not exist. The most common ones are the Copeland and the Simpson rules (Moulin, 1988, p. 233).

- *The Copeland rule*. Compare alternative ***a*** with every other alternative ***x***. Score +1 if a majority of criteria prefers ***a*** to ***x***, −1 if a majority prefers ***x*** to ***a***, and 0 if it is a tie. Summing up those scores over all ***x***, $x \neq a$, yields the Copeland score of ***a***. The alternative with the highest such score, called a Copeland winner, is selected.

- *The Simpson rule*. Consider alternative ***a***, and for every other alternative ***x***, compare the number *N(a, x)* of criteria preferring ***a*** to ***x***. The Simpson score of alternative ***a*** is the *minimum N(a, x)* over all ***x***, $x \neq a$. The alternative with the highest such score, called a Simpson winner, is selected. The Simpson rule is based on the same philosophy as the weakness/strength concept, with particular emphasis put on weakness. In fact, an alternative ***a,*** in the framework of the Simpson rule, can win only if no other alternative has a huge majority of criteria against ***a***.

Let us apply these two rules to the problem shown in Table 6.1. By using the outranking matrix of Table 6.3, the Copeland and Simpson rules can easily be computed for each alternative. Using the Copeland rule gives $a=-3$, $b=1$, $c=3$, $d=0$, thus the Copeland winner is alternative ***c***, which is also the Condorcet winner. Using the Simpson rule gives $a=\mathbf{8}$, $b=10$, $c=11$, $d=7$, again alternative ***c*** is selected.

Concluding, we can say that both Borda and Condorcet approaches solve the plurality rule paradox. However, the solutions offered are different. Arrow's theorem explains why this happens. At this point, the question arises: *in the framework of SMCE*, can we choose between Borda and Condorcet on some theoretical and/or practical grounds? A first, partial answer to this question will be offered in the next section.

6.2 Borda Versus Condorcet

The first question to be addressed is: do Borda and Condorcet rules normally lead to different solutions? (One might think in fact that the divergence of solutions we found in the previous section was a very special case.) The question has been answered by Fishburn (1973b) and Moulin (1988). Before we look at those results, however it is necessary to generalize the Borda approach such that we can compare any Borda consistent rule with any Condorcet consistent rule.

The Borda approach can be generalized by means of the concept of *scoring voting rules*, meaning that the Borda winner can always be found for any non-decreasing sequence of real numbers. $s_0 \leq s_1 \leq \ldots \leq s_{p-1}$, *with* $s_0 < s_{p-1}$. s_0 points are given to the alternative ranked last and so on; the leader in the ranking receives s_{p-1} points.

Fishburn (1973b) proves the following theorem: *there are profiles where the Condorcet winner exists and it is never selected by any scoring method*. Moulin (1988, p. 249) proves that "a Condorcet winner (loser) cannot be a Borda loser (winner)". In other words, Condorcet consistent rules and scoring voting rules are fundamentally different in nature. Their disagreement in practice is normal. We must therefore examine both approaches carefully and choose the one we consider more appropriate in a social multi-criteria evaluation framework.

Let us consider a numerical example with 60 criteria and three alternatives; the example shown in Table 6.4 is owed to Condorcet himself.

The corresponding frequency matrix is shown in Table 6.5.

By applying Borda's scoring rule, the following results are obtained: $a=58$, $b=69$, $c=53$, thus alternative ***b*** is univocally selected.

Let us now apply the Condorcet rule. The corresponding outranking matrix is shown in Table 6.6.

In this case, the concordance threshold is 31. It is: ***aPb***, ***bPc*** and ***cPa***, thus because of the transitive property a *cycle* exists and no alternative can be selected. Let us then try the application of Condorcet consistent rules. By applying the Copeland rule, it is $a=0$, $b=0$, $c=0$, thus no selection is possible here either. By means of the Simpson rule it is: $a=25$, $b=27$, $c=18$, alternative ***b*** is then selected which is also the Borda winner.

From this example we might conclude that the Borda rule (or any scoring rule) is more effective, since an alternative is always selected, while Condorcet sometimes leads to an irreducible indecision. It seems appropriate then to know more about the properties of the Borda rule.

Let us examine again the outranking matrix presented in Table 6.6. From this matrix we can see that 33 criteria favour alternative ***a***, while only 27 favour alternative ***b***. So a legitimate question is: why does the Borda rule rank ***b*** before ***a***? This is mainly due to the fact that the Borda rule is based on the concept of *intensity* of preference, while the Condorcet rule only uses the *number* of criteria.

Table 6.4 An original condorcet numerical example

Number of Criteria	23	17	2	10	8
	a	b	b	c	c
	b	c	a	a	b
	c	a	c	b	a

(Source: Condorcet, 1785)

Table 6.5 A frequency matrix derived from Table 6.4

Alternatives	***a***	***b***	***c***	
Ranking				Points
1st	23	19	18	2
2nd	12	31	17	1
3rd	25	10	25	0

Table 6.6 Outranking matrix derived from Table 6.4

$$\begin{bmatrix} & a & b & c \\ a & 0 & 33 & 25 \\ b & 27 & 0 & 42 \\ c & 35 & 18 & 0 \end{bmatrix}$$

In the context of the Borda rule, and all scoring methods in general, the intensity of preference is measured by the scores given according to the rank positions. This implies that *compensability* is allowed. Moreover, the rank position of a given alternative depends on the number of alternatives considered. This implies that the mutual preference relation of a given pair of alternatives may change according to the alternatives considered. As a consequence, *preference reversal* phenomena may easily occur and of course, the axiom of independence of irrelevant alternatives is not respected. This problem has been extensively studied by Fishburn (1984).

Let us examine the numerical example presented in Table 6.7. The corresponding frequency matrix is shown in Table 6.8.

By applying Borda's scoring rule, the following results are obtained: $a=13$, $b=12$, $c=11$, $d=6$; thus alternative ***a*** is chosen. Suppose now that alternative ***d*** is removed from the analysis. Since ***d*** was at the bottom of the ranking, nobody should have any reasonable doubt that alternative ***a*** is still the best alternative. Let us check if this assumption is correct. The new frequency matrix is presented in Table 6.9.

By applying Borda's scoring rule, the following results are obtained: $a=6$, $b=7$, $c=8$; thus alternative ***c*** is now preferred! Unfortunately, it seems that the Borda rule is wholly dependent on irrelevant alternatives and preference reversals can occur with an alarming frequency.

It is also interesting to remember that the Borda rule can sometimes lose a fundamental property in a ranking procedure, i.e. *monotonicity*. This happens with scoring rules based on successive elimination; that is, ascending procedures which first find

Table 6.7 Fishburn numerical example of the Borda rule

Number of Criteria	3	2	2
	c	b	a
	b	a	d
	a	d	c
	d	c	b

Table 6.8 Frequency matrix derived from Table 6.7

Alternatives	***a***	***b***	***c***	***d***	
Ranking					Points
1st	2	2	3	0	3
2nd	2	3	0	2	2
3rd	3	0	2	2	1
4th	0	2	2	3	0

Table 6.9 Frequency matrix without *d*

Alternatives	***a***	***b***	***c***	
Ranking				Points
1st	2	2	3	2
2nd	2	3	2	1
3rd	3	2	2	0

the worst alternative, then eliminate it and start again in search for the second worst and so on. Fishburn (1982) proves that *any rule based on successive elimination by scoring methods must violate monotonicity with some profiles*. A well-known example is the ranking algorithm of the ELECTRE III method (Roy, 1978).

At this point, we need to tackle the issue of when, in a multi-criteria decision framework, it is better to use a Condorcet consistent rule or a scoring method. There is consensus in the literature that the Condorcet theory of voting is non-compensatory while Borda's is fully compensatory. A first conclusion is that when one wishes to use weights as *importance coefficients*, a Condorcet approach is required while Borda's is desirable when weights are meaningful in the form of *trade-offs*. Moreover, a Condorcet approach is useful for generating a ranking of the available alternatives (γ-problem formulation) while Borda's is more useful for isolating one alternative, considered the best (α-problem formulation) (Moulin, 1988; Truchon, 1995; Young, 1988, 1995). An immediate implication is that all ranking algorithms based on the concept of intensity of preference are inconsistent since they should deal with the α-problem formulation only (e.g. PROMETHEE (Brans et al., 1986); PRAGMA (Matarazzo, 1988); NAIADE (Munda, 1995). Moreover, in all these methods weights should only be introduced in the form of trade-offs). Finally, in concordance-based outranking methods, if one introduces ranking algorithms based on scoring methods, a basic inconsistency exists: *one is using Borda searching rules in order to find Condorcet solutions*.

In the framework of SMCE, I clearly choose to follow the Condorcet tradition[5] (since non-compensability and a complete ranking of alternatives are considered desirable properties in SMCE (Munda, 2004)).

However, as we have seen, a basic problem inherent in the Condorcet approach is the presence of cycles, i.e. cases where $a\boldsymbol{P}b$, $b\boldsymbol{P}c$ and $c\boldsymbol{P}a$ arise. This problem has been studied by various scientists (e.g. Kemeny, 1959; Fishburn, 1973a, b; Young and Levenglick, 1978; Moulin, 1985; Truchon, 1995; Vidu, 2002; Weber, 2002). The probability π (N,M) of obtaining a cycle with N alternatives and M criteria increases with N as well as with the number of criteria. Estimations of probabilities of finding cycles according to N alternatives and M voters can be found in Fishburn (1973b, p. 95). One should note that these probabilities are estimated under the so-

---

[5] Arrow and Raynaud (1986, pp. 77–78) also arrive at the conclusion that a Condorcet aggregation algorithm must in general be preferred in a multi-criterion framework. They give more general justifications than those I give here; this is because I prefer to stay within the scope of public policy.

On the completely opposite side one finds Saari (1989, 2000, 2006). His main criticism against Condorcet-based approaches is based on two arguments: (1) if one wants to preserve relationships among pairs (e.g. to impose a side constraint to protect gender balance for candidates in a public competition) then it is impossible to use pair-wise voting rules; a Borda count must be used. It is important to note that, although desirable in some cases, preserving a relationship among pairs implies the loss of neutrality; this is not desirable on general grounds in SMCE. (2) The individual rationality property (i.e. transitivity) must necessarily be weakened if one wishes to adopt a Condorcet based voting rule. This criticism can be serious. It will be discussed in detail in Sect. 6.3 when dealing with the cycle issue in Condorcet approaches.

called "*impartial culture assumption*", i.e. voters' opinions do not influence each other. While this assumption is unrealistic in a mass election, it is fully respected in a multi-criterion problem since criteria are supposed to be non-redundant. In Fishburn et al. (1979), the issue of cycles was tackled specifically for the multi-criterion problem, it has indeed been proved that the cycle issue is a serious problem for the use of Condorcet's voting theory in a multi-criterion framework, since with many alternatives and criteria, cycles occur with an extremely high frequency. For this reason, the multi-criteria methods based on Condorcet ideas need rules of thumb to solve cycles. Unfortunately these rules of thumb normally imply the loss of neutrality (among tied alternatives the choice is by alphabetic order) or anonymity (among tied alternatives the choice is by the order of criterion).

One of the original suggestions of Condorcet was to delete successively all weakest pair-wise majorities until the point that all cycles have been eliminated. In the example of Table 5.4, by applying this rule the cycle is broken at the weakest majority of 33 (***a*** over ***b***) and the final ranking is $b \rightarrow c \rightarrow a$. However, Young (1986) has proved that this rule is not valid when the number of alternatives is greater or equal to four.

The question is now: is it possible to tackle the cycle issue in a broader, more general fashion?

6.3 The Cycle Issue in Condorcet Consistent Rules

Condorcet himself was aware of the problem of cycles in his approach; he built examples to explain it (as the one shown in Table 6.4) and he was even close to finding a consistent rule capable of ranking any number of alternatives when cycles are present. However, attempts to understand fully this part of Condorcet's voting theory have arrived at conclusions like "… the general rules for the case of any number of candidates as given by Condorcet are stated so briefly as to be hardly intelligible … and as no examples are given it is quite hopeless to find out what Condorcet meant" (Nanson, quoted in Black, 1958, p. 175). Or "The obscurity and self-contradiction are without any parallel, so far as our experience of mathematical works extends … no amount of examples can convey an adequate impression of the evils" (Todhunter, 1949, p. 352 as cited by Young, 1988, p. 1234).

Attempts to clarify, fully understand and axiomatize Condorcet's approach for solving cycles have mainly been done by Kemeny (1959), who made the first intelligible description of the Condorcet approach, and by Young and Levenglick (1978), who gave its clearest explication and most complete axiomatization. For this reason I call this approach the Condorcet–Kemeny–Young–Levenglick ranking procedure, in short the C–K–Y–L ranking procedure[6].

[6] However, one should note that this voting rule is normally referred to in the literature as Kemeny's method or Kemeny's rule.

Its main methodological foundation is the maximum likelihood concept. In fact, the C–K–Y–L ranking procedure may be considered one of its earliest applications. "Condorcet's argument proceeds along the following lines. People differ in their opinions because they are imperfect judges of which decision really is best. If on balance each voter is more often right than wrong, however, then the majority view is very likely to identify the decision that is objectively best." (Young, 1988, p. 1232).

The maximum likelihood principle selects as a final ranking the one with the maximum pair-wise support. The selected ranking involves the minimum number of pair-wise inversions. Since Kemeny (1959) proposes the number of pair-wise inversions as a distance to be minimized between the selected ranking and the other individual profiles, the two approaches are perfectly equivalent. Formal proofs of this equivalence can be found in Truchon (1998b, pp. 6– 10) and Saari and Merlin (2000). The selected ranking is also a median ranking for those composing the profile (in multi-criteria terminology it is the "compromise ranking" among the various conflicting points of view); for this reason the corresponding ranking procedure is often known as the Kemeny median order.

Condorcet made three basic assumptions:

1. Voters' opinions do not influence each other.
2. Voters all have the same competence, i.e. each voter chooses his/her best candidate with a fixed probability p, where $½<p<1$ and p is the same for all voters.
3. Each voter's judgement on any pair of candidates is independent of his/her judgement on any other pair.

As discussed earlier, the first assumption always applies in a multi-criterion framework. The second might be a serious limitation in our case, since it implies equal criterion weights, while we assume that criterion weights may be different. In social choice terms the *anonymity* property (i.e. equal treatment of all criteria) is thus broken. Fortunately this is not problematic, since what is interesting of the C–K–Y–L ranking procedure in a multi-criterion framework, is not its empirical probabilistic interpretation but its possibility of being characterized as a *reasonable ranking method on some theoretical grounds, in the framework of SMCE*.

Indeed, given that full decisiveness tends to dictatorship, Arrow's impossibility theorem forces us to make a trade-off between decisiveness (an alternative has to be chosen or a ranking has to be made) and anonymity. As a consequence, the loss of anonymity in favour of decisiveness in our case may even be a positive property.[7] In general, it is essential that no criterion weight is more than 50% of the total

[7] One should note that in general the opportunity cost for decisiveness is the loss of one of the basic requirements of a social choice rule, i.e. anonymity, neutrality or monotonicity. Saari's (2006) defense of the Borda rule is based on the fact that it is less dangerous, or even sometimes desirable, to eliminate neutrality, and if one eliminates neutrality, then only a Borda rule can be adopted. But if one wishes to maintain neutrality and eliminate anonymity (as is desirable in an SMCE framework), then a Condorcet voting rule is appropriate.

weights; otherwise the aggregation procedure would become lexicographic in nature, and this criterion would become a dictator in Arrow's terminology. Of course, when criterion weights are derived from different dimensions, the requirement is that no dimension should weigh more than 50% of the total weights.

The third assumption refers to the axiom of independence of irrelevant alternatives. Of course the C–K–Y–L ranking procedure does not respect this axiom. However, two considerations have to be made on this subject:

1. A Condorcet consistent rule always presents smaller probabilities of the occurrence of a rank reversal in comparison with any Borda consistent rule (Moulin, 1988; Young, 1995). This is again a strong argument in favour of a Condorcet approach in a multi-criteria framework.
2. Young (1988, p. 1241) claims that the C–K–Y–L ranking procedure is the "only plausible ranking procedure that is locally stable", where *local stability* means that the ranking of alternatives does not change if only an interval of the full ranking is considered. It is interesting to note that this property was also studied by Jacquet-Lagrèze (1969), one of the first researchers in multi-criteria analysis, who called it the *median procedure*.

Personally, I think it is desirable that the final ranking is dependent on the alternatives considered, provided of course, that the ranking procedure is *neutral*. However, I share Young's opinion that if the ranking depends on the alternatives considered, this may cause manipulation of results by introducing extraneous, strange, extreme alternatives into the evaluation process.[8] In this sense, his advice to consider closely-related alternatives looks wise and should be considered during the evolution of the decision process when alternatives are generated. For instance, in choosing a car, nobody will realistically compare a Ferrari and a Ford Focus (some multi-criteria methods will probably supply an incomparability relation for a pair like this). It is much more plausible that a decision-maker will compare therefore a Ford Focus with a Volkswagen Golf, or a Ferrari with a Porsche. The fact of considering homogeneous alternatives for evaluation has thus another consequence: incomparability relations are no longer very important.

Saari and Merlin (2000) explicitly state that the C–K–Y–L ranking procedure (cited by them as the Kemeny's rule) enjoys "remarkable properties"; one of these being "consistency in societal rankings when candidates are dropped ....

[8] Arrow and Raynaud, on the grounds that the C–K–Y–L ranking procedure does not respect the independence of irrelevant alternatives axiom, arrive at the conclusion that this method should be discarded since it is not "*prudent*" (Arrow and Raynaud, 1986, p. 96). However, the whole class of ranking procedure they propose needs to be free of any intransitivity and cyclical information (Arrow and Raynaud, pp. 83–110). Thus from an operational point of view, I think that the C–K–Y–L method has more general applicability. In any case, Arrow and Raynaud's argument is even stronger against the adoption of a Borda rule, since preference reversal phenomena are very probable in a Borda framework.

To underscore this Kemeny's rule property, recall how dropping candidates can cause the Borda count societal ranking to radically change … The unexpected, troubling fact is that Kemeny's rule achieves its consistency by weakening the crucial assumption about the individual rationality of the voters" (Saari and Merlin, 2000, p. 404). Thus they conclude that "The Kemeny's rule structure and the consistency of the Kemeny's rule words are impressive; the reasons why they occur are worrisome" (Saari and Merlin, 2000, p. 431).

Indeed the argument made by Saari and Merlin against the C–K–Y–L ranking procedure is a serious one. First of all, let us understand what is meant by individual rationality. In Saari's words (Saari, 2000, p. 35) "*Transitivity* is a sequencing condition which requires the pair-wise rankings to mimic the ordering properties of points on the line. For instance, if a voter prefers **X** to **Y** and **Y** to **Z**, then the voter must prefer **X** to **Z**. A voter with transitive preferences is called *rational;* a voter with non-transitive preferences is called *irrational*".

As we saw in Chap. 5, the underlying assumption of this definition is the identification of human rationality with consistency, and this can be criticized from many points of view[9]. In particular, in Chap. 4 we noted that a down-to-earth preference modelling should imply the use of indifference and preference thresholds; this implies exactly the loss of the transitivity property of indifference and preference relations. Surprisingly, we could conclude from this that an appropriate preference modelling should be based on the "*weakening the crucial assumption about the individual rationality*" and that this should be desirable!

Moreover, one should have a clear idea of why a C–K–Y–L ranking procedure is needed; it answers to a precise problem of the original Condorcet proposal, i.e. the issue of cycles. It is then clear that we have to evaluate this procedure in the light of the cycle issue. As we know, cycles originate precisely in the transitivity of the preference relations, thus it is clear that any attempt to solve cycles has to weaken this property. The point is to do this with as little arbitrariness as possible, and this is exactly what the C–K–Y–L ranking procedure does.

Concluding, we can state that if

1. One accepts high probabilities of the occurrence of rank reversals
2. Transitivity of preference and indifference relations is considered essential
3. Neutrality can be abandoned

Then to use a Borda rule is simply the only option left. Otherwise, if one wishes to adopt a Condorcet based approach (and then to maintain neutrality and to suffer fewer rank reversal phenomena), the only acceptable rule to break cycles is the C–K–Y–L ranking procedure. The price to pay is, of course, the loss of the "rationality

[9] In Chap. 5, we referred in particular to Herbert Simon's criticism of the individual rationality assumption. A corroboration of this criticism in the framework of social choice can be found in Kelsey (1986), where it is stated that because of social choice problems, an individual with multiple objectives may find it impossible to construct a transitive ordering.

assumption". Whether this is an acceptable price or not, depends on the framework of application. In voting theory, it may be a high price to pay, but in the framework of multi-criteria decision analysis, it is definitely an acceptable price or even, under certain conditions, desirable.

Other properties of the C–K–Y–L ranking procedure are the following (Young and Levenglick, 1978):

- *Neutrality*: it does not depend on the name of any alternative, all alternatives are equally treated.
- *Unanimity* (sometimes called *pareto optimality*): if all criteria prefer alternative ***a*** to alternative ***b*** then ***b*** should not be chosen.
- *Monotonicity*: if alternative ***a*** is chosen in any pair-wise comparison and only the criterion scores of ***a*** are improved, then ***a*** should still be the winning alternative. Monotonicity is an essential property when dominated alternatives are not advised to be deleted from the analysis.
- *Reinforcement*: if the set *A* of alternatives is ranked by two subsets G_1 and G_2 of the criteria set *G*, such that the ranking is the same for both G_1 and G_2, then $G_1 \cup G_2 = G$ should still produce the same ranking. This general consistency requirement is very important in a multi-criteria framework where one may wish to apply the criteria belonging to each dimension first and then to pool them in the general model (see Munda, 1997b, for an example).

Although one can see that, the theoretical characterization of the C–K–Y–L ranking procedure is not easy, the algorithm *per se* is very simple. The maximum likelihood ranking of alternatives, in a multi-criteria framework, is that ranking supported by the maximum number of criteria for each pair-wise comparison, summed over all pairs of alternatives (for more details see the next chapter). By applying the C–K–Y–L ranking procedure to the numerical example of Table 6.4, the following six possible rankings with the corresponding scores are obtained.

a	*b*	*c*	100
b	***c***	***a***	**104**
c	*a*	*b*	86
b	*a*	*c*	94
c	*b*	*a*	80
a	*c*	*b*	76

The ranking $b \rightarrow c \rightarrow a$ is the final result. The original Condorcet problem has been solved satisfactorily.

To conclude let us examine the example of Table 6.10, with three alternatives and three criteria (this is the textbook example of the Condorcet paradox).

By applying the Borda rule, all alternatives receive a score equal to 3; no selection is possible. By applying the Condorcet rule, the majority being equal to 2/3, the cycle *a**P**b, b**P**c* and *c**P**a* is obtained. By applying both Simpson and Copeland

Table 6.10 Example of an insoluble ranking problem

Number of Criteria	1	1	1
	a	c	b
	b	a	c
	c	b	a

rules, all alternatives receive a score equal to 1; no selection is possible. By applying the C–K–Y–L ranking procedure, three rankings have the greatest support. These are: $a \to b \to c$, $b \to c \to a$, $c \to a \to b$ the cycle remains unsolved[10].

This example is a perfect materialization of Arrow's theorem; no decisiveness is possible. To eliminate ties, a larger number of criteria or some criterion weights are necessary. This is why I defend, when meaningful, the use of criterion weights, anonymity is lost but decisiveness improves enormously.

6.4 Arrow–Raynaud's Ranking Procedure

Arrow and Raynaud (1986) developed an original procedure explicitly designed to solve the discrete multi-criterion problem. Such a procedure is based on a set of axioms mainly built on previous research done by Köhler (1978). These axioms can be summarized as follows:

1. Alternatives are ranked through a step-by-step process.
2. At each step the information used refers only to alternatives not yet ranked.
3. The axiom of independence of irrelevant alternatives must apply. No preference reversal is possible. This is called the"*prudence*" axiom by Arrow and Raynaud (1986, p. 95).

The prudence axiom clearly discards both scoring methods and Condorcet consistent rules; Arrow and Raynaud thus proposed a new ranking algorithm. If no cycles exist, the ranking algorithm is the following. Given an outranking matrix, "Step r: Identify the maximum a_{ij} along each row of the current matrix. One at least from among these maxima is smaller than the others. If there are ties, one from among them is chosen arbitrarily. The row of this minimum corresponds to an alternative that will be ranked at the $(n - r + 1)th$ rank in the multicriterion ranking.

[10] It is important to note that the so-called ranking wheel, discussed by Saari (2006, p. 116–117), is simply this classical voting paradox as obtained when considering four alternatives and four criteria (voters). Of course Saari's criticism that a Condorcet approach is not adequate in this case, since it produces a cycle instead of a tie is perfectly correct. However, this criticism is not valid for the C–K–Y–L ranking procedure since in this case, the four rankings $a \to b \to c \to d$, $b \to c \to d \to a$, $c \to d \to a \to b$, $d \to a \to b \to c$ all receive the maximum score of 14, thus producing a tie correctly.

If $r<n$, delete the corresponding row and column of the outranking matrix, in order to obtain the current outranking matrix for the $(r+1)th$ step. The algorithm stops when the outranking matrix becomes void." (Arrow and Raynaud, 1986, p. 105).

Let us develop a numerical example starting with the outranking matrix presented in Table 6.11.

For each row (starting from the first to the fifth) the maxima are: 2.5, 4, 2.5, 4, 4. Since there is a tie, the first row can be chosen arbitrarily and alternative ***a*** is put into the lowest position in the ranking. The next is obviously alternative ***c***. At this stage the new outranking matrix is as shown in Table 6.12.

Now the corresponding maxima are 1.5, 3.5 and 3.5. The best alternative is unambiguously ***b***. After having eliminated ***b***, the only remaining comparison to be made is between ***d*** and ***e***. Obviously ***d*** is in first position and ***e*** in second. Finally, by applying Arrow and Raynaud's algorithm two rankings are produced:

$$d \to e \to b \to c \to a \text{ and } d \to e \to b \to a \to c.$$

The reader could at this point practice applying Arrow-Raynaud's algorithm to the outranking matrix of Table 6.6, which is derived from the original Condorcet approach. One may easily see that the ranking obtained is $b \to c \to a$, which is the same as that obtained through the application of the C–K–Y–L procedure but with a much shorter computation time. Now the question arises as to why not to use the Arrow–Raynaud procedure if even cycles can sometimes be solved so efficiently? To answer this question let us look at another numerical example.

Table 6.11 Outranking matrix for the Arrow–Raynaud algorithm

	a	*b*	*c*	*d*	*e*
a	0	1.5	2.5	1	1
b	3.5	0	4	1.5	1.5
c	2.5	1	0	1	1
d	4	3.5	4	0	3
e	4	3.5	4	2	0

Table 6.12 Outranking matrix after deleting alternatives *a* and *c*

	b	*d*	*e*
b	0	1.5	1.5
d	3.5	0	3
e	3.5	2	0

Table 6.13 Outranking matrix for a problem with three criteria and four alternatives

	a	*b*	*c*	*d*
a	0	2	2	1
b	1	0	1	2
c	1	2	0	2
d	2	1	1	0

Table 6.13 presents a numerical example of an outranking matrix with four alternatives and three criteria.

By considering the *4!* possible rankings, the C–K–Y–L procedure gives an unambiguous solution: the final ranking is $a \rightarrow c \rightarrow b \rightarrow d$. One should note however, that this clear-cut solution comes at an analytical cost, i.e. the loss of a feasible ranking. In fact, as noted by Arrow and Raynaud (1986, p. 110): "Rows 1 and 3 on one hand and 2 and 4 on the other hand are identical up to a permutation of their coefficients. Hence, in both cases, there should be a tie in the choice of a first element, and Kemeny's method should have given at least two solutions".

On the other hand, Arrow and Raynaud do not mention that, in this case, their procedure cannot produce any robust ranking at all given the presence of so many ties. Unfortunately, ties are very common when cycles are present. This diminishes the applicability of Arrow–Raynaud's algorithm significantly.

6.5 Discussion

Summarizing the results presented thus far, we can draw the following conclusions. Scoring methods present the advantage of always selecting one final solution (but not a final ranking), thus their degree of decisiveness is very high. However, one has to accept that a scoring method always implies transforming (arbitrarily) an original ordinal scale of measurement into a quantitative one, and this implies always having a compensatory aggregation rule. Compensability, which is based on the concept of intensity of preference, causes a high probability of preference reversal phenomena. Weights should always be in the form of trade-offs, such as in MAUT (see Chap. 5). Monotonicity is sometimes lost and neutrality can be relaxed. A strong argument in favour of a Borda scoring rule is that transitivity of the preference relation is never weakened, thus the assumption of individual rationality always applies.

All concordance-based methods should implement Condorcet consistent rules. These rules are adequate for finding rankings of alternatives. They present a lower probability of rank reversal than any scoring method. They are not compensatory, therefore weights can be treated as importance coefficients. A weak point is the high probability of the presence of cycles; their solution normally implies *ad hoc* rules of thumb. By means of the C–K–Y–L approach cycles can be tackled in a general way with no arbitrariness. Reinforcement is always respected by this ranking procedure. The axiom of independence of irrelevant alternatives is not fulfilled by the C–K–Y–L rule. In any case, this rule is much more stable than any Borda count; however the cost of this stability is the weakening of the individual rationality assumption. Moreover feasible rankings are sometimes lost. Neutrality cannot be relaxed, but anonymity can; this increases decisiveness greatly. The criticism that "… by concentrating on a particular pair, the pair-wise decision rule adopts a highly myopic perspective making it incapable of recognizing, or reacting to, the profile information manifesting the ranking wheel symmetry" (Saari, 2006, p. 117), is correct if referring to the original Condorcet approach (since it gives a cycle instead of an indifference relation), but is not valid at all if the C–K–Y–L rule is used. In fact, this rule correctly produces an indifference relation among the rankings with the maximum likelihood.

Arrow–Raynaud's method is the only ranking procedure fully respecting the axiom of independence of irrelevant alternatives; no preference reversal can occur. It is useful for the ranking decision problem formulation. It is not compensatory. However, it does not respect reinforcement, which as noted by Arrow and Raynaud themselves, is a very important characteristic when social decisions have to be made.[11] As a consequence the Arrow–Raynaud method can be considered more useful in the framework of private business decisions, while the C–K–Y–L ranking procedure is more appropriate in a social multi-criteria framework. Moreover, when cycles are present, the Arrow–Raynaud approach often cannot offer any clear-cut solution.

In the framework of social multi-criteria evaluation, we know that compensability should be limited and that a ranking should be produced; furthermore, the transitivity of preference relation can be weakened and neutrality should in principle always be maintained. Scoring methods are then, in the framework of SMCE, less appropriate than Condorcet based approaches to ranking feasible alternatives. Since reinforcement is very important in a social context and since cycles are very likely to occur, Arrow–Raynaud's method looks slightly less useful than the C–K–Y–L ranking procedure.

At this point, we have found a consistent ranking procedure to use in a SMCE framework. However, a serious problem is the computation of the C–K–Y–L ranking scores when many alternatives exist. One should note that the number of

[11] Arrow and Raynaud (1986, pp. 95–96) took into consideration the paper by Young and Levenglick (1978), but they arrive at the conclusion that reinforcement "… has definite ethical content and is therefore relevant to welfare economics and political science. But here our aim is operations research, of use to businessmen. We are unable to see why the "consistency" criterion has any compelling justification when efficiency is the prime consideration." (Arrow and Raynaud, 1986, p. 96).

permutations can easily become unmanageable; for example when ten alternatives exist, the number of permutations is 10! = 3,628,800. Moulin (1988, p. 312) clearly states that the Kemeny method (that I call the C–K–Y–L approach) is "the correct method" for ranking alternatives, and that the "only drawback of this aggregation method is the difficulty in computing it when the number of candidates grows". Indeed this computational drawback is very serious, since the Kemeny median order is NP-hard to compute.[12]

This NP-hardness has discouraged the development of algorithms to search for exact solutions; thus the majority of algorithms which have been proposed in the literature; are heuristics based on artificial intelligence, branch and bound approaches and multi-stage techniques (see e.g. Barthelemy et al., 1989; Charon et al., 1997; Truchon, 1998a; Cohen et al., 1999; Dwork et al., 2001; Davenport and Kalagnam, 2004). Thanks to all these efficient approaches, the computational problem can easily be dealt with in a multi-criteria framework.

Issues such as how to deal with mixed criterion scores or how to introduce weights that are definitively importance coefficients remain to be discussed and will thus be taken into consideration in Chap. 7. An attempt to combine the C–K–Y–L ranking procedure with a Borda count will also be illustrated in Chap. 7. This will be done by using the C–K–Y–L approach to obtain a ranking of alternatives and the Borda rule to implement the minority principle.

[12] The *complexity class* of *decision problems* that are intrinsically harder than those that can be solved by a *non-deterministic Turing machine* in *polynomial time*. When a decision version of a combinatorial *optimization problem* is proved to belong to the class of *NP-complete* problems, then the optimization version is NP-hard (definition given by the National Institute of Standards and Technology, *http://www.nist.gov/dads/HTML/nphard.html*).

Part C
Mathematical Procedures to Search for Technical and Social Compromise Solutions

"Salomon saith: There is no new thing upon the earth. So that as Plato had an imagination, that all knowledge was but remembrance; so Salomon giveth his sentence, that all novelty is but oblivion."

Francis Bacon *– Essays, LVIII, cited by Jorge Luis Borges – El Aleph, Alianza Editorial, Madrid, 1971, p. 7.*

"El final de la historia solo es referible en metáforas, ya que pasa en el reino de los cielos, donde no hay tiempo.….
… Mas correcto es decir que en el paraíso, Aureliano supo que para la insondable divinidad, el y Juan de Panonia (el ortodoxo y el hereje, el aborrecedor y el aborrecido, el acusador y la victima) formaban una sola persona."

Jorge Luis Borges *– Los teólogos in El Aleph, Alianza Editorial, Madrid, 1971, p. 48.*

Chapter 7
Searching for the "Technical Compromise Solution": Solving the Discrete Multi-Criterion Problem in an SMCE Framework

7.1 Pair-Wise Comparison of Alternatives

Given a set of evaluation criteria $G = \{g_m\}$, $m=1, 2,\ldots, M$, and a finite set $A=\{a_n\}$, $n=1, 2,\ldots, N$ of potential alternatives (actions), let us start with the simple assumption that the performance (i.e. the criterion score) of an alternative a_n with respect to a judgement criterion g_m is based on an *interval or ratio* scale of measurement. For simplicity's sake, it is assumed that a higher value of a criterion is preferred to a lower one (i.e. the higher, the better). The pair-wise comparison of alternatives proposed here is a preference modelling structure based on the so-called threshold model and fuzzy preference relations.

As shown in Chap. 4, by introducing a positive indifference threshold q the resulting preference model is the *threshold model:*

$$\left\{\begin{array}{lll} a_j P a_k & \Leftrightarrow & g_m(a_j) > g_m(a_k)+q \\ a_j I a_k & \Leftrightarrow & \left|g_m(a_j) - g_m(a_k)\right| \le q \end{array}\right\} \tag{7.1}$$

where a_j and a_k belong to the set A of alternatives and g_m to the set G of evaluation criteria.

The *double threshold model* is a preference relation where indifference and preference thresholds have been introduced, i.e.:

$$\left\{\begin{array}{lll} a_j P a_k & \Leftrightarrow & g_m(a_j) \; > \; g_m(a_k) \; + \; p\big(g_m(a_k)\big) \\ a_j Q a_k & \Leftrightarrow & g_m(a_k) \; + \; p\big(g_m(a_k)\big) \ge g_m(a_j) > g_m(a_k) + q\big(g_m(a_k)\big) \\ a_j I a_k & \Leftrightarrow & \left\{\begin{array}{l} g_m(a_k) \; + \; q\big(g_m(a_k)\big) \ge g_m(a_j) \\ g_m(a_j) \; + \; q\big(g_m(a_j)\big) \ge g_m(a_k) \end{array}\right. \end{array}\right\} \tag{7.2}$$

for any $m = 1, 2,\ldots, M$, p being a positive preference threshold.

A *pseudo-order structure* is a double threshold model upon which the following consistency condition is imposed

$$g_m(a_j) > g_m(a_k) \Leftrightarrow \begin{cases} g_m(a_j) + q\big(g_m(a_k)\big) > g_m(a_k) + q\big(g_m(a_k)\big) \\ g_m(a_j) + p\big(g_m(a_k)\big) > g_m(a_k) + p\big(g_m(a_k)\big) \end{cases} \tag{7.3}$$

A problem is that the modelling procedure based on the notion of a pseudo-criterion may display a serious lack of stability. Such undesirable discontinuities make a sensitivity analysis (or robustness analysis) necessary; however, this important analysis step is very complex in its execution because of the combinatorial nature of the various sets of data (Saltelli et al., 2004). One should combine variations of two thresholds (indifference and preference) and k possible scores of the M criteria. A solution to this problem may be found in the concept of *valued preference relations*, that is a preference relation in which it is necessary to assign to each ordered pair of alternatives (a_j, a_k) a value $v(a_j, a_k)$ representing the "strength" or the "degree of preference" (Fishburn, 1970, 1973a; Roubens and Vincke, 1985; Ozturk et al., 2005).

In this framework, an interesting concept is the one of a *fuzzy preference relation* (Kacprzyk and Roubens, 1988). If A is assumed to be a finite set of N alternatives, a *fuzzy preference relation* is an element of the $N{\times}N$ matrix $R = (r_{jk})$, i.e.:

$$r_{jk} = \mu_R(a_j, a_k) \qquad \text{with } j,k = 1,2,\ldots,N \text{ and } 0 \le r_{jk} \le 1 \tag{7.4}$$

$r_{jk} = 1$ indicates the maximum credibility degree of the preference of a_j over a_k; each value of r_{jk} in the open interval (0.5, 1) indicates a definite preference of a_j to a_k (a higher value means a stronger credibility); $r_{jk} = 0.5$ indicates the indifference between a_j and a_k. This definition implies that fuzzy preference relations can be used as mathematical models of intensity of preference.
Usually, fuzzy preference relations are assumed to satisfy two properties:

(1) Reciprocity, i.e. $r_{jk}+r_{kj} = 1$
(2) Max–min transitivity, i.e. if a_i is preferred to a_j and a_j is preferred to a_k, then a_i should be preferred to a_k with at least the same credibility degree, i.e.:

$$r_{ij} \ge 0.5, \quad r_{jk} \ge 0.5 \Rightarrow r_{ik} \ge \min\ (r_{ij}, r_{jk}) \tag{7.5}$$

Since small variations of input data (scores and thresholds) are modelled by means of a continuous membership function, by using fuzzy preference modelling as developed in (7.6), the combinatorial drawbacks of the pseudo-criterion model can be avoided.

Let us now consider any criterion g_m belonging to the set G and any pair of alternatives a_j and a_k belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are measured on an interval or ratio scale. Let p_m be a constant preference threshold and q_m a constant indifference threshold for the criterion g_m. Then the credibility degree $\boldsymbol{\mu}$ of preference (P) and indifference (I) relations between a_j and a_k can be computed as follows:

$$\begin{cases} \mu(a_j\, P\, a_k) = \left[1 + c_{pm} \left(g_m(a_j) - g_m(a_k)\right)^{-2}\right]^{-1} \\ \mu(a_j\, I\, a_k) = e^{-c_{qm}\left|g_m(a_j) - g_m(a_k)\right|} \\ \mu(a_k\, P\, a_j) = \left[1 + c_{pm} \left(g_m(a_k) - g_m(a_j)\right)^{-2}\right]^{-1} \end{cases} \tag{7.6}$$

where $\mu(a_j\, I\, a_k)\ \forall g_m(a_j)$ *and* $g_m(a_k)$ and

$$\mu(a_j\, P\, a_k) \qquad \textit{if} \quad g_m(a_j) - g_m(a_k) > 0 \tag{7.7}$$

$$\mu(a_k\, P\, a_j) \qquad \textit{if} \quad g_m(a_j) - g_m(a_k) < 0 \tag{7.8}$$

In (7.6) the parameters (c_{pm}) and (c_{qm}) are derived in function of the cross-over point, i.e. the value of the difference between two criterion scores where the credibility degree of the corresponding indifference/preference relation is equal to 0.5; see Figs. 7.1 and 7.2 for an example[1] (in these figures in the *y*-axis the credibility degrees and in the *x*-axis the thresholds are represented respectively). The relations $\mu(a_j P\ a_k)$and $\mu(a_k P\ a_j)$ are derived from values satisfying the condition of additive transitivity, thus it is trivial to prove that all these relations are max-min transitive (Kacprzyk and Roubens, 1988); however the property of reciprocity does not hold, thus these are not fuzzy preference relations in a strict sense. The relation $\mu(a_j I\ a_k)$ is a *resemblance relation*, which is reflexive and symmetrical but no transitivity is implied (thus the Luce paradox cannot occur).

It has to be admitted that the shape of the function representing the credibility degrees of the preference and indifference relations is arbitrary. However, there do exist some consistency requirements e.g. that the functions be continuous and monotonic and that $p_m > q_m$ thereby reducing considerably the degree of arbitrariness.

---

[1] Algebraically the parameters are the solution of this equation

$$\begin{cases} 0.5 = \left[1 + c_{pm} \left(p_m\right)^{-2}\right]^{-1} \\ 0.5 = e^{-c_{qm}\left|q_m\right|} \\ 0.5 = \left[1 + c_{pm} \left(p_m\right)^{-2}\right]^{-1} \end{cases}$$

where p_m and q_m are the preference and indifference thresholds respectively.

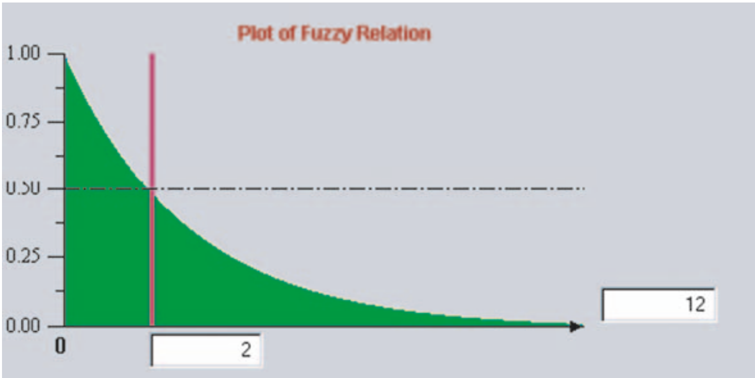


Fig. 7.1 Example of credibility degrees of a fuzzy indifference relation

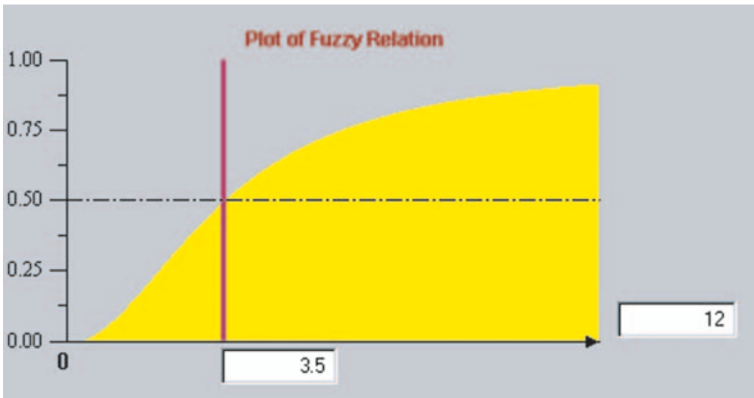


Fig. 7.2 Example of credibility degrees of a fuzzy preference relation

7.2 Extensions: The Case of Mixed Information on Criterion Scores

Ideally the information available for a policy problem should be precise, certain, exhaustive and unequivocal. But in real-world situations, it is often necessary to use information which lacks these characteristics and thus to deal with uncertainty of a stochastic and/or fuzzy nature in the data. Let us then introduce a more realistic assumption, i.e. that the set of evaluation criteria $G = \{g_m\}$, $m = 1, 2,\ldots, M$, on the set $A = \{a_n\}$, $n = 1, 2,\ldots, N$ of potential alternatives may include either crisp (i.e. impacts measured on interval or ratio scales), stochastic and fuzzy criterion scores.

The treatment of mixed information on criterion scores proposed here is mainly based on the semantic distance I developed some years ago (Munda, 1995, Chap. 6). This because this semantic distance allows us to deal consistently with an impact (or evaluation) matrix which may include crisp, stochastic or fuzzy measurements of the performance of an alternative with respect to an evaluation criterion. As a consequence the multi-criterion problem is considered in its more general form (the next section will show that ordinal criterion scores can also be considered). The only restriction is in the case of fuzzy information, when continuous, convex membership functions allowing for a definite integration are required.

Let us start with the case of fuzzy criterion scores (to complete the axiomatic system in Appendix 7.1 it is proved that this distance satisfies the property of triangle inequality):

if $\mu_1(x)$ and $\mu_2(x)$ are two fuzzy numbers, one can write (see Ragade and Gupta, 1977 for a formal proof):

$$f(x) = k_1\, \mu_1(x) \quad \text{and} \quad g(y) = k_2\, \mu_2(x) \tag{7.10}$$

where $f(x)$ and $g(y)$ are two functions obtained by rescaling the ordinates of $\mu_1(x)$ and
$\mu_2(x)$ through k_1 and k_2, such that

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} g(y)dy = 1 \tag{7.11}$$

The distance between all points of the membership functions is computed as follows:

If $f(x)$ is defined on $X = [x_L, x_U]$ and $g(y)$ is defined on $Y = [y_L, y_u]$
where sets X and Y can be non-bounded from one or either sides, then

$$S_d\big(f(x),\ g(y)\big) = \iint_{x,\,y} |x-y| f(x)g(y)dydx \tag{7.12}$$

If the intersection between the two membership functions is empty, it is $x > y\ \forall x \in X$ *and* $\forall y \in Y$, it follows that a continuous function in two variables is defined over a rectangle. Therefore the double integral can be calculated as iterated single integrals; the result is

$$S_d\big(f(x),\ g(y)\big) = |E(x)\ -\ E(y)| \tag{7.13}$$

where $E(x)$ and $E(y)$ are the expected values of the two membership functions.

When the intersection between two fuzzy sets is not empty, their distance is greater than the difference between the respective expected values, since $|x - y|$ is always greater than $(x - y)$. In this case one finds:

$$S_d\big(f(x),\ g(y)\big) = \int_{-\infty}^{+\infty} \int_{x}^{+\infty} (y-x) f(x)g(y)dydx \ + \\ + \int_{-\infty}^{+\infty} \int_{-\infty}^{x} (x-y) f(x)g(y)dydx \tag{7.14}$$

This is the case of a double integral over a general region; since this is not vertically or horizontally simple, its computation is not possible by means of iterated integration; it is necessary to take the limit of the Rieman sum. This problem can easily be overcome by means of numerical analysis (in Munda, 1995 a Monte Carlo type numerical algorithm for the computation of this distance was developed. This is presented in Appendix 7.2).

As an example, we will compute the semantic distance between a symmetrical and a LR fuzzy number. Let us assume:

$$\tilde{A} = e^{-0.077(x-143)^2} \qquad \text{and}$$

$$\tilde{B} = \begin{cases} 1 - e^{-1500\left(\frac{x-140}{140}\right)^2} & if\ 140 \le x < 147 \\ 1 & if\ x = 147 \\ \left[e^{-470\left(\frac{x-147}{147}\right)^2} \right]^2 & if\ 147 < x < +\infty \end{cases}$$

Then we have

$$S_d(\tilde{A},\tilde{B}) = \int_{147}^{+\infty}\int_{-\infty}^{+\infty} |x-y| \frac{\left[e^{-470\left(\frac{x-147}{147}\right)^2} \right]^2}{7.965} \bullet \frac{e^{-0.077(y-143)^2}}{6.349}\, dydx +$$

$$+ \int_{140}^{147}\int_{-\infty}^{+\infty} |x-y| \frac{1 - e^{-1500\left(\frac{x-140}{140}\right)^2}}{7.965} \bullet \frac{e^{-0.077(y-143)^2}}{6.349}\, dydx \cong 5$$

while the difference between their expected values is about 4.368.

From a theoretical point of view, the following main conclusions can be drawn:

(1) The absolute value metric is a particular case of the semantic distance
(2) The comparison between a fuzzy number and a crisp number is equal to the difference between the expected value of the fuzzy number and the value of the crisp number
(3) Stochastic information can also be taken into account

In sum the semantic distance allows us to deal with fuzzy numbers, probability distributions and crisp numbers with the theoretical guarantee that all these sources of information are tackled equivalently, thus solving an open problem for multi-criteria methods dealing with mixed information. Of course, this search for an equivalent treatment of available information implies a trade-off with precision. For example, if stochastic information only is available, a stochastic dominance approach is more effective (see e.g. Markowitz, 1989; Martel and Zaras, 1995), or if fuzzy numbers only have to be compared, Matarazzo and Munda (2001) present a more sophisticated approach based on area comparison. However, in the case of mixed information in a multi-criteria framework, the semantic distance illustrated here is probably the best available compromise solution between generality and precision. Moreover, the use of this semantic distance allows a homogeneous preference modelling on all the criteria, otherwise impossible; this can be illustrated as follows.

Going back to the pair-wise comparison of alternatives, let us assume $f(x) = g_m(a_j)$ and $g(y) = g_m(a_k)$, where g_m is any criterion belonging to the set G and a_j and a_k any pair of alternatives belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are *fuzzy or stochastic* in nature. Let p_m be a preference threshold and q_m an indifference threshold for the criterion g_m. Then we have:

$$\begin{cases} \mu(a_j\,P\,a_k) = \left[1 + c_{pm}\left(\iint_{x,y}(x-y)\,f(x)g(y)\,dydx\right)^{-2}\right]^{-1} \\ \mu(a_j\,I\,a_k) = e^{-c_{qm}\left(\iint_{x,y}|x-y|\,f(x)g(y)\,dydx\right)} \\ \mu(a_k\,P\,a_j) = \left[1 + c_{pm}\left(\iint_{y,x}(y-x)\,f(x)g(y)\,dydx\right)^{-2}\right]^{-1} \end{cases} \tag{7.15}$$

where

$\mu(a_j\,I\,a_k) \quad \forall \; x,y$ and

$$\mu(a_j\,P\,a_k) \;\; if \;\; \iint_{x,y}(x-y)f(x)g(y)\,dydx > 0 \tag{7.16}$$

$$\mu(a_k P\,a_j) \;\; if \;\; \iint_{x,y}(x-y)f(x)g(y)\,dydx < 0 \tag{7.17}$$

One should note that the comparison between the criterion scores of each pair of actions is carried out by means of the semantic distance. Since the absolute value metric is a particular case of this distance, fuzzy, stochastic and crisp criterion scores are dealt with equivalently.

7.3 Extensions: Introducing Weights as Importance Coefficients

At this point, a very delicate step has still to be tackled, i.e. the exploitation of the inter-criteria information in the form of weights. Let us then assume the existence of a set of criterion weights $W = \{w_m\}$, $m = 1, 2, \ldots, M$, with $\sum_{m=1}^{M} w_m = 1$ derived as importance coefficients. The problem is the theoretical guarantee that weights are really treated as importance coefficients and not as trade-offs. The point is that no connection can be made between criterion weights and the corresponding criterion intensity of preference. Our objectives are then:

(1) To find a way to combine weights with credibility degrees without a direct interpretation of the latter as intensity of preference
(2) To divide each criterion weight into two parts proportionally to the credibility degrees of the indifference and preference fuzzy relations. In doing so, the requirement that $\sum_{m=1}^{M} w_m = 1$ should not be lost.

Let us define μ_p as the fuzzy preference relation between a pair of alternatives and μ_I as the fuzzy indifference relation between the same pair. Let us put $\mu_{min} = \min(\mu_p, \mu_I)$ and $\mu_{max} = \max(\mu_p, \mu_I)$.

Clearly, it is $\mu_p = \mu_{min}$ on the left of the intersection point between the indifference and the preference fuzzy relations and vice versa on the right. I propose that a criterion weight w_m be divided proportionally to μ_p and μ_I according to (7.18).

$$\begin{cases} w_{m1} = w_m \dfrac{\mu_{\min}}{\mu_{\max} + \mu_{\min}} \\ \\ w_{m2} = w_m \dfrac{\mu_{\max}}{\mu_{\max} + \mu_{\min}} \end{cases} \tag{7.18}$$

(7.18) presents the following properties:

$$w_{m1} + w_{m2} = w_m \tag{7.19}$$

$$if \ \mu_{\min} = 0 \ \Rightarrow w_{m2} = w_m \tag{7.20}$$

$$if \ \mu_{\min} = \mu_{\max} = 0 \ \Rightarrow \ w_m = 0 \tag{7.21}$$

$$if \ \mu_{\min} = \mu_{\max} \ \Rightarrow \ w_{m1} = w_{m2} = \frac{1}{2} w_m \tag{7.22}$$

As a consequence (7.18) meets our objective of keeping the sum of weights perfectly equal to one. Moreover, in (7.18) no direct use of the concept of intensity of preference is made; as a result we can be sure that criterion weights are being used consistently with their nature as importance coefficients. Finally if a criterion score is *ordinal in nature*, it can be considered a particular case where $\mu_{\min} = 0$. Again *the treatment of crisp, fuzzy, stochastic and ordinal criterion scores is perfectly* *equivalent*. Moreover, when indifference and preference thresholds are not used, the corresponding criteria can be dealt with as ordinal criteria[2], where

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \\ a_j I a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \end{cases} \tag{7.23}$$

Now an $N \times N$ matrix E can be built, where any generic element e_{jk} with $j \neq k$ is the result of the pair-wise comparison between alternatives j and k according to all the M criteria. Such a global pair-wise comparison is obtained by means of (7.24):

$$e_{jk} = \sum_{m=1}^{M} \left(w_m(P_{jk}) + \frac{1}{2} w_m(I_{jk}) \right) \tag{7.24}$$

where $w_m(p_{jk})$ and $w_m(I_{jk})$ are derived from μ_P and μ_I through (7.18). It is

$$e_{jk} + e_{kj} = 1 \tag{7.25}$$

Property 7.25 is very important since it allows us to consider matrix E as a *voting matrix* i.e. a matrix where instead of using criteria, alternatives are compared by means of voters' preferences (on the principle of one agent, one vote). This analogy between the multi-criterion and the social choice problem, as noted by Arrow and Raynaud (1986), is very useful for tackling the step of ranking the N alternatives in a consistent axiomatic framework: a Condorcet consistent rule can now be used to exploit the pair-wise comparisons to order alternatives.

7.4 Ranking of Alternatives in a Complete Pre-Order

The issue here is whether it is possible to find a ranking algorithm consistent with the desirable properties of social multi-criteria evaluation. And conversely, given the results of Arrow's impossibility theorem (Arrow, 1963), whether it is possible to ensure that no essential property is lost. Both social choice literature and multi-criteria decision theory agree that whenever the majority rule can be operationalized,

---

[2] If criterion scores are used with an ordinal meaning only, as we saw in Chap. 6, the following axiomatic conditions must be added (adapted from Arrow and Raynaud, 1986, p. 81–82).

Axiom 1: Diversity. Each criterion is a total order on the finite set A of alternatives to be ranked, and there is no restriction on criteria; they can be any total order on A.

Axiom 2: Symmetry. Since criteria have incommensurable scales, the only preference information they provide is the ordinal pair-wise preferences they contain.

Axiom 3: Positive Responsiveness. The degree of preference between two alternatives ***a*** and ***b*** is a strictly increasing function of the number and weights of criteria that rank ***a*** before ***b***.

it should be applied. However, majority rule often produces undesirable intransitivities, thus "more limited ambitions are compulsory. The next highest ambition for an aggregation algorithm is to be Condorcet" (Arrow and Raynaud, 1986, p. 77). As we have discussed in Chap. 6, in the framework of SMCE the C–K–Y–L ranking procedure seems the most appropriate.

According to this ranking procedure, the maximum likelihood ranking of alternatives, in a social multi-criterion framework, is that ranking supported by the maximum number of criteria for each pair-wise comparison, summed over all pairs of alternatives. More formally, the C–K–Y–L ranking procedure can be adapted to a multi-criteria framework as follows.

All the $N(N-1)$ pair-wise comparisons compose the matrix E, in which we remember that $e_{jk} + e_{kj} = 1$, with $j \neq k$. Let us call R the set of all the $N!$ possible complete rankings of alternatives: $R = \{r_s\}$, $s = 1, 2,\ldots, N!$. For each r_s, let us compute the corresponding score φ_s as the sum of e_{jk} over all the $\binom{N}{2}$ pairs jk of alternatives, i.e.

$$\varphi_s = \sum e_{jk} \tag{7.26}$$

where $j \neq k$, $s = 1, 2, \ldots N!$ and $e_{jk} \in \mathrm{r}_s$

The final ranking (r^*) is that[3] which maximizes (7.26), which is:

$$r_* \Leftrightarrow \varphi_* = \max \sum e_{jk} \qquad \text{where } e_{jk} \in R \tag{7.27}$$

A final issue to be discussed is the matter of ties, i.e. the case that in some individual profiles alternatives can be ranked in the same position. This does not constitute a problem since such an event can easily be taken into account in the concordance index used for the construction of an outranking matrix (see (7.24)).

However, if some ties occur in the outranking matrix E, this might sometimes create a problem for the interpretation of the final results. In this case, one has to choose a tie-breaking rule, thus neutrality is necessarily lost. The question is then: which is the probability of finding ties in the outranking matrix E? Proposition 7.1 states that this probability is approximately zero; as a consequence ties in the outranking matrix are not a serious problem. The proof of this proposition can be found in Appendix 7.1.

Proposition 7.1 *In the outranking matrix E the event of obtaining ties, that is,* $e_{jk} = e_{kj} = \frac{M}{2}$, *is possible but its probability is approximately zero.*

---

[3] It is important to remember that sometimes the final ranking is not unique. This is a desirable property since it can be considered a measure of the *robustness* of the results provided.

7.5 Introducing the Minority Principle: A Borda Approach

At this point, we refer to the normative tradition in political philosophy, which also has an influence in modern social choice (Moulin, 1981) and public policy (Mueller, 1978). The fundamental idea is that any coalition controlling more than 50% of the votes may be converted into an actual dictator. As a consequence, the "remedy to the tyranny of the majority is the minority principle, requiring that all coalitions, however small, should be given some fraction of the decision power. One measure of this power is the ability to veto certain subsets of outcomes.… "(Moulin, 1988, p. 272). The introduction of a veto power in a multi-criteria framework can be further justified in the light of the so-called "prudence" axiom (Arrow and Raynaud, 1986, p. 95), whose principle is that it is not prudent to accept alternatives whose degree of conflictuality is too high (and thus might make the final decision very vulnerable[4]). The point is then how to implement this idea of veto power in a multi-criteria framework.

Historically, the first attempt was made by Roy (1985, 1996) in the so-called ELECTRE methods. Basically, Roy proposed that for any pair of alternatives one should look at the majority principle expressed as a concordance index and to the minority principle in the form of the discordance index. The discordance index is calculated according to the intensity of preference any single criterion has against the concordance coalition. This means that for each single criterion a veto threshold must be defined.

In my opinion, the implementation of veto power in an SMCE framework presupposes three desirable characteristics:

1. To be independent of arbitrary ad hoc thresholds.
2. To consider the global opposition to the final ranking and not to a pair of alternatives (Roy's approach), or any specific possible ranking (Paelinck, 1978).
3. No specific intensity of preference should be considered (if a weight is combined with a veto threshold for each criterion, the resulting concept of criterion importance also depends on the intensity of preference; this means that weights probably can no longer be considered importance coefficients).

It is interesting to note that an approach meeting these requirements can again be found in classical social choice theory, in particular, in the Borda approach. The Borda rule is normally used to find a Borda winner, where the winner is the alternative which receives the highest score in favour (an alternative receives $N - 1$ points if it ranks first, and so on until *0* score if it ranks last on a given criterion). In the same way, a *Borda loser* can be defined as the alternative which receives the highest score against (where $N - 1$ points are assigned to the last alternative in the ranking and so on until 0 points are given to the option which ranks first).

---

[4] It should be noted that mitigating the vulnerability of the C–K–Y–L ranking procedure is very important since this is one of the main criticisms of the method.

Formally, the procedure I am proposing can be described as follows by taking inspiration from the concept of frequency matrices (Hinloopen et al., 1983; Matarazzo, 1988). Let us call F the matrix where any element f_{ij} means that a given criterion g_m scores alternative a_j in the i-th ordinal position. Now it is possible to define the $N \times N$ matrix Φ where any element ϕ_{ij} represents the summation of the weights of criteria which score alternative j at the i-th position; that is

$$\phi_{ij} = \sum_{m \in G_i} w_m \tag{7.28}$$

$$where \;\; G_i = \left\{ g_m : g_m(a_j) = f_{ij} \right\} \quad with \quad G_i \subset G \tag{7.29}$$

$i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, N$

It is obviously:

$$\sum_{i=1}^{N} \phi_{ij} = 1 \quad \forall \; a_j \in A \quad \text{and} \tag{7.30}$$

$$\sum_{j=1}^{N} \phi_{ij} = 1 \quad with \;\; j = 1, 2, \ldots, N \tag{7.31}$$

Now for any alternative a_j let us apply the Borda rule in search for the Borda loser, i.e.

$$B(a_j) = \sum_{i=1}^{N} (\phi_{ij} \times b_i) \tag{7.32}$$

$$where \quad b_i = N-1, \; N-2, \ldots, 0 \quad with \; i = N, \; N-1, \ldots, 1$$

The vetoed alternative $\bar{a}_j$ is the Borda loser, i.e. the a_j for which $B(a_j) = \max$.

One should note that by means of this procedure weights are never combined with intensities of preference and no *ad hoc* parameter is needed. Consistently with the Borda approach, only one alternative, considered the one with the greatest opposition, is selected to be vetoed. It must be remembered that the Borda procedure respects all the properties of the C–K–Y–L, except local stability. This is the main reason why Borda consistent rules are more appropriate for the selection of one alternative only and not for the generation of rankings.

Finally a question to be answered is: do Borda and Condorcet rules normally lead to different solutions? One might in fact believe that the divergence of solutions is a very special case and thus the value added of introducing the Borda loser is very limited. As we have seen in Chap. 6, this question can be answered very easily. Fishburn (1973b) proves the following theorem: *there are profiles where the Condorcet winner exists and it is never selected by any scoring method*. Moulin (1988, p. 249) proves that “a Condorcet winner (loser) cannot be a Borda loser (winner)”. In other words, Condorcet consistent rules and scoring voting rules are deeply different in nature.

7.6 Numerical Examples

Let us consider an evaluation problem concerning three types of publicly provided goods (A, B, C). Let us assume that it has been agreed that these goods have to be evaluated by taking into account three dimensions, i.e. economic, social and environmental, and that each dimension has the same weight.

These dimensions are operationalized by means of the following evaluation criteria:

1. ***Financial cost*** (economic dimension), weight = 0.167; its criterion scores are in millions of Euro measured in crisp terms, indifference threshold = €250,000, preference threshold = €500,000 (see Fig. 7.3).
2. ***Employment*** (economic dimension), weight = 0.167; its criterion scores are in number of persons/year, measured by means of symmetric fuzzy numbers, indifference threshold = 30 persons/year, preference threshold = 50 persons/year (see Fig. 7.4).
3. ***Avoidance of social exclusion*** (social dimension), weight = 0.333; its criterion scores are qualitative, measured by means of an ordinal scale of measurement (good better than moderate).

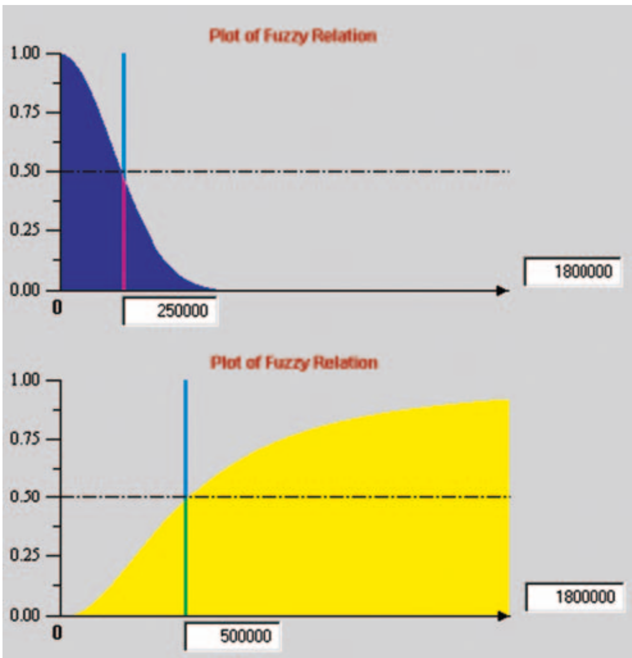


Fig. 7.3 Indifference and preference fuzzy relations on the criterion "financial cost"

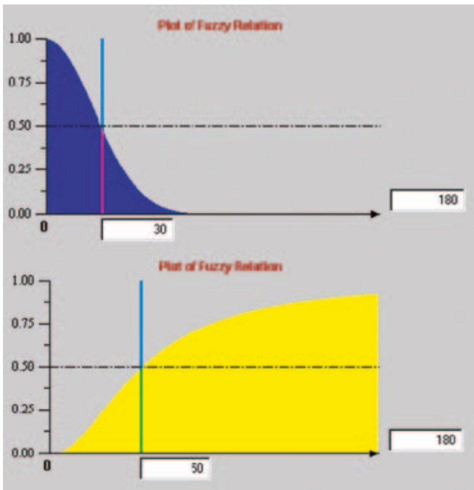


Fig. 7.4 Indifference and preference fuzzy relations on the criterion "employment"

Table 7.1 Evaluation matrix of a hypothetical public policy problem

Alternatives	A	B	C
Criteria			
financial cost	13.2	13.5	15
employment	approx. 100	approx. 135	approx. 200
avoidance of social exclusion	Very Good	Moderate	Good
environmental impact	2-nd	1-st	3-rd

Table 7.2 Values of the semantic distance for the criterion "financial cost"

	Semantic distance
(A,B)	0.30
(A,C)	1.80
(B,C)	1.50

4. ***Environmental impact*** (environmental dimension), weight = 0.333; its criterion scores are qualitative, measured by means of an ordinal scale of measurement (1° better than 2°).

This policy problem can be summarized in the evaluation matrix described in Table 7.1.

Let us now compare each pair of alternatives according to each criterion. For the ordinal criterion scores the comparison is obvious. For the other criteria, let us apply the semantic distance. The results are presented in Tables 7.2 and 7.3.

Now it is possible to compute the fuzzy preference and indifference relations. Values are given in Tables 7.4 and 7.5.

Table 7.3 Values of the semantic distance for the criterion "employment"

	Expected value difference	Semantic distance
(A,B)	–35.0081	56.2391
(A,C)	–99.9977	107.9972
(B,C)	–64.9896	87.8337

Table 7.4 Values of the fuzzy relations for the criterion "financial cost"

	μ_P	μ_I
(A,B)	0.2647	0.5946
(A,C)	0.9284	0.0442
(B,C)	0.9000	0.0743

Table 7.5 Values of the fuzzy relations for the criterion "employment"

	μ_P	μ_I
(B,A)	0.3290	0.3827
(C,A)	0.8000	0.1570
(C,B)	0.6282	0.2224

Through (7.18), let's now compute the relative weights on each criterion for any pair of alternatives.

Criterion 1: "Financial Cost".

$$w_1(P)_{(A,B)} = \frac{0.2647}{0.2647 + 0.5946} \times 0.167 = 0.051$$

$$w_1(I)_{(A,B)} = \frac{0.5946}{0.2647 + 0.5946} \times 0.167 = 0.116$$

$$w_1(P)_{(A,C)} = \frac{0.9284}{0.9284 + 0.0442} \times 0.167 = 0.159$$

$$w_1(I)_{(A,C)} = \frac{0.0442}{0.9284 + 0.0442} \times 0.167 = 0.0075$$

$$w_1(P)_{(B,C)} = \frac{0.9000}{0.9000 + 0.0743} \times 0.167 = 0.154$$

$$w_1(I)_{(B,C)} = \frac{0.0743}{0.9000 + 0.0743} \times 0.167 = 0.013$$

Criterion 2: "Employment".

$$w_2(P)_{(B,A)} = 0.0762$$
$$w_2(I)_{(A,B)} = 0.08872$$
$$w_2(P)_{(C,A)} = 0.1379$$
$$w_2(I)_{(A,C)} = 0.0270$$
$$w_2(P)_{(C,B)} = 0.1218$$
$$w_2(I)_{(B,C)} = 0.04319$$

Criterion 3: “Avoidance of Social Exclusion”.

$$w_3(P)_{(A,B)} = 0.333$$
$$w_3(P)_{(A,C)} = 0.333$$
$$w_3(P)_{(C,B)} = 0.333$$

Criterion 4: “Environmental Impact”.

$$w_4(P)_{(B,A)} = 0.333$$
$$w_4(P)_{(A,C)} = 0.333$$
$$w_4(P)_{(B,C)} = 0.333$$

By applying (7.24) the following results are obtained (see Table 7.6):

By applying the C–K–Y–L rule to the *3!* possible rankings it is:

$$\text{ABC}\,\varphi_1 = 0.485 + 0.513 + 0.841 = 1.83$$
$$\text{BCA}\,\varphi_2 = 0.513 + 0.159 + 0.515 = 1.18$$
$$\text{CAB}\,\varphi_3 = 0.159 + 0.485 + 0.487 = 1.13$$
$$\text{ACB}\,\varphi_4 = 0.841 + 0.487 + 0.485 = 1.81$$
$$\text{BAC}\,\varphi_5 = 0.515 + 0.841 + 0.513 = 1.86$$
$$\text{CBA}\,\varphi_6 = 0.487 + 0.515 + 0.159 = 1.16$$

The final ranking r^* is then BAC

Let us now look for the Borda loser. Matrix F is presented in Table 7.7:

Computing the elements ϕ_{ij} of matrix Φ, we obtain:

Table 7.6 Matrix E of a hypothetical public policy problem

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.485 & 0.841 \\ B & 0.515 & 0 & 0.513 \\ C & 0.159 & 0.487 & 0 \end{bmatrix}$$

Table 7.7 Matrix *F* of a hypothetical public policy problem

Alternatives	A	B	C
Criteria			
financial cost	1-st	2-nd	3-rd
employment	3-rd	2-nd	1-st
avoidance of social exclusion	1-st	3-rd	2-nd
environmental impact	2-nd	1-st	3-rd

$$\phi_{1A} = 0.167 + 0.333 = 0.5$$
$$\phi_{2A} = 0.333$$
$$\phi_{3A} = 0.167$$
$$\phi_{1B} = 0.333$$
$$\phi_{2B} = 0.167 + 0.167 = 0.33333$$
$$\phi_{3B} = 0.333$$
$$\phi_{1C} = 0.167$$
$$\phi_{2C} = 0.333$$
$$\phi_{3C} = 0.333 + 0.167 = 0.5$$

Then matrix Φ is the following $\Phi = \begin{bmatrix} 0.5 & 0.333 & 0.167 \\ 0.333 & 0.333 & 0.333 \\ 0.167 & 0.333 & 0.5 \end{bmatrix}$

By applying(7.32), we have:
$B(\text{A}) = 0.333 \times 1 + 0.167 \times 2 = 0.666$
$B(\text{B}) = 0.333 \times 1 + 0.333 \times 2 = 1$
$\boldsymbol{B(C)} = 0.333 \times 1 + 0.5 \times 2 = \mathbf{1.333}$
The Borda loser is alternative C which, in this case, is also the C–K–Y–L loser.

Let us now look at a completely ordinal example. As discussed in Box 4.1, with composite indicators it is essential to use weights as importance coefficients. Moreover, indifference and preference thresholds would increase the degree of arbitrariness too much, thus a proper ranking procedure for composite indicators should be ordinal in nature (Munda and Nardo, 2003). Let us then apply the C–K–Y–L ranking procedure to the urban sustainability assessment example with the criterion weights illustrated in Sect. 4.5. The corresponding outranking matrix is presented in Table 7.8.

The 24 possible rankings and the corresponding scores φ_s are the following (where A is Budapest, B is Moscow, C is Amsterdam and D is New York):

Table 7.8 Weighted outranking matrix

	Budapest	Moscow	Amsterdam	New York
Budapest	0	0.3	0.4	0.4
Moscow	0.7	0	0.5	0.6
Amsterdam	0.6	0.5	0	0.3
New York	0.6	0.4	0.7	0

B	**D**	**C**	**A**	**3,6**	B	C	A	D	2,9
D	B	C	A	3,5	C	B	A	D	2,9
D	C	B	A	3,5	A	B	D	C	2,9
B	D	A	C	3,5	B	A	C	D	2,8
D	B	A	C	3,4	A	D	B	C	2,8
B	A	D	C	3,3	A	D	C	B	2,8
B	C	D	A	3,2	C	D	A	B	2,7
C	B	D	A	3,2	C	A	B	D	2,6
D	C	A	B	3,2	C	A	D	B	2,5
C	D	B	A	3,1	A	B	C	D	2,5
D	A	B	C	3,1	A	C	B	D	2,5
D	A	C	B	3,1	A	C	D	B	2,4

In comparison with the results obtained by applying the linear aggregation rule, without any criterion weights as described in Chap. 1, Moscow is still in top position, but this time Budapest is at the bottom. New York again scores better than Amsterdam.

Note that the use of weights and the improvement of the mathematical aggregation procedure (in comparison with the simple linear aggregation rule) do not change the results spectacularly. The structuring process, and in this case above all, the input information used for the indicator scores clearly determine the final ranking. "*Garbage in, garbage out*" phenomena are almost impossible to avoid (Funtowicz and Ravetz, 1990). This is a fundamental lesson to bear in mind in real-world applications of SMCE. Good mathematical algorithms guarantee consistency with the problem structuring and nothing else. Of course, *ceteris paribus*, the mathematical properties of a ranking algorithm may make an important difference.

Let us conclude by examining thereafter examples from the field of composite indicators (Munda and Nardo, 2003). Let us take into consideration a simple hypothetical example with three countries (A, B, C) to be ranked according to a composite sustainability indicator. Let us assume that three dimensions have to be considered, i.e. economic, social and environmental, and that each dimension should have the same weight, i.e. 0.3333.

The following individual indicators are used:

Economic dimension

Indicator: GDP per capita. Weight: 0.167. Objective: maximization of economic growth. Variable: US dollar per year.

Indicator: Unemployment rate. Weight: 0.167. Objective: minimization of unemployed people. Variable: percentage of population.

Environmental dimension

Indicator: Solid waste generated per capita. Weight: 0.333. Objective: minimization of environmental impact. Variable: tons per year.

Social dimension

Indicator: Income disparity. Weight: 0.167. Objective: minimization of distributional inequity. Variable: Q5/Q1.

Indicator: Crime rate. Weight: 0.167. Objective: minimization of criminality. Variable: robberies per 1000 inhabitants.

The impact matrix described in Table 7.9 can then be constructed.

The pair-wise comparison results can be summarized in the following outranking matrix:

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.666 & 0.333 \\ B & 0.333 & 0 & 0.333 \\ C & 0.666 & 0.666 & 0 \end{bmatrix}$$

By applying the C–K–Y–L rule to the 3! possible rankings we obtain:

$$\text{ABC}\ \varphi_1 = 0.666 + 0.333 + 0.333 = 1.333$$
$$\text{BCA}\ \varphi_2 = 0.333 + 0.333 + 0.666 = 1.333$$
$$\textbf{CAB}\ \varphi_3 = 0.666 + 0.666 + 0.666 = 2$$
$$\text{ACB}\ \varphi_4 = 0.333 + 0.666 + 0.666 = 1.666$$
$$\text{BAC}\ \varphi_5 = 0.333 + 0.333 + 0.333 = 1$$
$$\text{CBA}\ \varphi_6 = 0.666 + 0.666 + 0.333 = 1.666$$

The final ranking r^* is then CAB.

Note that using one of the standard ways to produce a composite indicator would produce a different result. If the composite indicator for each country is calculated in terms of the difference from the group leader (which assigns 100 to the leading country and ranks the others in percentage points away from the leader), the impact matrix becomes as shown in Table 7.10.

The index will be calculated by averaging each indicator (with the same weights as in the multi-criterion matrix), obtaining $\boldsymbol{I}_A = 69.8$, $\boldsymbol{I}_B = 79.7$, and $\boldsymbol{I}_C = 81.9$. The ranking would be CBA, different from the ranking obtained with the other algorithm.

Table 7.9 Impact matrix of the illustrative numerical example

Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries					
A	22,000	0.17	0.4	10.5	40
B	45,000	0.09	0.45	11.0	45
C	20,000	0.08	0.35	5.3	80

Table 7.10 Impact matrix: distance from the leader

Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries					
A	48.9	47.05	87.5	50.5	100
B	100	88.9	77.8	48.2	88.9
C	44.4	100	100	100	50

Table 7.11 Performance in the knowledge-based economy: a tentative indicator

Indicators	BERD	Researchers	Patents	HT
Countries				
Australia	23.1	22.0	8.3	64.5
Austria	38.7	27.5	24.4	76.3
Belgium	46.5	40.2	32.0	22.3

Consider another example, the *composite indicator of Industrial innovation* (OECD, 2003). This composite indicator is based on four sub-indicators: Business enterprise R&D as percentage of GDP (*BERD*), the number of business researchers per 10,000 labour force (*Researchers*), the number of patents per million population (*Patents*), and the share of firms having introduced at least one new or improved product or process on the market *(HT)*. For sake of simplicity let us take the first three countries of OECD classification (see Table 7.11) (see OECD, 2003, p. 19)[5]:

The composite index is the simple average of indicators (thus we have a case of equal weights): 41.7 for Austria, 35.2 for Belgium, and 29.5 for Australia.

However, the outranking matrix with weight equal to ¼ for each index is as follows:

$$E = \begin{bmatrix} & AU & A & B \\ AU & 0 & 0 & 0.25 \\ A & 1 & 0 & 0.25 \\ B & 0.75 & 0.75 & 0 \end{bmatrix}$$

From the comparison of 3! possible combinations it turns out that the one with the highest score is (B, A, AU) with 2.25[6]. Again the ranking produced with the standard methods of summing up normalized variables is different from that produced with the C–K–Y–L ranking procedure.

[5] The indicators in the matrix shown in Table 7.11 have been normalized with the min-max method which ranks each country with respect to the global maximum (the leader = 100) and the global minimum (the laggard = 0). The index is calculated as: (actual value – minimum value)/(maximum value – minimum value)*100. Note that none of the countries chosen is either a maximum or a minimum.

[6] The outranking matrix is the same for the original data and for the normalized indicators.

Let us conclude with a real-world example which shows the importance of a computational algorithm: the "Environmental Sustainability Index" (ESI). The index for 2005 was produced by a team of environmental researchers from Yale and Columbia Universities, in co-operation with the World Economic Forum and the Joint Research Centre of the European Commission.

The aim of the ESI is to benchmark the ability of 146 nations to protect the environment over the next decades, by integrating 76 data sets into 21 indicators of environmental sustainability (see Esty et al., 2005). The database used to construct the ESI covers a wide range of aspects of environmental sustainability ranging from the physical state and stress of the environmental systems (like natural resource depletion, pollution, ecosystem destruction) to the more general social and institutional capacity to respond to environmental challenges. Poverty, short-term thinking and lack of investment in capacity and infrastructure committed to pollution control and ecosystem protection thus compete to determine the measure of a country's sustainability.

Although the official ESI ranking is based upon the linear aggregation of 21 equally weighted indicators, an attempt has been made, in the methodological appendix, to apply the non-compensatory approach presented in this chapter, in order to tackle the issues of weights as "importance measure" and the compensability of different and crucial dimension of environmental sustainability (see the Methodological Appendix in Esty et al., 2005).

Figure 7.5 compares the ranking obtained by means of the non-compensatory aggregation rule with that of the ESI2005. In both cases all 21 indicators are equally weighted. From this figure it is clear that the aggregation method used affects principally the middle-of-the-road and, to a lesser extent, the leader and the laggard countries. Overall, for the set of 146 countries, the assumption underlying the

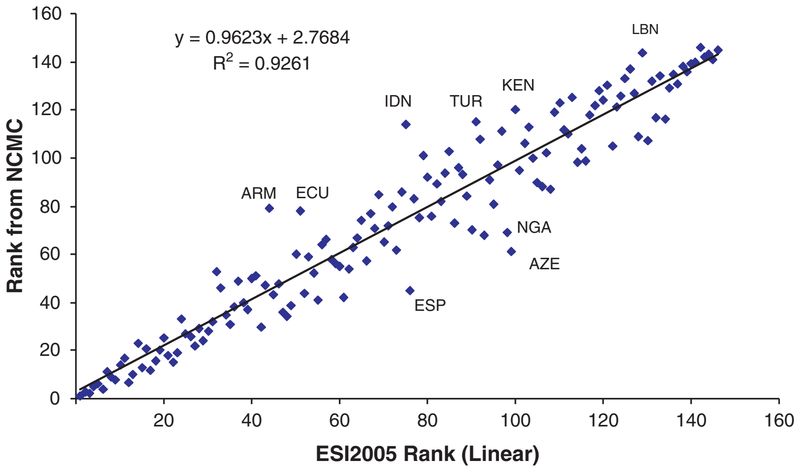


Fig. 7.5 Comparison of rankings obtained by the linear aggregation (ESI2005 on the *x*-axis) and the non-linear/non-compensatory –C–K–Y–L– (NCMA on the *y*-axis) rules

Table 7.12 ESI rankings obtained by linear aggregation (LIN) and the C–K–Y–L ranking procedure: countries that greatly improve or greatly worsen their rank position

	Aggregation	ESI rank with LIN	Rank with NCMC	Change in rank
Improvement	Azerbaijan	99	61	38
	Spain	76	45	31
	Nigeria	98	69	29
	South Africa	93	68	25
	Burundi	130	107	23
Deterioration	Indonesia	75	114	39
	Armenia	44	79	35
	Ecuador	51	78	27
	Turkey	91	115	24
	Sri Lanka	79	101	22
	Average change over 146 countries			8

aggregation scheme has an average impact of eight ranks and a rank-order correlation coefficient of 0.962. In particular, while the top 50 countries move on average only five positions, the next 50 countries on average move twelve positions and the remaining 46 countries eight positions.

It is important to underline that although both aggregation schemes seem to produce consistent rankings (the R^2 is 0.92), those rankings do not nevertheless coincide. Using the non-compensatory approach, 43 out of 146 countries experience a change in rank greater than ten positions (none before the 30th ESI rank). When compensability among indicators is not allowed, countries with very poor performance in some indicators, such as Indonesia or Armenia, worsen their rank with respect to the linear yardstick, whereas countries that have less extreme values improve their ranking, such as Azerbaijan or Spain. Table 7.12 shows the countries with the largest variation in their ranks.

7.7 Conclusion

This chapter has presented a new mathematical aggregation convention for the solution of the so-called discrete multi-criterion problem in a SMCE context. This multi-criterion aggregation convention can be divided into two main steps:

- Pair-wise comparison of alternatives
- Ranking of alternatives in a complete pre-order

Throughout the pair-wise comparison step it is guaranteed that ordinal, crisp, stochastic and fuzzy criterion scores are tackled equivalently. The double threshold model, generating a pseudo-order structure, is used for preference modelling; as a consequence the so-called Luce paradox is avoided. To deal with the lack of stability of the pseudo-order structure, valued preference relations modelled by means of fuzzy preference relations are introduced. Given the requirement of consistency between indifference and preference thresholds, the functional form of these fuzzy relations looks descriptively reasonable. Weights are never combined with intensities of preference; as a consequence the theoretical guarantee that they are importance coefficients holds. Given that the sum of weights is equal to one, the pair-wise comparisons can be synthesized in a matrix, which can be interpreted as a voting matrix. Thanks to Proposition 7.1, it is known that ties are possible but that the probability of coming across one is approximately zero; neutrality is then in general respected.

The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule. The Condorcet tradition has been chosen for four main reasons:

- Non-compensability is implied, since intensities of preference are never used.
- Manipulation rules of weights guarantee that they are importance coefficients.
- It is the most consistent approach for generating a complete ranking.
- There is a low probability of obtaining rank reversals.

A problem connected with the use of Condorcet consistent rules is the occurrence of cycles. A cycle-breaking rule normally demands some arbitrary choices, such as eliminating the cycle with the lowest support, and so on. In search of a non-arbitrary cycle breaking rule the Condorcet–Kemeny–Young–Levenglick ranking procedure was chosen; no arbitrary choice is called for with this procedure. Given the fact that criterion weights are used, anonymity is necessarily lost. However, given that Arrow's impossibility theorem forces us to make trade-offs between decisiveness and anonymity, the loss of anonymity in favour of decisiveness in our framework is a positive feature. An important advantage of the C–K–Y–L procedure is that its properties are completely known and meet the requirements of social multi-criteria evaluation. A problem connected with the C–K–Y–L procedure is its computational complexity. Given that this problem can be solved by the numerical algorithm presented, its implementation in a multi-criteria framework is possible without any restriction on the number of alternatives considered. Consistently with the normative tradition in political philosophy and following the prudence axiom, the minority principle is introduced by means of a veto power. A vetoed alternative, the Borda loser, is found by means of the original Borda approach, implemented through a frequency matrix. This approach has been chosen because:

- It is independent of arbitrary ad hoc thresholds.
- It considers the global opposition to the final ranking.
- No specific intensity of preference is considered, thus weights continue to be importance coefficients.

The issue that makes multi-criterion aggregation conventions intrinsically complex, is the fact they are simultaneously *formal, descriptive* and *normative* models (Munda, 1993). As a consequence, the properties of an approach have to be evaluated at least in the light of these three dimensions. In the framework of the debate on the maximization assumption in microeconomics, Musgrave (1981) made a very useful classification of the assumptions made in economic theory. He makes a distinction between *negligibility assumptions, domain assumptions and heuristic assumptions*. The first type is required to simplify and focus on the essence of the phenomena studied. The second type of assumption is needed when applying a theory to specify the domain of applicability. The third type is needed either when a theory cannot be directly tested or when the essential assumptions give rise to such a complex model that successive approximation is required.

Let us then try to clarify the properties of the approach I am proposing in the light of these considerations.

Descriptive domain assumptions:

- Mixed information is tackled in the form of ordinal, crisp, stochastic and fuzzy criterion scores.
- The preference model is a pseudo-order structure with constant indifference and preference thresholds.
- The most useful result for policy-making is a complete ranking of alternatives.

Normative domain assumptions:

- Simplicity is desirable and means the use of as few ad hoc parameters as possible.
- Weights are meaningful only as importance coefficients and not as trade-offs. As a consequence, complete compensability cannot be implemented.
- A minority principle must be implemented for ethical and prudential reasons.

Formal domain assumptions:

- Unanimity
- Monotonicity
- Neutrality
- Reinforcement

Heuristic descriptive assumptions:

- Criteria can always be derived from the higher dimensions to which they univocally belong.
- Valued preference relations (in the form of fuzzy relations) are useful for solving the problem of lack of stability of a pseudo-order structure.

Heuristic formal assumptions:

- Local stability
- Cycle-breaking without losing neutrality.
- Semantic distance as a compromise solution between generality and precision

Negligibility formal assumptions:

- Anonymity
- Independence of irrelevant alternatives

In conclusion, we may state that the main characteristic of the multi-criterion aggregation convention I am proposing is that all the steps are fully justified and all the properties made explicit. Of course this is not to imply that it is the "best" possible approach to the discrete multi-criterion problem. It is a "reasonable" approach based on theoretical and empirical grounds, all of them explicit and thus open to evaluation in relation to a particular purpose.

Appendix 7.1

Proposition 7.1 *In the outranking matrix E the event of finding ties, that is* $e_{jk} = e_{kj} = \frac{M}{2}$ *is possible but its probability is approximately zero.*

Proof

The probability that $e_{jk} = e_{kj} = \frac{M}{2}$ always depends at least on the number of criteria (voters or individual indicators) in favour of a_j and a_k. Let us also assume the existence of a set of criterion weights $W = \{w_m\}$, $m = 1, 2, \ldots, M$, with $\sum_{m=1}^{M} w_m = 1$, which is very common in multi-criteria analysis and in the construction of composite indicators. Let us then look at the specific probabilities of each single factor.

Given S criteria in favour of a_j and T criteria in favour of a_k, the probability of a specific combination is $\frac{1}{\binom{M}{S}}$.

A specific value on the vector of weights depends on how many numbers we are using for the definition of each weight. Let us assume that each weight is defined by two integer numbers (e.g. 0.02, 0.10, 0.25,….,..), then the probability of a specific weight is $\frac{1}{10^2}$. Since the value 0.00 does not make sense, the probability is $\frac{1}{99}$. Thus a specific vector has a probability of $\left(\frac{1}{99}\right)^M$.

At this point it is possible to compute the probability $\pi(v)$ of finding a tie on the outranking matrix E. The probability is:

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M$$

It is evident that $\pi_v \simeq 0$.

Let us now assume that no criterion weight is used. In this case, to have a tie, it is necessary that the number of criteria M be even. The probability is

$\pi_v = \dfrac{1}{\binom{M}{M/2}} \times \left(\dfrac{1}{2}\right)$, where $\dfrac{1}{2}$ is the probability of having an even number of criteria (on the grounds of Laplace insufficient reason principle). In this case too, the probability is very low.[7]

One should note finally that the probability to get ties is always close to zero in the case where indifference and preference thresholds are used in the preference modelling.

In fact, given the vector of intensities I_m, of criteria g_m, the values $\mu(P)_m$ and $\mu(I)_m$ depend on the thresholds q_m and p_m defined on each criterion g_m. Let us denote with $\pi(q_m)$ and $\pi(p_m)$ the respective probabilities of getting a precise value of q_m and p_m on a criterion g_m. Then the probability of getting a specific vector of values on q_m and p_m is $\bigcap_{m=1}^{M} \pi\,(q)_m$ and $\bigcap_{m=1}^{M} \pi\,(p)_m$. The question now concerns the values of $\pi(q_m)$ and $\pi(p_m)$? In theory, the thresholds may vary a priori on any point of the intensity of preference I_m, I_m being a set which is not finite or countable. Thus, it is[8]

$$\pi(q_m) = \pi(p_m) = \lim_{n \to +\infty} \frac{1}{n} = 0$$

At this point it is possible to compute the probability $\pi(v)$ of having a tie in the matrix E in the most general case, i.e. where thresholds and weights are defined on all criteria. It is

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M \times \pi(q)_m{}^M \times \pi(p)_m{}^M$$

Let us put $\pi(q)_m = \pi(p)_m = \lambda$, then

[7] In the case with the smallest M, which makes sense with a ranking problem, i.e. four criteria, the probability is $\pi_v = 1/2 \simeq 0.083$; for $M = 6$, it is: $\pi_v = 1/40 \simeq 0.025$.

[8] One could argue that, from a descriptive point of view, it is not very realistic to assume that all thresholds have the same probability along the set I_m. Let us make the VERY optimistic assumption that only ten thresholds of each type are possible for each criterion and four criteria exist. In this case the probability of a specific vector on q_m and p_m is $\left(\frac{1}{10}\right)^8 = 0.00000001$. As one can see, even in this optimistic case the probability is very close to zero.

$$\pi_v = \frac{1}{\binom{M}{S}} \times \left(\frac{1}{99}\right)^M \times \lambda^{2M}$$

It is evident that $\pi_v \simeq 0$

Appendix 7.2

It is a trivial matter to prove that the semantic distance satisfies the properties of non-negativity and symmetry: the fulfilment of the property of triangle inequality can be proven as follows (Munda, 1995). Let us assume three functions:

$$f(x): X \to R^+,\ g(y): Y \to R^+,\ h(z): Z \to R^+$$

Let's also assume that $X \cap Y \cap Z \neq \varnothing$
We first prove that $\forall\ x \in X, \forall\ y \in Y$ *and* $\forall\ z \in Z$; the relationship $|x-y|+|y-z| \geq |x-z|$ is always true.

The total number of possible cases is *3!*

$$\begin{aligned}
&x \geq y \geq z \to (x-y)+(y-z)-(x-z)=0\\
&x \geq z \geq y \to (x-y)+(-y+z)-(x-z)= 2(z-y) \geq 0\\
&y \geq x \geq z \to (-x+y)+(y-z)-(x-z)= 2(y-x) \geq 0\\
&y \geq z \geq x \to (-x+y)+(y-z)-(-x+z)= 2(y-z) \geq 0\\
&z \geq x \geq y \to (x-y)+(-y+z)-(-x+z)= 2(x-y) \geq 0\\
&z \geq y \geq z \to (-x+y)+(-y+z)-(-x+z)=0
\end{aligned}$$

therefore

$$|x-y|+|y-z| \geq |x-z| \geq 0 \quad \forall\, x \in X,\ \forall\, y \in Y \text{ and } \forall\, z \in Z.$$

Since $f(x) \geq 0,\ g(y) \geq 0$ and $h(z) \geq 0$, it is:

$$\int_x \int_y \int_z \left[|x-y|+|y-z|-|x-z|\right] f(x)\, g(y)\, h(z)\, dz\, dy\, dx \geq 0$$

This integral can be decomposed as follows:

$$\begin{aligned}
&\int_x \int_y \int_z |x-y|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx + \int_x \int_y \int_z |y-z|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx -\\
&\int_x \int_y \int_z |x-z|\, f(x)\, g(y)\, h(z)\, dz\, dy\, dx = \int_x \int_y |x-y| f(x)\, g(y)\, dy\, dx +\\
&+\int_y \int_z |y-z|\, g(y)\, h(z)\, dz dy - \int_x \int_z |x-z| f(x)\, h(z)\, dz\, dx
\end{aligned}$$

This is because the triple integrals can be computed by means of iterated integrals and because it is: $\int_x f(x)\,dx = \int_y g(y)\,dy = \int_z h(z)\,dz = 1$

Therefore, it is:

$$S_d\left[f(x),g(y)\right] + S_d\left[g(y),h(z)\right] - S_d\left[f(x),h(z)\right] \geq 0 \qquad \text{or}$$
$$S_d\left[f(x),g(y)\right] + S_d\left[g(y),h(z)\right] \geq S_d\left[f(x),h(z)\right]$$

When both variables x and y are defined in the same interval, i.e. $X = Y = [x_L, x_U] = [y_L, y_U]$, for reasons of consistency it is necessary to prove that the value of the semantic distance is other than zero. For simplicity let us say $x_L = y_L = a$ and $x_U = y_U = b$ where $a < b$. Let us now consider the following algebraic expression: $|x-y|\,f(x)\,g(y)$. This product assumes the value of zero if at least one of its three elements is zero. Concerning *f(x)* and *g(y)*, the following two assumptions are made:

$$(1)\begin{cases} f(x) > 0 & \forall\, x \in (a,b) \\ f(x) = 0 & \text{if } x = a \ or\ x = b \end{cases}$$

$$(2)\begin{cases} g(y) > 0 & \forall\, y \in (a,b) \\ g(y) = 0 & \text{if } y = a \text{ or } y = b \end{cases}$$

For the factor $|x-y|$ it is:

$$(3)\begin{cases} |x-y| > 0 & \forall\, x \neq y \in \left[a,b\right] \\ |x-y| = 0 & \text{if } x = y \end{cases}$$

On the basis of these assumptions it is possible to conclude that

$$\begin{cases} |x-y|\,f(x)\,g(y) > 0 & \forall\, x \neq y \in (a,b) \\ |x-y|\,f(x)\,g(y) = 0 & \text{elsewhere} \end{cases}$$

If we take into consideration the sum of all

$$|x-y|\,f(x)\,g(y) \qquad \forall\, x \in \left[a,b\right] \text{ and } \forall\, y \in \left[a,b\right], \quad i.e.$$
$$\int_x\int_y |x-y|\,f(x)\,g(y)\,dy\,dx, \quad \text{given that it is not } |x-y|\,f(x)\,g(y) = 0,$$
$$\forall\, x \in \left[a,b\right] \text{ and } \forall\, y \in \left[a,b\right], \quad \text{it is}$$
$$\int_x\int_y |x-y|\,f(x)\,g(y)\,dy\,dx > 0$$

Note that

$$S_d(f(x),g(y)) = 0 \ \ iff\ \ x = y\ \forall\, x \in \left[a,b\right] \ and\ \ \forall\, y \in \left[a,b\right],$$
$$i.e.\ iff\ a = x = y = b.$$

This is true only if x and y are two equal, crisp real numbers.

Finally, one should note that if the distance between a fuzzy number and itself is computed by definition, the condition S_d ($f(x)$, $f(x)$) = 0 has to be imposed.

Appendix 7.3

To make the semantic distance presented in Sect. 7.2 operational, a Monte Carlo type numerical algorithm is required (Munda, 1995).

The initial *assumptions* are:

(1)
$$f(x): X = [x_L, x_U] \to \mathrm{M}$$
$$g(y): Y = [x_{L'}, x_{U'}] \to \mathrm{M}$$

where M is the membership space.

(2) All $x \in X$ and all $y \in Y$ can be obtained by means of a random generator that supplies uniformly distributed numbers $r \in [0,1]$.

We have:

$$x = rx_L + (1-r)x_U \quad \text{and}$$
$$y = rx_{L'} + (1-r)x_{U'}$$

(3) The probability of obtaining a point p inside e.g. $f(x)$, whose value on the x-axis is x_0 depends on the shape of the function. An auxiliary variable Z, with $z \in [0, \max f(x)]$, is then introduced by means of a random generator.

Now the *procedure* is as follows:

STEP 1: draw a random number r_0
STEP 2: $x_0 = r_0 x_L + (1 - r_0)\, x_U$
STEP 3: draw a random number Z_0
STEP 4: if $Z_0 \le f(x_0)$ then go to next step
if $Z_0 > f(x_0)$ then return to step 1.
STEP 5: draw a random number r_1
STEP 6: $y_1 = r_1 x_{L'} + (1 - r_1) x_{U'}$
STEP 7: draw a random number Z_1
STEP 8: if $Z_1 \le g(y_1)$ then compute $|x_0 - y_1|$
if $Z_1 > g(y_1)$ then return to step 5.

By repeating this procedure N times, N values of $|x_i - y_i|$ are obtained. The semantic distance between two fuzzy sets is approximately equal to the arithmetic mean of all the points bounded by their respective membership functions obtained by drawing random numbers. In more formal terms it is:

$$S_d\big(f(x),\ g(y)\big) = \iint_{x,\,y} |x - y| f(x) g(y) dy dx \simeq \frac{\sum_{i=1}^{N} |x_i - y_i|}{N}$$

with $i = 1, 2, \ldots, N$.

Of course, the greater N, the more precise the computation.

Chapter 8
Searching for the "Social Compromise Solution": A Conflict Analysis Procedure for Illuminating Distributional Issues

8.1 Introduction

As discussed in Chap. 1, one of the most interesting research directions in modern public economic policy is the explicit attempt to take political constraints, interest groups and collusion effects into account. The issue of "distributional coalitions" has been considered of key importance in determining growth factors (Olson, 1982). In the framework of sustainability policies, the need to deal with conflicts among various social actors is even more unavoidable.

Ecosystems are used in several ways at the same time by a number of different users. Such situations almost always lead to conflicts of interest and damage to the environment. Any social decision problem is characterized by conflict between competing values and interests and the different groups and communities that represent them. In sustainability policies, biodiversity goals, landscape objectives, the direct functioning of different environments as resources, the historical and cultural meanings that places have for communities, the recreational options environments provide, etc., are a constant source of conflict. Any policy option always implies winners and losers, thus it is imperative to check if a policy option seems preferable just because some dimensions (e.g. the environmental) or some social actors (e.g. the lower income groups) have not been taken into account. This is what was defined in Chap. 2 as the *social incommensurability* issue.

In operational terms, one of the classical evaluation tools is cost–benefit analysis. It is generally considered that cost–benefit analysis focuses on *efficiency* criteria; but, any policy decision affects the welfare of individuals, regions or groups in different ways; consequently, public support for any policy decision very much depends on the *distributional effects* of such a decision. Some revisions of cost–benefit analysis try to include distribution issues directly in the analysis (see e.g. Helmers, 1979). However, all these revisions may sometimes present such theoretical and operational difficulties (see Box 4.2) that it is rather tempting to ignore distributional aspects without further comment. This attitude is rarely defended theoretically, but unfortunately often practiced (Bojö et al., 1990).

A well-known approach for dealing with distributional issues in land-use planning is the so-called *planning balance sheet method* which can be considered an extension of conventional cost–benefit analysis (Lichfield, 1964, 1988, 1993). This approach aims to provide a broader framework for the assessment of the gains and losses of a plan by constructing detailed socio-economic accounts of all project effects and by taking into account the different groups in society which are affected in their well-being by the plan. A weakness of this method is that it is primarily meant to present in a systematic way a description of all the distributive impacts, but no elaboration with *normative* purposes is generally made.

This chapter presents a possible way of overcoming this drawback of the planning balance sheet method; introducing concepts coming mainly from fuzzy set theory and social choice. Most of the results presented here proceed from the empirical experience of various real-world applications of the NAIADE conflict analysis procedure (Munda, 1995) over a decade. First a fuzzy coalition formation algorithm will be developed, followed by the introduction of a ranking procedure and a veto index founded on the minority principle.

8.2 Do Similarities Exist Among Social Actors? A Fuzzy Cluster Analysis

As in the planning balance sheet method, the proposed approach requires as a first step the construction of a social impact matrix showing the various policy options and their impact on the social actors. From an empirical point of view, the construction of this matrix requires sophisticated field work based mainly on participative techniques (see e.g. Kasemir, 2003). However, the results obtained with such techniques are qualitative in nature and often presented in an unstructured manner. As a consequence, formal techniques helping the structuring, synthesis and further elaboration of this information are operationally very useful (Funtowicz et al., 1999; Munda, 2004). The following main assumptions are made:

(1) Only a set of well defined policy options has to be taken into account.
(2) The impact of these policy options on different social actors are evaluated by means of "linguistic variables" (good, not very good, etc.).
(3) The semantic distance between any pair of social actors is used as a conflict indicator.
(4) A fuzzy cluster algorithm is used to synthesize similarities/diversities among social actors.

In more formal terms, the problem faced can be described in the following way:

A is a finite set of N feasible policy options; B is the set of different social actors,

$B = \{b_p\}$ $p = 1, 2, \ldots, P$ considered relevant in a policy problem,

$\Lambda = \{\lambda_p\}$, $p = 1, 2, \ldots, P$, with $\sum_{p=1}^{P} \lambda_p = 1$ being the vector of weights attached to the set of the P social actors, indicating their relative importance, i.e. given two non-equal numbers to construct a vector in R^2, then the greatest number must be placed in the

Table 8.1 Example of a social impact matrix

	Policy options			
Social Actors	a_1	a_2	a_3	a_4
b_1	$b_1(a_1)$	$b_1(a_2)$	·	$b_1(a_4)$
b_2	.	.	.	.
b_3	.	.	.	.
b_4	.	.	.	.
b_5	.	.	.	.
b_6	$b_6(a_1)$	$b_6(a_2)$	·	$b_6(a_4)$

position corresponding to the most important social actor (Podinovskii, 1994). In this framework, the policy option a_1 is evaluated to be better than a policy option a_2 (both belonging to the set A) according to the *p-th* point of view if $b_p(a_1) > b_p(a_2)$, where the social actor scores $b_p(.)$ are measured on a linguistic variable scale of measurement.

As we have seen in Chap. 4, fuzzy set theory provides a framework for representing "*qualitative information*" by means of the concept of "*linguistic variable*". Human judgements, especially in linguistic form, appear to be plausible and natural representations of cognitive observations. We can explain this phenomenon through *cognitive distance*. A linguistic representation of an observation may require a less complicated transformation than a numerical representation, and therefore less distortion may be introduced by the former than by the latter.

The formal problem we are dealing with can be summarized in a table form, called a *social impact matrix*, as shown in Table 8.1 (where $P = 6$ and $N = 4$).

As discussed in Chap. 7, the semantic distance can be used to compare fuzzy sets in general and linguistic variables in particular. In short if $\mu_1(x)$ and $\mu_2(x)$ are two linguistic variables, one can write:

$$f(x) = k_1\mu_1(x) \; and \; g(y) = k_2\mu_2(x) \tag{8.1}$$

where $f(x)$ and $g(y)$ are two functions obtained by rescaling the ordinates of $\mu_1(x)$ and $\mu_2(x)$ through k_1 and k_2, such that

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} g(y)dy = 1 \tag{8.2}$$

The distance between all points of the membership functions is computed as follows:

$$S_d\big(f(x),\ g(y)\big) = \iint_{x,\,y} |x - y| f(x)g(y)dydx \tag{8.3}$$

In the problem at hand, between any pair of social actors b_i, b_j with $i \neq j$, their relative distance can be computed by considering the set of N linguistic evaluations given to the set of policy options (i.e. the rows i and j in the matrix). In more formal terms, if $\overline{x}$ is the vector of the linguistic evaluations of b_i and $\overline{y}$ is that of b_j, both belonging to R^N, the generalization of the Minkowski metric described in (8.4) can be applied:

$$d(b_i, b_j) = \left\| \bar{x} - \bar{y} \right\| = \sum_{i=1}^{N} \left[\left(\iint_{x,\,y} \left| x - y \right| f_i(x) g_i(y) dy dx \right)^{\beta} \right]^{\frac{1}{\beta}} \tag{8.4}$$

For

$\beta = 1$ an absolute value metric (completely compensatory)
$\beta = 2$ a Euclidean metric (partially compensatory)
$\beta \to \infty$ the Tchebycheff metric (completely non-compensatory) can be obtained.

By using the distance described in (8.4) as a conflict indicator, a similarity matrix (achieved by means of the simple transformation $s_{ij} = \frac{1}{1+d_{ij}}$) for all possible pairs of groups can be obtained, so that a clustering procedure is meaningful.

On an axiomatic basis, cluster analysis can be differentiated into deterministic, stochastic and fuzzy. By taking into consideration the "clustering criteria", the following distinction exists (Anderberg, 1973: Hartigan, 1975; Bezdek, 1980):

- Hierarchical methods
- Graph theoretic methods
- Objective functional methods

The hierarchical clustering approach, in particular, allows an evolutionary view of the aggregation process and can easily be dealt with within fuzzy terms. However, in a fuzzy environment a problem exists, i.e. the relation between the concepts of *partition* and *equivalence class*.

In a crisp environment, the choice of treatment of data in terms of partitions or equivalence relations is a matter of convenience, since the two models are fully equivalent (philosophically and mathematically). On the contrary, fuzzy equivalence relations and partitions are philosophically similar, but their mathematical structures are not isomorphic (e.g. the notion of transitivity is unique for crisp relations but may take any of several forms in the fuzzy case).

We begin the discussion of fuzzy cluster analysis with the definition of a crisp equivalence relation. Let $B=\{b_1, b_2, \ldots, b_p\}$ be the finite set of social actors. Then a $P \times P$ matrix $S = [s_{ij}] = [s(b_i, b_j)]$ is a *crisp equivalence relation* for $B \times B$ if

$$s_{ij} = 1 \qquad 1 \le i \le P \qquad \text{(reflexivity)}$$

$$s_{ij} = s_{ji} \qquad 1 \le i \ne j \le P \qquad \text{(symmetry)}$$

$$\begin{cases} s_{ij} = 1 \\ \\ s_{jk} = 1 \end{cases} \Rightarrow s_{ik} = 1 \quad \forall i, j, k \qquad \text{(transitivity)}$$

Let S be a fuzzy binary relation with $\mu_s(b_i, b_j)$ indicating the degree to which two social actors b_i and b_j are similar (similarity matrix). The relation S is obviously reflexive and symmetrical, thus it is called a *resemblance relation*.

A fuzzy relation is a *similitude relation* if it has the following properties:

$$\mu_S(b_i,b_i)=1 \qquad \forall (b_i,b_i)\in B\times B \qquad \text{(reflexivity)}$$

$$\mu_S(b_i,b_j)=\mu_S(b_j,b_i) \qquad \forall (b_i,b_j)\in B\times B \qquad \text{(symmetry)}$$

$$\mu_S(b_i,b_k)\geq \max\min\left[\mu_S(b_i,b_j),\mu_S(b_j,b_k)\right]$$
$$\forall (b_i,b_j),(b_j,b_k),(b_i,b_k)\in B\times B \qquad \text{(max-min transitivity)}$$

Note that compared to the notion of transitivity in conventional analysis, the present concept defines a weaker transitivity of similarity.

If one wants to derive a set of *equivalence classes* (and not simple partitions) it is necessary for the similarity matrix to be at least max–min transitive. As is well-known (Leung, 1988), an intransitive similarity matrix can be transformed into transitive by deriving the *transitive closure* $\hat{S}$ of S. The *max–min transitive closure* of a fuzzy binary relation S is

$$\hat{S}=S\cup S^2\cup S^3\cup\ldots \qquad (8.5)$$

where $S^2 = S \circ S$ is the *max–min composition of S* (more technical details can be found in Appendix 8.1).

Knowing that any fuzzy set $\tilde{A}$ can always be decomposed into a series of α-level sets $\tilde{A}_\alpha$, the similitude relation $\hat{S}$ can be decomposed into

$$\hat{S}=\bigcup_{\alpha\in[0,1]}\alpha\hat{S}_\alpha \qquad (8.6)$$

Since $\hat{S}_\alpha$ is reflexive, symmetrical and transitive in the sense of ordinary sets, it is an equivalence class of level α. Within each *α-level equivalence class*, the similarity of any two social actors is no less than α.

Note that the equivalence classes obtained are ordinary disjoint sets. In fact, in order to have non-mutually exclusive equivalence classes, it is necessary to assume the use of a *min-addition transitive similarity matrix* (which is a stronger assumption than max–min transitivity). Consider the social impact matrix described in Table 8.2.

By applying the semantic distance described in (8.4) with β= 2, after the transformation $s_{ij}=\dfrac{1}{1+d_{ij}}$, the similarity matrix for all possible pairs of social actors shown in Table 8.3 is obtained:

This means, for example, that the greatest similarity is found between social actors b_1 and b_2, and between b_4 and b_5. These social actors have a relatively high correspondence of goals, accordingly. The reverse holds true for social actors b_2 and b_4, between which the lowest degree of similarity is found.

By using the notion of max–min composition, the following new fuzzy relations are derived:

S^2

	b_1	b_2	b_3	b_4	b_5	b_6
b_1	1	0.729	0.426	0.426	0.426	0.426
b_2	0.729	1	0.426	0.410	0.410	0.410
b_3	0.426	0.426	1	0.675	0.675	0.672
b_4	0.426	0.410	0.675	1	0.729	0.672
b_5	0.426	0.410	0.675	0.729	1	0.672
b_6	0.426	0.410	0.672	0.672	0.672	1

S^3

	b_1	b_2	b_3	b_4	b_5	b_6
b_1	1	0.729	0.426	0.426	0.426	0.426
b_2	0.729	1	0.426	0.426	0.426	0.426
b_3	0.426	0.426	1	0.675	0.675	0.672
b_4	0.426	0.426	0.675	1	0.729	0.672
b_5	0.426	0.426	0.675	0.729	1	0.672
b_6	0.426	0.426	0.672	0.672	0.672	1

S^4

	b_1	b_2	b_3	b_4	b_5	b_6
b_1	1	0.729	0.426	0.426	0.426	0.426
b_2	0.729	1	0.426	0.426	0.426	0.426
b_3	0.426	0.426	1	0.675	0.675	0.672
b_4	0.426	0.426	0.675	1	0.729	0.672
b_5	0.426	0.426	0.675	0.729	1	0.672
b_6	0.426	0.426	0.672	0.672	0.672	1

Table 8.2 Illustrative example of a social impact matrix

	Policy options						
Social Actors	**a1**	**a2**	**a3**	**a4**	**a5**	**a6**	**a7**
b_1	Very good	good	Moderate	bad	fairly good	fairly bad	very bad
b_2	Very good	good	Moderate	bad	fairly good	very bad	Very bad
b_3	Very bad	fairly bad	Moderate	good	very good	good	moderate
b_4	very bad	fairly bad	fairly bad	good	fairly good	good	very good
b_5	Very bad	bad	fairly bad	moderate	fairly good	good	very good
b_6	Very bad	good	Bad	good	good	good	very good

Table 8.3 Similarity matrix between the social actors of the illustrative example

	b_1	b_2	b_3	b_4	b_5	b_6
b_1	1	0.729	0.426	0.399	0.403	0.403
b_2	0.729	1	0.410	0.386	0.390	0.390
b_3	0.426	0.410	1	0.675	0.584	0.569
b_4	0.399	0.386	0.675	1	0.729	0.672
b_5	0.403	0.390	0.584	0.729	1	0.595
b_6	0.403	0.390	0.569	0.672	0.595	1

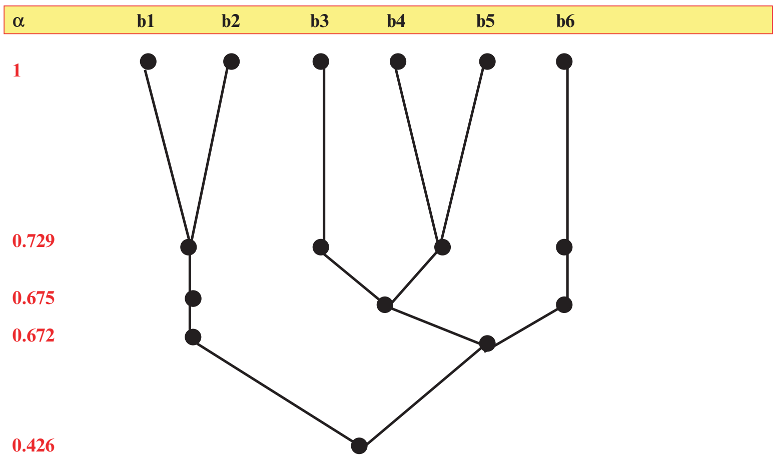


Fig. 8.1 Dendrogram of the cluster formation process

Since in the series of max–min compositions $S^3 = S^4$, the transitive closure is

$$\hat{S} = S \cup S^2 \cup S^3 = S^3 \tag{8.7}$$

Since $\hat{S}$ is a similitude relation, it can be decomposed into equivalence classes with respect to the degree of similarity α.

Thus the application of the clustering procedure leads to the following results (see Fig. 8.1). As long as the similarity degree α required for convergence is higher than 0.729, there will be no cluster formation. Two groups will be formed when α is between 0.729 and 0.675 (b_1 and b_2), and (b_4 and b_5). When the similarity degree is reduced to 0.675 and 0.672, social actors b_3 and b_6 join the last group respectively. The conflict of interest between the remaining groups (b_1, b_2) versus (b_3, b_4, b_5, b_6) is considerable, as can be inferred from the low degree of similarity associated with a grand coalition.

It can be proved that the following four algorithms generate the same partition (Miyamoto, 1990):

- The single linkage method
- The connected components of an undirected fuzzy graph
- The transitive closure of a reflexive and symmetric fuzzy relation
- The maximal spanning tree of a weighted graph

Thus the following conclusions can be drawn:

1. Since the connected components are independent of the numbering of the vertices, the algorithm is independent of the ordering of the inputs, and is therefore *stable*.
2. *No reversal* exists in the dendrogram ("reversal" meaning that the merging levels are not monotonically decreasing, and thus a cut of the dendrogram might produce ambiguous results).
3. One is not obliged to use only the Euclidean metric (e.g. as in the "*centre of gravity*" procedures), any distance measure (even if it does not respect the triangular inequality property) can be used, thus the method is *general*.

In real-world applications, when the actors involved in a policy process look at dendrograms, they can generally make little sense of them. Clearly, further elaboration is then needed. In particular, information on rankings of policy options according to each cluster of social actors seems very desirable.

8.3 Ranking Policy Options

As discussed in Chaps. 6 and 7, the maximum likelihood ranking of policy options is that ranking supported by the maximum number of social actors for each pair-wise comparison, summed over all pairs of policy options. More formally, the C–K–Y–L ranking procedure, adapted to the problem at hand, can be described as follows.

For each α-level equivalence class, let $C_\alpha=\{c_1, c_2,\ldots,c_z\}$ be the finite set of possible groups of social actors, with $|c_1|\cup|c_2|\cup\ldots\cup|c_z| = P$. Then, $\forall c_i \in C_\alpha$, with $i=1, 2,\ldots,Z$, a pair-wise comparison of the N policy options needs to be carried out.

For carrying out such a pair-wise comparison the following axiomatic system is required (adapted from Arrow and Raynaud, 1986, pp. 81–82).

Axiom 1: *Diversity*. Each social actor b_p defines a total order on the finite set A of policy options to be ranked.

Axiom 2: *Symmetry*. The only preference information social actors provide is the ordinal pair-wise preferences.

Axiom 3: *Positive Responsiveness*. The degree of preference between two policy options $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ is a strictly increasing function of the number $|c_i|$ and weights λ_p, of the social actors who rank $\boldsymbol{a}_1$ before $\boldsymbol{a}_2$.

Clearly all three axioms are fulfilled by giving an ordinal meaning to the linguistic variables contained in the social impact matrix (i.e. no intensity of preference is used). Thanks to these three axioms an $N\times N$ *outranking matrix* E can be built. Any generic element of E: $e_{jk}, j\neq k$ is the result of the pair-wise comparison, according to all the $|c_i|$ social actors, between policy options *j* and *k*. Such a global pair-wise comparison is obtained by means of (8.8).

$$e_{jk} = \sum_{p=1}^{|c_i|} \left(\lambda_p(P_{jk}) + \frac{1}{2}\lambda_p(I_{jk}) \right) \tag{8.8}$$

where $\lambda_p\,(P_{jk})$ and $\lambda_p\,(I_{jk})$ are the weights of the social actors expressing a preference and an indifference relation respectively. All the $N(N-1)$ pair-wise comparisons of policy options N compose the matrix E.

Let us call T the set of all the *N!* possible complete rankings, of policy options, $T = \{\tau_s\}$, $s = 1, 2, \ldots, N!$. For each τ_s, we compute the corresponding score φ_S as the sum of e_{jk} over all the $\binom{N}{2}$ pairs *jk* of policy options, i.e.

$$\varphi_S = \sum e_{jk} \tag{8.9}$$

where $j \neq k$, $s = 1, 2, \ldots N!$ and $e_{jk} \in \tau_s$

The final ranking (τ_*) is the one[1] which maximises(8.9):

$$\tau_* \Leftrightarrow \varphi_* = \max \sum e_{jk} \qquad \text{where } e_{jk} \in \mathrm{T}. \tag{8.10}$$

As we know from Chap. 6, other properties of the C–K–Y–L ranking procedure are as follows.

- *Neutrality:* it does not depend on the name of any policy option, all policy options are treated equally.
- *Unanimity* (sometimes called *Pareto Optimality*): if all social actors prefer policy option $\boldsymbol{a_1}$ to policy option $\boldsymbol{a_2}$ than $\boldsymbol{a_2}$ should not be chosen.
- *Monotonicity:* if policy option $\boldsymbol{a_1}$ is chosen in any pair-wise comparison and only the social actors' linguistic evaluations of $\boldsymbol{a_1}$ are improved, then $\boldsymbol{a_1}$ should still be the winning policy option.
- *Reinforcement:* if the set A of policy options is ranked by two subsets B_1 and B_2 of the social actors set B, such that the ranking is the same for both B_1 and B_2, then $B_1 \cup B_2 = B$ should still produce the same ranking.

At this point, we refer to the *normative tradition* in political philosophy, which has also an influence in modern social choice (Moulin, 1981) and public policy (Mueller, 1978). The basic idea is that any coalition controlling more than 50% of the votes may be converted in an actual dictator. As a consequence, the "remedy to the tyranny of the majority is the minority principle, requiring that all coalitions, however small, should be given some fraction of the decision power. One measure of this power is the ability to veto certain subsets of outcomes.…"

[1] It is important to remember that sometimes the final ranking is not unique. This is a desirable property since it can be considered a measure of the *robustness* of the results provided.

(Moulin, 1988, p. 272). As discussed in Chap. 7, the introduction of a veto power can be further justified in the light of the so-called "*prudence*" axiom (Arrow and Raynaud, 1986, p. 95), whose main idea is that it is not prudent to accept alternatives whose degree of conflictuality is too high (and thus the decision taken might be very *vulnerable*).

Note that allocating veto power across the various groups of social actors has profound ethical implications, since it entails attaching different weights to different groups. Moreover, if too much veto power is allowed, cooperatively stable solutions may disappear; on the other hand, if too little veto power is allowed, stable solutions are too numerous. This problem has a unique *mathematical* solution attributable to Moulin (1981). The philosophy behind Moulin's theorem is that any group with $x\%$ of social actors must be able to veto any subset containing less than $x\%$ of policy options.

Formally, Moulin's theorem can be adapted to our problem as follows. Given P social actors, N policy options and $C_\alpha=\{c_1, c_2,\ldots,c_z\}$ possible groups of social actors, $|c_1|\cup|c_2|\cup\ldots\cup|c_z| = P$, $\forall\ c_i \in C_\alpha$, with $i = 1,2,\ldots,Z$, the corresponding *proportional veto function* is defined in (8.11):

$$V_{P,N}(c_i) = \left(N \bullet \frac{|c_i|}{P}\right) - 1 \tag{8.11}$$

where (x) is the smallest integer bounded below by x, with

$$x = \left(N \bullet \frac{|c_i|}{P}\right).$$

In the case that weights are attached to the social actors, the proportional veto function is presented in (8.12).

$$V_{P,N}(c_i) = \left(N \bullet \lambda_i\right) - 1 \tag{8.12}$$

$$\text{where } \lambda_i = \sum_{p \in c_i} \lambda_p \tag{8.13}$$

Let us continue with the example described in Table 8.2, by applying the ranking method and the veto function just discussed. Let us choose the equivalence class obtained with α=0.672. For $C_{0.672}$, groups c_1 with (b_1, b_2) and c_2 with (b_3, b_4, b_5, b_6) exist. By applying the computations described in Equations from (8.8) to (8.10), with the assumption of equal weighting of social actors, the following rankings are obtained.

For c_1 the permutation with the highest score is unique:

$$a_1 \rightarrow a_2 \rightarrow a_5 \rightarrow a_3 \rightarrow a_6 \rightarrow a_4 \rightarrow a_7$$

For c_2 the ranking is also unique:

$$a_7 \to a_6 \to a_5 \to a_4 \to a_3 \to a_2 \to a_1$$

The application of Moulin's proportional veto function produces the following results:

$V_{6,7}(c_1) = \left(7 \bullet \frac{2}{6}\right) - 1 \cong 1$, only a_7 can be vetoed.

$V_{6,7}(c_2) = \left(7 \bullet \frac{4}{6}\right) - 1 \cong 3$, options a_1, a_2 and a_3 can be vetoed.

From analysing these results it is clear that social compromise solutions could be options a_6 and a_5. Any other choice would imply "*strong value judgements*" such as attaching an enormous weight to group c_1, which would be the only way to defend options a_1 or a_2; while choosing a_7 would imply a complete "dictatorship" of the majority.

It is important to highlight that I do not maintain that a policy-maker should not be free to use these "strong value judgements". What I want to emphasize here is that, when she/he uses them, this fact should be transparent and responsibility of doing so clearly assumed. As discussed in the tradition of public choice (Buchanan and Musgrave, 1999) not necessarily a public policy-maker is always benevolent; this is why I stated that the objective of the proposed approach is to illuminate distributional issues and corresponding ethical (or un-ethical) positions. This call for transparency in modern public economics is widely shared by various contemporary authors (see e.g. Stiglitz, 2002).

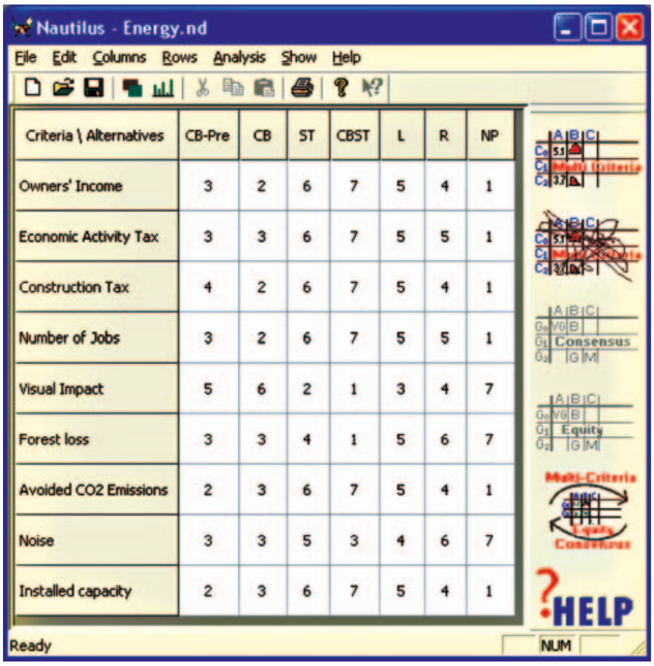
As a final example, let us consider again the Catalan wind park location problem introduced in Chap. 3. The ordinal multi-criteria evaluation matrix for this problem is described in Table 8.4 (the higher the criterion score the better the evaluation). The ordinal criterion scores are obtained by applying a positive indifference threshold q to each quantitative criterion score (for more details see Gamboa and Munda, 2007).

By considering the information contained in the impact matrix shown in Table 8.4, the outranking matrix presented in Table 8.5 is obtained. All criteria are considered under the equal weighting assumption.

By applying the C–K–Y–L ranking procedure, among the 5,040 possible rankings, the following four present the maximum score (see Table 8.6) (where the extreme left alternatives are the top ones and the extreme right alternatives are the bottom ones):

As we know, criteria and criterion scores are not determined directly by social actors. The impact matrix is a result of a technical translation operationalized by the scientific team. Even if the criteria are exactly the ones agreed with the social actors the determination of the criterion scores is independent of their preferences. For example, an interest group can accept the use of a criterion measuring the effects of the various alternatives on the employment, but the determination of the figure cannot

Table 8.4 Ordinal multi-criteria evaluation matrix for the wind park location problem



Criteria \ Alternatives	CB-Pre	CB	ST	CBST	L	R	NP
Owners' Income	3	2	6	7	5	4	1
Economic Activity Tax	3	3	6	7	5	5	1
Construction Tax	4	2	6	7	5	4	1
Number of Jobs	3	2	6	7	5	5	1
Visual Impact	5	6	2	1	3	4	7
Forest loss	3	3	4	1	5	6	7
Avoided CO2 Emissions	2	3	6	7	5	4	1
Noise	3	3	5	3	4	6	7
Installed capacity	2	3	6	7	5	4	1

Table 8.5 Outranking matrix for the wind park location problem

	CB Pre	***CB***	***ST***	***CBST***	***L***	***R***	***NP***
CB Pre	0,00	0,50	0,30	0,40	0,30	0,30	0,70
CB	0,50	0,00	0,10	0,20	0,10	0,10	0,70
ST	0,70	0,90	0,00	0,40	0,70	0,65	0,70
CBST	0,60	0,80	0,60	0,00	0,60	0,60	0,70
L	0,70	0,90	0,30	0,40	0,00	0,65	0,70
R	0,70	0,90	0,35	0,40	0,35	0,00	0,70
NP	0,30	0,30	0,30	0,30	0,30	0,30	0,00

Table 8.6 Multi-criteria maximum likelihood rankings for the wind park location problem

First	Second	Third	Fourth	Fifth	Sixth	Seventh
CBST	ST	L	R	CBPre	CB	NP
CBST	ST	L	R	CB	CBPre	NP
ST	CBST	L	R	CBPre	CB	NP
ST	CBST	L	R	CB	CBPre	NP

Table 8.7 Social impact matrix for the wind park location

Groups \ Alternatives	CB-Pre	CB	ST	CBST	L	R	NP
CAT. GOVERNM (G1)	More or Less Good	More or Less Bad	Very Good	Perfect	Good	Good	Extremely Bad
MUN. VALLBONA (G2)	More or Less Good	More or Less Bad	Very Good	Perfect	Good	Good	Extremely Bad
MUN. OMELLS(G3)	Very Bad	Good	Bad	Good	Bad	Bad	Bad
MUN. ROCALLAURA (G4)	More or Less Good	More or Less Bad	Very Good	Perfect	Good	Good	Extremely Bad
MUN. SENAN (G5)	Very Bad	Very Bad	Very Bad	Extremely Bad	More or Less Bad	Moderate	Perfect
"DEFEND LAND" (G6)	Very Bad	Very Bad	Extremely Bad	Very Bad	Very Bad	Bad	Perfect
PLAT. SENAN (G7)	Very Bad	Very Bad	Extremely Bad	Extremely Bad	More or Less Bad	Moderate	Perfect
ASS. MONTBLANQUET (G8)	Extremely Bad	Extremely Bad	Very Bad	Extremely Bad	Very Bad	More or Less Bad	Perfect
EHN (G9)	Extremely Bad	Extremely Bad	Perfect	Perfect	More or Less Good	Moderate	Extremely Bad
GERSA (G10)	Very Good	Perfect	Extremely Bad	Perfect	Extremely Bad	Extremely Bad	Extremely Bad

be (at least completely) controlled by them. This is one of the main reasons why it is desirable to combine a social impact matrix with the technical impact matrix.

As we have seen in this chapter, the first step is the construction of the *Social Impact Matrix* i.e. the evaluation every social actor gives to each option (see Table 8.7).

By applying the fuzzy clustering procedure introduced in Sect. 8.2 to the social impact matrix presented in Table 8.7 (by using the assumption of equal weighting for the various social actors), the dendrogram presented in Fig. 8.2 is obtained.

- The proximity of aims between the Municipality of Senan (G5) and the Platform per Senan (G7) are reflected in the dendogram. Also the Municipalities of Vallbona de les Monges (G2) and Rocallaura (G4) are working together in looking for their benefits.
- The Association of friends and neighbours of Montblanquet (G8) joints to the first mentioned coalition (G5 + G7) with a medium-high degree of credibility. They meet with other actors in the Coordinating committee to defend the land (G6). Most of them working in independently.
- On the other side, EHN (G9) has been negotiating with the municipalities and with the Catalonian government in order to push their project forward. This coalition (G2 + G4 + G1 + G9) has a medium degree of credibility.
- A coalition between the municipality of Els Omells de Na Gaia (G3) and Gerrsa (G10) shares a medium degree of proximity with the previous coalition. Nowadays this coalition depends more or less in the amount of money that can be received from Gerrsa as benefit tax revenue.

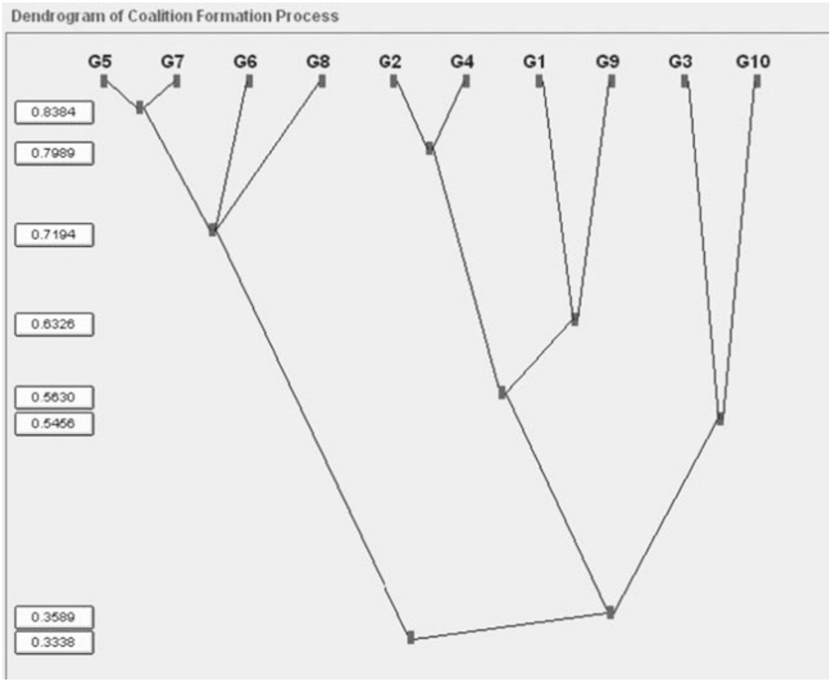


Fig. 8.2 Dendrogram of the coalition formation process for the wind park location problem

Table 8.8 Maximum likelihood rankings for coalition C_1

First	Second	Third	Fourth	Fifth	Sixth	Seventh
NP	R	L	CB Pre	CB	ST	CBST
NP	R	L	CB	CB Pre	ST	CBST
NP	R	L	CB Pre	CB	CBST	ST
NP	R	L	CB Pre	ST	CB	CBST

In real-world applications, when the actors involved in a policy process look at dendrograms generally have a question like: *and so what?* Clearly further elaborations are then needed. In particular, information on rankings of policy options according to each coalition of social actors seems very desirable. This can easily been done by applying again the C–K–Y–L ranking procedure (already used on the multi-criteria impact matrix). The coalitions obtained with the degree of credibility 0.7194 (thus a very high one) are considered.

The coalition C_1, with Municipality of Senan (G5), Platform per Senan (G7), Association of friends and neighbours of Montblanquet (G8) and Coordinating committee to defend the land (G6) present the following rankings as the maximum likelihood ones (see Table 8.8):

Table 8.9 Maximum likelihood rankings for coalition C_2

First	Second	Third	Fourth	Fifth	Sixth	Seventh
CBST	ST	L	R	CB Pre	CB	NP
CBST	ST	R	L	CB Pre	CB	NP
ST	CBST	L	R	CB Pre	CB	NP
ST	CBST	R	L	CB Pre	CB	NP
CBST	ST	L	CB Pre	R	CB	NP
CBST	ST	L	R	CB Pre	NP	CB
CBST	ST	L	R	CB	CB Pre	NP
CBST	ST	R	CB Pre	L	CB	NP
CBST	ST	R	L	CB Pre	NP	CB
CBST	ST	R	L	CB	CB Pre	NP
CBST	L	ST	R	CB Pre	CB	NP
CBST	R	ST	L	CB Pre	CB	NP

For coalition C_2, (including Municipalities of Vallbona (G2) and Rocallaura (G4)) the following rankings receive the maximum score (see Table 8.9):

Moreover by looking at the social impact matrix (Table 8.7), it is clear that for the Catalonian Government, option CBST is the best one. Anyway all the other alternatives are also more o less OK, except for NP that is considered as extremely bad. For the Municipality of Els Omells, the only acceptable alternatives are CB Pre, CB and CBST, all the others are considered bad. For EHN, alternatives ST and CBST are good options. L and R are more or less acceptable but NP is considered as extremely bad. For Gersa, alternatives CB Pre, CB and CBST are at least very good options, all the other possibilities are considered as extremely bad.

By applying Moulin's theorem the only coalition that can veto one option is C_1, which vetoes option CBST. However, it is important to remember that veto power is not a technical decision only. For instance, the alternatives as well as the social actors to be considered are defined in the problem structuring phase, which is mainly a technical, political and social process.

Concluding we can say that technically speaking, the most defensible alternatives are CBST, ST and L. From a social conflict analysis point of view, it seems that alternative CBST is the one which might generate the maximum conflict. Even if CBST seems acceptable for the majority of the social actors involved, coalition C_1 always ranks it in low positions. R has good evaluations, except by GERRSA which would be excluded in this case. L is always ranked in medium positions by all social actors. It might also be a social compromise. NP is not acceptable for most of social actors. In summary, we can state that alternatives L and R seem the only ones defensible from both technical and social points of view. All other options might maximize the social conflict or are not technically acceptable. It is interesting to note that business as usual is definitely not a desirable situation.

8.4 Concluding Remarks

In the area of environmental and resource management and in policies aiming at sustainable development, conflicting issues and interests are the normal state of affairs. Formal approaches like the one proposed here cannot resolve all conflict, but they can help to provide more insight into the nature of the conflict by providing systematic information; and to arrive at political compromises by making a complex situation more transparent to policy-makers and lay people.

In the present chapter, distribution issues have been taken into consideration by means of an eclectic approach using concepts from land-use planning, fuzzy cluster analysis and social choice. Starting with a matrix showing the impact of different courses of action on each social actor, a fuzzy clustering procedure indicating the groups whose interests are closer is used. This is more or less in agreement with the hypotheses underlying the "*minimal range theory*" in coalition formation literature. Rankings for each "credible" group of social actors are obtained by means of the majority principle implemented using a Condorcet voting principle. The issue of cycles has been tackled thoroughly. The minority principle has also been considered by means of Moulin's proportional veto function.

The approach proposed aims to be a *normative* model based on a set of *formal* properties with some *descriptive* meaning. As a consequence, the properties of this approach have to be evaluated at least in the light of these three dimensions (descriptive, normative and formal). By adding Musgrave's distinction between *negligibility assumptions, domain assumptions and heuristic assumptions* the following set of properties is obtained.

Descriptive domain assumptions:

- Evaluation scores are considered in the form of linguistic variables.
- The preference model is a complete pre-order structure.
- The most useful result for policy-making is considered to be a complete ranking of policy options.

Normative domain assumptions:

- Simplicity is desirable and means the use of a few ad hoc parameters as possible.
- Weights are meaningful only as importance coefficients.
- A minority principle must be implemented for ethical and prudential reasons.

Formal domain assumptions:

- α*-level* equivalence classes obtained by using max–min composition operations arriving at a max–min transitive closure.
- Stability of the clustering algorithm.
- Generality of the clustering algorithm.
- No reversal in the dendrogram.
- Monotonicity.
- Diversity.

- Symmetry.
- Positive Responsiveness.
- Unanimity.
- Neutrality.
- Reinforcement.

Heuristic descriptive assumptions:

- Transparency is a desirable feature of policy processes.
- Social actors and policy options can always be identified in a satisfactory way.
- The social impact matrix is a consistent and meaningful representation of the qualitative field work (institutional analysis, interviews, questionnaires, focus groups, and so on).
- A fuzzy cluster algorithm is a good tool for forming an idea of the credibility of similarities/diversities among social actors.

Heuristic formal assumptions:

- Semantic distance as a conflict indicator.
- C–K–Y–L ranking procedure as a proper tool for implementing the majority principle.
- Local stability of the ranking method.
- Cycle-breaking without losing neutrality.
- Proportional veto function as a proper tool for implementing the minority principle.

Negligibility formal assumptions:

- Anonymity.
- Independence of irrelevant alternatives.

Appendix 8.1

Let X and Y be two universes of discourse, a fuzzy binary relation R in the Cartesian product $X \times Y$ is a fuzzy set in $X \times Y$ defined by the membership function

$$\mu_R : X \times Y \rightarrow [0,1]$$
$$(x, y) \rightarrow \mu_R(x, y), \qquad x \in X \text{ and } \quad y \in Y$$

where the grade of membership μ_R (x, y) indicates the degree of relationship between x and y.

The *max–min composition* is a standard operation for two fuzzy relations: given two relations $R(x, y)$, $Q(y, z)$ defined on $X \times Y$ and $Y \times Z$, respectively, the max–min composition of R and Q, denoted as $R \circ Q$, is defined by

$$\mu_{R \circ Q}(x, z) = \max_{y \in Y} \min\left[\mu_R(x, y), \mu_Q(y, z)\right] \quad x \in X, y \in Y \text{ and } z \in Z.$$

By using the notion of max–min composition, one can derive new fuzzy relations. A transitive closure can be obtained by means of the following theorem (Leung, 1988, p. 125):

Theorem 1

Let R be any fuzzy binary relation. If for some k, the max–min composition $R^{k+1} = R^k$, *then the max–min transitive closure is*

$$\hat{R} = R \cup R^2 \cup R^3 \cup \ldots \cup R^k .$$

Chapter 9
Conclusions

A proper evaluation of sustainability options needs to deal with the plurality of legitimate values and interests coexisting in society. In empirical evaluations of public projects and publicly provided goods, multi-criteria decision analysis seems to be an appropriate policy tool, since it makes it possible to take into account a wide range of assessment criteria (e.g. environmental impact, distributional equity, and so on) and not simply profit maximization, as a private economic agent might do. However, the management of a policy process involves many layers and kinds of decisions, and requires the construction of a *dialogue process* between many social actors, individual and collective, formal and informal, local and beyond.

In general, these concerns were not considered terribly relevant in scientific research as long as the basic implicit assumption held that time was an infinite resource. However, the unprecedented nature of the policy problems faced in this third millennium (e.g. mad cow disease, genetically modified organisms, etc.), implies that there may very often be long term consequences of using science for policy-making. Scientists and policy-makers are confronting issues where, as stated by Funtowicz and Ravetz "*facts are uncertain, values in dispute, stakes high and decisions urgent*". In these circumstances, scientists cannot provide any useful input without interacting with the rest of society, and vice versa, the rest of the society cannot make any sound decision without interacting with scientists. That is, the question of "*how to improve the quality of a social decision process*" must be put, rather urgently, on the agenda of scientists, decision makers and indeed the society as a whole.

One outcome of this discussion is that the political and social framework must find a place in multi-criteria decision analysis. An effective policy exercise should consider not merely the measurable and contrastable dimensions of the simple parts of the system, which even if complicated, may be technically simulated (technical incommensurability). To be realistic it should also deal with the higher dimensions of the system, i.e. those dimensions in which power relations, hidden interests, social participation, cultural constraints, and other soft values become relevant and unavoidable variables that strongly, but not deterministically, affect the possible outcomes of the strategies to be adopted (social incommensurability).

The idea of *social multi-criteria evaluation* (SMCE) has thus been developed, in which the pitfalls of the technocratic approach are avoided by applying different methods of sociological research. For example, "*institutional analysis*", performed mainly on historical, legislative and administrative documents, can produce a map of the relevant social actors. By means of focus groups it is possible to form an idea of people's desires and then to develop a set of policy options and evaluation criteria. The main limitations of focus group techniques are that they are not supposed to be a representative sample of the population and that sometimes people are not willing to participate or to state publicly what they really think (above all in small towns and villages). For this reason anonymous questionnaires and personal interviews are an essential part of the participatory process.

One should note that policy evaluation is not a one-shot activity. On the contrary, it evolves as a *learning process* which is usually highly dynamic, so that judgements regarding the political relevance of elements, alternatives or impacts may undergo sudden changes. Hence a policy analysis must be flexible and adaptive in nature. This is why evaluation processes have a *cyclical nature*. By this is meant the possible adaptation and modification of elements of the evaluation process due to continuous feedback loops among the various steps and consultations among the actors involved (see Fig. 9.1). Of course, the steps of the process are not rigidly set out. On the contrary, flexibility and adaptability to real-world situations are some of the main advantages of social multi-criteria evaluation.

At this stage a question arises: what is the role of mathematical aggregation procedures in a social evaluation process for sustainability policies? Mathematical aggregation conventions naturally play an important role, i.e. to ensure that the rankings obtained are *consistent* with the information and the assumptions used

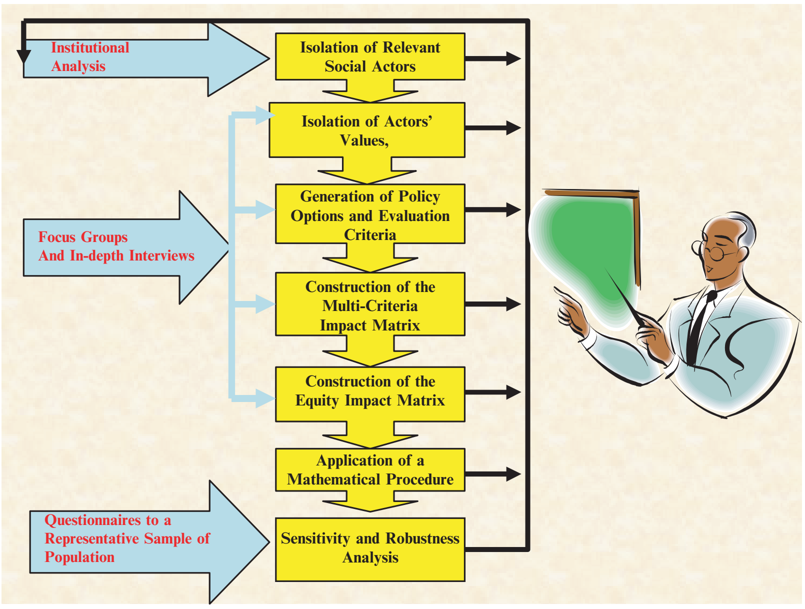


Fig. 9.1 The ideal problem-structuring in SMCE

throughout the structuring process. It can be stated quite safely that a problem of multi-criteria decision theory is the many different mathematical aggregation conventions (or methods) that exist. In the case of SMCE, the following considerations can help in selecting appropriate methods.

The idea of *social incommensurability* makes the following properties desirable in a social multi-criteria method:

- Weights in this framework are clearly meaningful only as importance coefficients and not as trade-offs. As a consequence, complete compensability cannot be implemented.
- Conflict analysis procedures explicitly looking for social compromises should integrate an SMCE exercise.
- In a policy framework, having a ranking of all the alternatives is more useful than selecting one alternative only.

The idea of *technical incommensurability* makes the following properties desirable in a social multi-criteria method:

- Partial or complete non-compensability is an essential consistency requirement.
- Indifference and preference thresholds should be explicitly taken into account.
- Mixed information of the widest range should be addressed in a consistent way.
- Simplicity, meaning the use of as few parameters as possible, is a very desirable property.
- The hierarchical dimension of a policy problem should be explicitly considered.

Since none of the most commonly used multi-criteria methods can boost of having all the properties considered desirable in an SMCE framework, a new mathematical aggregation convention for the solution of the so-called discrete multi-criterion problem in an SMCE context has been developed. This multi-criterion aggregation convention can be divided into two main steps:

- Pair-wise comparison of alternatives
- Ranking of alternatives in a complete pre-order

Throughout the entire pair-wise comparison step, it is guaranteed that ordinal, crisp, stochastic and fuzzy criterion scores are tackled equivalently. The double threshold model, generating a pseudo-order structure, is used for preference modelling, thus avoiding the so-called Luce paradox. To deal with the lack of stability of the pseudo-order structure, valued preference relations modelled by means of fuzzy preference relations, are introduced. Weights are never combined with intensities of preference, thereby establishing the theoretical guarantee that they are importance coefficients. Given that the sum of weights is equal to one, the pair-wise comparisons can be synthesized in a matrix, which can be interpreted as a voting matrix. The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule.

A problem connected with the use of Condorcet consistent rules is that of cycles. A cycle-breaking rule normally demands some arbitrary choices, such as deleting the cycle with the lowest support, and so on. In search of a non-arbitrary cycle-breaking rule the so-called Condorcet–Kemeny–Young–Levenglick ranking procedure was chosen; no arbitrary choice is required with this procedure. Given the fact that criterion weights are used, anonymity is necessarily lost. However, given that Arrow's impossibility theorem forces us to make trade-offs between decisiveness and anonymity, the loss of anonymity in favour of decisiveness in this framework is a positive feature. Consistently with the normative tradition in political philosophy and following the prudence axiom, the minority principle is introduced by means of a veto power.

In the area of environmental and resource management and in policies aiming at a sustainable development, conflicting issues and interests are the normal state of affairs. Distributional issues have been taken into consideration by means of an eclectic approach using concepts from land-use planning, fuzzy cluster analysis and social choice. Starting with a matrix showing the impact of different courses of action on each different social actor, a fuzzy clustering procedure is used to identify the groups whose interests are relatively closer to each other. This is more or less in agreement with the hypotheses underlying the "*minimal range theory*" in coalition formation literature. Rankings for each "credible" group of social actors are obtained by means of the majority principle implemented by using a Condorcet voting principle. The issue of cycles has been tackled adequately. The minority principle has also been considered by means of Moulin's proportional veto function.

In order to use these mathematical procedures (developed in Chaps. 7 and 8), in real-world applications a user friendly software package has been developed at the Joint Research Centre of the European Commission at Ispra. It is called SOCRATES (*SOcial multi-CRiteriA for The Evaluation of* S*ustainability*).

To conclude it is possible to state that social multi-criteria evaluation provides a powerful framework for the implementation of the incommensurability principle. In fact it meets the requirements of being *inter-multi-disciplinary* (with respect to the research team), *participatory* (with respect to the local community) and *transparent* (since all criteria are presented in their original form without any transformations into money, energy or whatever common measurement rod). As a consequence SMCE looks as an adequate assessment framework for (micro and macro) sustainability policies.

However, one should remember that we live in a "second-best" world. A useful analogy here is with *Flatland*, the classic Victorian science fiction and social parody (Abbott, 1935). There, the inhabitants of spaces with three dimensions had a richer awareness of themselves, and could also see beyond and through the consciousness of the simpler creatures inhabiting fewer dimensions. At this stage it is not unfair to reveal the dénouement of the story, ultimately the "*Sphere*" of three-dimensional space shows himself to be just another Flatlander at heart, when he angrily refuses to credit the existence of higher dimensions of being …

Annex: A Sustainability Composite Indicator Based on Multi-Criteria and Sensitivity Analysis[1]

A.1 The Empirical and Measurement Framework

The scope of this application is to measure sustainability at a regional level. The regions chosen are mainly Spanish, although some Mediterranean regions have been considered for the sake of comparison. The three dimensions of sustainability chosen – environment, society, economy – are described by 29 indicators or variables (Table A.1). The data used are the most recent (2004, or nearest year) data from the EUROSTAT regional database. This choice was made so as (a) to ensure high data quality, since EUROSTAT has given particular care to the issue of data quality in recent years; and (b) to allow benchmarking between Spanish and other Mediterranean regions.

The proposed framework builds on the best data available at regional level for the 17 Spanish regions, as these derive from the REGIO database of EUROSTAT and the Spanish National Statistical Office. We furthermore found sufficient data in the REGIO database for four Greek and four Italian regions which have similar climatic and economic conditions to the Spanish.

The ranking of these regions can easily be obtained by using the non- compensatory multi-criteria approach described in Chap. 7. The results obtained provide firm ground for the analysis of regional-level sustainability performance. The findings and a review of the Spanish leaders and laggards in sustainability performance confirm some common perceptions about the determinants of policy success. But they also reveal some surprises and otherwise unexpected relationships among regions.

The overall ranking is based on equal weights for the indicators within each dimension, as well as equal weights for each dimension. This decision is due to the fact that all dimensions have approximately the same number of individual indicators. Table A.2 presents the final non-compensatory ranking and the ranking for each of the three sustainability dimensions. Among the 17 Spanish regions, the *top*

[1] This Annex has been written in collaboration with Michaela Saisana. Financial support form *fBBVA* is acknowledged.

Table A.1 Dimensions and indicators used

Dimensions	Indicators
Environmental	Agricultural area
	Forest area
	Distances driven by trucks
	Municipal waste collected
	Forest area affected by fires
	Non-differentiated urban waste
	Differentiated urban waste
	Cement (consumption &sales)
	Abstraction of total fresh water by public water supply
	Investments in waste water collection and treatment facilities
	Population affected by diseases of the respiratory system
Social	Ageing population
	Infant mortality rate
	Population density
	Crude birth rate
	Number of hospital beds
	Number of physicians/doctors
	Population affected by mental and behavioural disorders
	Population affected by alcohol abuse
	Participation in General Elections
	Population in prison
Economic	Employment
	Gross Domestic Product
	Households with minimum subsidy or no income
	Foreigners with tertiary education
	Lifelong learning
	Population employed in hi-tech
	Patent application to the EPO
	Total intramural R&D expenditure (GERD)

Table A.2 Ranking of the Spanish regions in sustainability and its three sub-dimensions

	Overall rank	Environment rank	Society rank	Economy rank
Madrid	1	5	4	1
Navarra	2	1	1	2
Catalonia	3	7	7	3
Rioja	4	3	5	4
Balearic islands	5	2	14	9
Pais Vasco	6	6	8	5
Murcia	7	4	3	11
Valencia	8	17	6	6
Aragon	9	10	12	7
Cantabria	10	9	10	12
Castilla y Leon	11	15	13	8
Canary Islands	12	13	15	10
Asturias	13	14	17	14
Andalucia	14	11	9	16
Castilla la Mancha	15	12	11	15
Extremadura	16	8	2	17
Galicia	17	16	16	13

five are Madrid, Navarra, Catalonia, Rioja and the Balearic islands. The *lowest five* are Asturias, Andalucia, Castilla la Mancha, Extremadura and Galicia. Mid-ranking performers include the remaining seven regions – Pais Vasco, Murcia, Valencia, Aragon, Cantabria, Castilla y Leon and the Canary Islands. It is interesting to note that the top performing regions do not necessarily perform exceptionally in all three dimensions. In fact, Catalonia has middling performance in two dimensions (environment and society), whilst the Balearic Islands have middling performance in economy and low performance in society. On the other hand, the bottom five regions do not necessarily have the worst performance in all three sub-dimensions. For example, Extremadura has a top-five performance in society, while Andalucia and Castilla la Mancha have mid-table performance in two dimensions (environment and society). Among the mid-ranking regions, performance is medium in all three sub-dimensions for Aragon and Cantabria. Exceptionally, Murcia, despite its top-five performance in two dimensions (environment and society), is ranked in the middle of the table. The opposite is observed for three regions – Castilla y Leon, the Canary Islands and Asturias.

While each region has unique socio-economic and geographic characteristics, environmental policy priorities, and development goals, cross-regional comparisons

Table A.3 Sustainability ranking of the Spanish and selected Italian and Greek regions

	All regions studied rank	Only Spanish regions rank
Lombardy (IT)	1	
Madrid	2	1
Catalonia	3	3
Tuscany (IT)	4	
Navarra	5	2
Rioja	6	4
Balearic Islands	7	5
Veneto (IT)	8	
Pais Vasco	9	6
Aragon	10	9
Valencia	11	8
Murcia	12	7
Cantabria	13	10
Sicily (IT)	14	
Castilla y Leon	15	11
Andalucia	16	14
Castilla la Mancha	17	15
Canary Islands	18	12
Kriti (GR)	19	
Attiki (GR)	20	
Asturias	21	13
Extremadura	22	16
Galicia	23	17
St. Ellada (GR)	24	
Thessalia (GR)	25	

between various countries can nevertheless yield useful insights. To this end, we selected four Italian and four Greek regions which are similar to the Spanish regions in terms of socio-economic development, climate, land area, and population density. Table A.3 shows the ranking of the Spanish regions before and after the inclusion of the Italian and Greek regions.

Among the 25 regions studied, Lombardy (the Italian region hosting Milan) performs best in overall sustainability. Madrid and Catalonia follow, with Tuscany (IT) arriving in the 4th place. In general, three of the four Italian regions we analysed are among the top eight. Sicily (IT) performs lower than the other three Italian regions, although it has a midtable position (14th) in the overall sustainability ranking. All four Greek regions, including Attiki (the most urbanized region of Greece, which hosts Athens and is recognized as the "business capital" of the country), rank 19th or lower.

It is interesting to note that once Italian and Greek regions enter into the non-compensatory ranking system the order of the Spanish regions is slightly affected. There are two pairs of regions where this is observed: Catalonia and Navarra (shift between second and third rank), and Aragon and Murcia (shift between 7th and 9th rank). This result can be explained in part by the fact that the non-compensatory ranking procedure, as we saw in Chap. 7, does not respect the property of the independence of irrelevant alternatives; moreover the indicator framework for the Greek and Italian regions is slightly reduced by the seven indicators that were available only for the Spanish regions.

A.2 Sensitivity Analysis

By acknowledging a variety of methodological assumptions that are intrinsic to policy research, "sensitivity analysis" can determine whether the main results of a ranking system change substantially when those assumptions are varied over a reasonable range of possibilities (Saltelli et al., 2000, 2004; Saisana et al., 2005). Here the robustness of the ranking can be assessed by studying its sensitivity to three types of technical uncertainties: (a) the indicator set, (b) the aggregation rule and (c) the set of weights.

To begin, we compared the impact on the ranking of excluding a single indicator from the framework and using either a linear/compensatory or a non-linear/non-compensatory multi-criteria aggregation rule, while maintaining equal weights for all the indicators included. Normalization is not needed in case of the multi-criteria approach, while a min-max scaling in (0, 1) was undertaken prior to the linear aggregation. Thus 30 scenarios were analysed for each type of aggregation, one with the entire set of 29 indicators, and 29 indicator sets of 28 indicators each. Table A.4 provides statistics for the regions' "rank range", i.e. the width of the interval between the worst and the best case scenario, in either the linear or the multi-criteria approach.

Table A.4 Statistics for the rank range (difference between best and worst case scenario) of the 25 regions studied

	Multi-criteria	Additive (Linear)
Minimum	0 (St. Ellada, Galicia, Thessalia)	2 (Galicia, Navarra)
Average	3	7
Maximum	10 (Tuscany)	14 (Canary Islands)
Standard deviation	2.4	3.0
Less than (≤) 2 positions	10	2
More than (≥) 5 positions	5	20

An interesting feature revealed by Table A.4 is the sensitivity of the linear system to the exclusion of a single indicator, as opposed to the more stable results produced by the multi-criteria procedure. The significant impact of such a small structural change on the ranking produced by the linear aggregation is due to the compensation effects among indicators.

In fact, in the linear system only two regions are not affected by the exclusion of a single indicator (i.e. shift ≤2 positions), while 20 regions shift more than (≥) five positions. On the contrary, in the non-linear/non-compensatory multi-criteria approach ten regions are not affected, while only five regions (as opposed to 20) shift more than five positions. To complement these results, Figs. A.1 and A.2 present the median, best and worst rank across the 30 scenarios in the multi-criteria and the additive aggregation, respectively. The four regions whose multi-criteria rank is affected by the selection of indicators are Veneto (IT), Tuscany (IT), Sicily (IT) and Catalonia. The wide rank range for Veneto and Catalonia is due to several indicators, while only two indicators influence the rank of Tuscany and Sicily. To be more specific, Sicily's rank is sensitive to "Infant mortality rate" and "Patent application", while Tuscany's is sensitive to "Employment" and "Foreigners with tertiary education". In the linear ranking system, the regions that display the widest rank range (more than ten positions) are Attiki (GR), Sicily (IT) and the Canary Islands. These results have shown that the non-linear/non-compensatory ranking system is robust to small changes in the indicator set (i.e. the exclusion of one indicator at-a-time).

After having studied the impact of the exclusion of a single indicator on the multi-criteria ranking and confronted it with that of the linear aggregation, we next analyse which regions would be affected by the choice of the aggregation system when all indicators are included, and why. Figure A.3 presents the scatterplot of the multi-criteria ranks versus those of the linear aggregation. In both cases, all indicators are weighted equally. This graph allows one to see immediately which regions are compensating for their deficiencies in some indicators with a relatively good performance in other indicators under a linear/compensatory aggregation rule. All those regions are found at the bottom-right of Fig. A.3, e.g. Attiki, Kriti, Extremadura and Thessalia. Another apparent feature is that the aggregation method primarily affects the mid-ranking regions and, to a lesser extent, the most

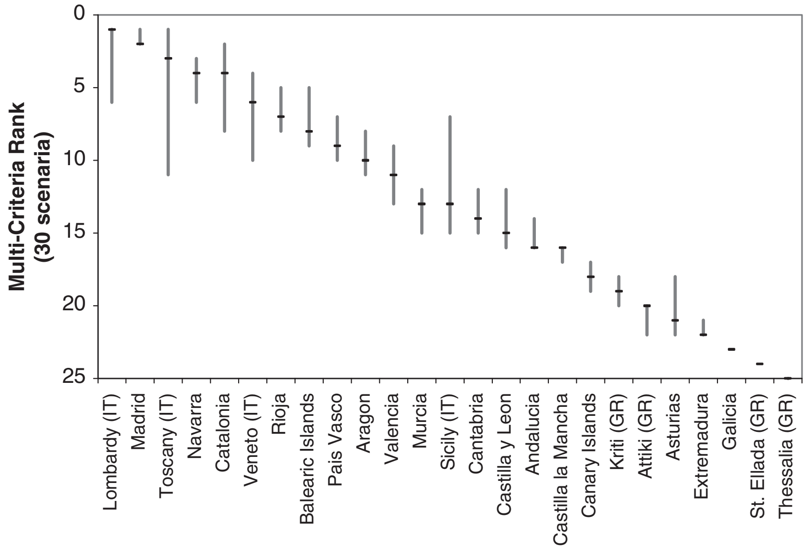


Fig. A.1 Multi-criteria based ranking. Black Marks Correspond to the median simulation rank. Whiskers show the best and worst rank of 30 scenarios produced either by considering all 29 indicators or by excluding one indicator at a time

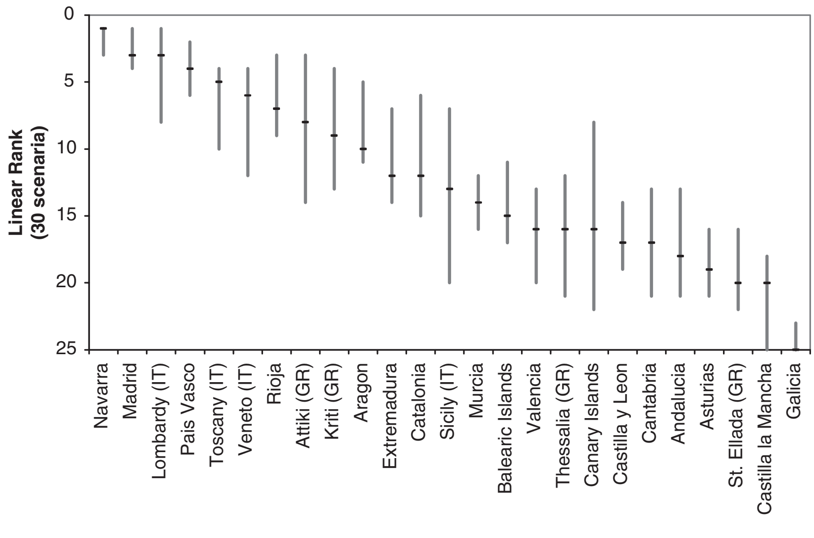


Fig. A.2 Additive (linear) based ranking. black marks correspond to the median simulation rank. Whiskers show the best and worst rank of 30 scenarios produced either by considering all 29 indicators, or by excluding one indicator at a time

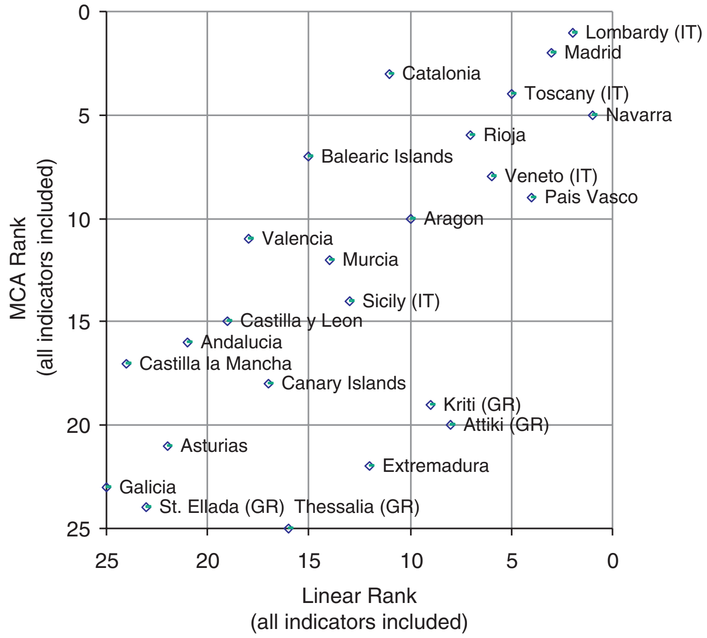


Fig. A.3 Non-compensatory multi-criteria aggregation (MCA) of indicators versus linear aggregation of indicators

or least sustainable regions. The two aggregation approaches have a Spearman correlation coefficient, $r = 0.643$.

We finally study the impact on the ranking of the set of weights to be assigned to the indicators. We employ data envelopment analysis (Charnes et al., 1978; Charnes and Cooper, 1985), which uses linear optimization rules to calculate region-specific weights for the indicators, accepting that there is no (expert) consensus on the appropriate set of weights for the indicators. Moreover, several authors have argued that differential weighting may be desirable in composite indicators, e.g. because of different environments or political attitudes in different countries or regions, or because the very idea of imposing weights may be inconsistent with the subsidiarity principle (Cherchye et al., 2004). Basically, such worries may be then overcome by rendering the weight selection problem endogenous for each observation. That is, the relative weight assigned to each indicator is endogenously determined in this type of performance evaluation model, so as to reflect the associated relative performance for the region under evaluation (Melyn and Moesen, 1991).

Figure A.4 presents the scatterplot of the multi-criteria based ranking with that produced by data envelopment analysis (DEA). A common feature of the two approaches is that the normalization stage is not required, as both methods are unit

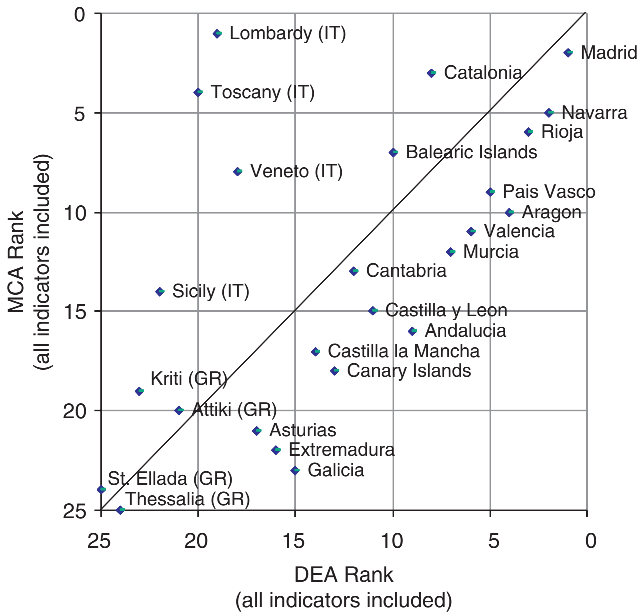


Fig. A.4 Non-compensatory multi-criteria aggregation (MCA) of indicators versus data envelopment analysis (DEA) aggregation of indicators

invariant. In our DEA application we require that the relative pie-share of each indicator (i.e. the product of indicator value and the respective weight) should lie between 3% and 20% of the aggregate score. These constraints, which are recommended by recent DEA applications (Cherchye et al., 2007) were added to avoid allowing regions to achieve a high score simply by assigning zero weight to all indicators for which they have a low performance, or by assigning an unreasonably high weight to a single indicator. The rank order correlation coefficient is slightly lower than before, i.e. $r = 0.564$. The four Italian regions are those that are most affected by the DEA system, landing in a much lower position than their multi-criteria equivalent. Again, given the resemblance of the data envelopment analysis to a linear aggregation system, the impact on the regions' ranks of excluding a single indicator from the data-set when using DEA is pronounced (results not presented here).

Sensitivity analysis helps to gauge the robustness of the results obtained, to increase the transparency of the ranking system, to identify the regions that improve or deteriorate under certain assumptions, and to help the in framing of debate with the proposed framework.

Bibliography and References

Abbott, E.A. (1935) *Flatland. A romance of many dimensions*. Boston, Little, Brown & Co.
Ackoff, R.L. (1978) *The art of problem solving*. Wiley, New York.
Allais, M. (1979) The so-called Allais paradox and rational decision under uncertainty, in Hagen O. and Allais M. (eds.) *Expected utility hypotheses and the Allais paradox*. Reidel, Dordrecht, pp. 437–681.
Allen, T.F.H., Hoekstra, T.W. (1992) *Toward a unified ecology*. Columbia University Press, New York.
Allen, T.F.H., Tainter, J.A., Hoekstra, T.W. (2003) *Supply-side sustainability*. Columbia University Press, New York.
Anderberg, M.R. (1973) *Cluster analysis for applications*. Academic Press, New York.
Anderson, N.H., Zalinski, J. (1988) Functional measurement approach to self-estimation in multiattribute evaluation. *Journal of Behavioural Decision Making*, 1, pp. 191–221.
Arrow, K.J. (1963) *Social choice and individual values*. Second edition, Wiley, New York.
Arrow, K.J. (1971) *Essays in the theory of risk bearing*. North-Holland, Amsterdam.
Arrow, K.J. (1997) Invaluable Goods. *Journal of Economic Literature*, 35(2), pp. 757–763.
Arrow, K.J., Raynaud, H. (1986) *Social choice and multicriterion decision making*. MIT Press, Cambridge.
Bana e Costa, C.A. (1990) An additive value function technique with a fuzzy outranking relation for dealing with poor intercriteria preference information, in Bana e Costa C.A. (ed.) *Readings in multiple criteria decision aid*. Springer, Berlin, Heidelberg, New York, pp. 351–382.
Bana e Costa, C.A., De Corte, J.M., Vansnick, J.C. (2005) On the mathematical foundations of MACBETH, in Figueira J., Greco S. and Ehrgott M. (eds.) *Multiple-criteria decision analysis. State of the art surveys*. Springer International Series in Operations Research and Management Science, New York, pp. 409–442.
Banville, C., Landry, M., Martel, J.M., Boulaire, C. (1998) A stakeholder approach to MCDA. *Systems Research and Behavioral Science*, 15, pp. 15–32.
Barbier, E.B. (1987) The concept of sustainable economic development. *Environmental Conservation*, 14(2), pp.101–110.
Barbier, E.B., Markandya, A. (1990) The conditions for achieving environmentally sustainable growth. *European Economic Review*, 34, pp. 659–669.
Barbiroli, G. (1993) Direct and indirect indicators of sustainable development. Proceedings International Conference on Environmental Pollution, Barcelona, pp. 489–496.
Barthelemy, J.P., Guenoche, A., Hudry, O. (1989) Median linear orders: heuristics and a branch and bound algorithm. *European Journal of Operational Research*, 42, pp. 313–325.
Baumol, W.J. (1969) *Welfare Economics and the Theory of the State*. Bells and Sons, London.
Baumol, W.J., Oates, W.E. (1972) The use of standards and prices for protection of the environment. *Swedish Journal of Economics*, 73, pp. 42–54.
Beckerman, W., Pasek, J. (1997) Plural values and environmental valuation. *Environmental Values*, 6(1), pp. 65–86.

Begon, M., Harper, J.L., Townsend, C.R. (1996) *Ecology: individuals, population and communities*. Third edition, Blackwell, Oxford.

Beinat, E. (1997) *Value functions for environmental management*. Kluwer, Dordrecht.

Beinat, E., Nijkamp, P. (eds.) (1998) *Multicriteria evaluation in land-use management: methodologies and case studies*. Kluwer, Dordrecht.

Bell, D.E., Raiffa, H., Tversky, A. (eds.) (1988) *Decision making: descriptive, normative and prescriptive interactions*. Cambridge University Press, Cambridge.

Bell, M.L., Hobbs, B., Elliott, E.M., Ellis, H., Robinson, Z. (2001) An evaluation of multi-criteria methods in integrated assessment of climate policy. *Journal of Multi-Criteria Decision Analysis*, 10(5), pp. 229–256.

Belton, V., Stewart, T.J. (2002)*Multiple criteria decision analysis: an integrated approach*. Kluwer, Dordrecht.

van den Bergh, J.C.J.M., Verbruggen, H. (1999) Spatial sustainability, trade and indicators: an evaluation of the ecological footprint. *Ecological Economics*, 29, pp. 61–72.

Bezdek, J.C. (1980) *Pattern recognition with fuzzy objective functions algorithms*. Plenum, New York.

Black, D. (1958) *The theory of committees and elections*. Cambridge University Press, Cambridge.

Bogetoft, P., Pruzan, P. (1991) *Planning with multiple criteri*. North-Holland, Amsterdam.

Bojö, J., Mäler, K.G., Unemo, L. (1990) *Environment and development: an economic approach*. Kluwer, Dordrecht.

Borda, J.C. de (1784) Mémoire sur les élections au scrutin, in *Histoire de l' Académie Royale des Sciences*. Paris.

Bordes, G., Tideman, N. (1991) Independence of irrelevant alternatives in the theory of voting. *Theory and Decision*, 30(2), pp. 163–186.

Bouyssou, D. (1986) Some remarks on the notion of compensation in MCDM. *European Journal of Operational Research*, 26, pp. 150–160.

Bouyssou, D. (1990) Building criteria: a prerequisite for MCDA, in Bana e Costa C.A. (ed.) *Readings in multiple criteria decision aid*. Springer, Berlin, Heidelberg, New York, pp. 58–80.

Bouyssou, D., Vansnick, J.C. (1986) Non-compensatory and generalized non-compensatory preference structures. *Theory and Decision*, 21, pp. 251–266.

Bouyssou, D., Marchant, Th., Pirlot, M., Tsoukiàs, A., Vincke, Ph. (2000) *Evaluation and decision models: a critical perspective*. Kluwer, Dordrecht.

Brans, J.P., Mareschal, B. Vincke, Ph. (1984) PROMETHEE: a new family of outranking methods in multicriteria analysis, in Brans J.P. (ed.) *Operational research '84*. Elsevier, Amsterdam, pp. 408–421.

Brans, J.P., Mareschal, B., Vincke, Ph. (1986) How to select and how to rank projects. The PROMETHEE method. *European Journal of Operational Research*, 24, pp. 228–238.

Bresso, M. (1993) *Per un'Economia Ecologica*. La Nuova Italia Scientifica.

Bromley, D. (1998) Searching for sustainability: the poverty of spontaneous order. *Ecological Economics*, 24, pp. 231–240.

Buchanan, J.M., Musgrave, R.A. (1999) *Public finance and public choice*. The MIT Press, Cambridge.

CEROI *Cities Environment Report on the Internet*. *http://www.ceroi.org*

Charnes, A., Cooper, W.W. (1985) Preface to topics in data envelopment analysis. *Annals of Operations Research*, 2, pp. 59–94.

Charnes, A., Cooper, W.W., Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operations Research*, 2, pp. 429–444.

Charon, I., Guenoche, A., Hudry, O., Woirgard, F. (1997) New results on the computation of median orders. *Discrete Mathematics*, 165/166, pp. 139–153.

Cherchye, L., Moesen, W., Van Puyenbroeck, T. (2004) Legitimately diverse, yet comparable: on synthesising social inclusion performance in the EU. *Journal of Common Market Studies*, 42, pp. 919–955.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., Tarantola, S. (2007) Creating composite indicators with data envelopment analysis and robustness analysis: the case of the technology achievement index. *Journal of Operational Research Society* forthcoming

Chernoff, H. (1954) Rational selection of decision functions. *Econometrica*, 22(4), pp. 422–443.

Chichilnisky, G. (1996) An axiomatic approach to sustainable development. *Social Choice and Welfare*, 13(2), pp. 219–248.

Climaco, J.C.N. (2000) *Uma reflexao critica da decisao óptima*. Abertura Solene das Aulas na Universidade de Coimbra, Coimbra.

Cocossis, H., Nijkamp, P. (eds.) (1995) *Planning for our cultural heritage*. Averbury, London.

Cohen, W., Schapire, R., Singer, Y. (1999) Learning to order things. *Journal of Artificial Intelligence Research*, 10, pp. 213–270.

Condorcet, Marquis de (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la probabilité des voix*. De l' Imprimerie Royale, Paris.

Coombs, C.H. (1958) On the use of inconsistency of preferences in psychological measurement. *Journal of Experimental Psychology,* 55, pp. 1–7.

Copeland, A.H. (1951) *A reasonable social welfare function*. University of Michigan, mimeo.

Corral-Quintana, S.A. (2000) *Una metodología integrada de exploración y comprensión de los procesos de elaboración de políticas públicas*. University of La Laguna, Department of Economics, Spain, ISBN 84-688-4314-8.

Corral-Quintana, S.A., De Marchi, B., Funtowicz, S., Gallopín, G., Guimarães-Pereira, Â., Maltoni, B. (2001) *The Visions Project at the JRC*. European Commission, Joint Research Centre, Institute for the Protection and Safety of the Citizens, EUR 19926 EN (*http://alba.jrc.it/visions)*.

Costanza, R. (ed.) (1991) *Ecological economics: the science and management of sustainability*. Columbia University Press, New York.

Cropper, M.L., and Oates, W.E. (1992) Environmental economics: a survey. *Journal of Economic Literature*, 30, pp. 675–740.

Daly, H.E. (1977) *Steady-state economics*. Freeman, San Francisco, CA.

Daly, H.E., Cobb, J.J. (1989) *For the common good: redirecting the economy toward community, the environment and a sustainable future*. Beacon Press, Boston.

Dasgupta, A.K., Pearce, D.W. (1972) *Cost–benefit analysis*. MacMillan, London.

Dasgupta, A.K., Marglin, S., Sen, A. (1972) *Guidelines for project evaluation*. UNIDO, New York.

Davenport, A., Kalagnanam, J. (2004) *A computational study of the Kemeny rule for preference aggregation*, in McGuinness D.L., Ferguson G. Proceedings of the Nineteenth National Conference on Artificial Intelligence. Sixteenth Conference on Innovative Applications of Artificial Intelligence, July 25–29, 2004, San Jose, California. AAAI Press/The MIT Press, USA.

D'Avignon, G., Vincke, Ph. (1988) An outranking method under uncertainty. *European Journal of Operational Research*, 36, 311–321.

Debreu, G. (1960) Topological methods in cardinal utility theory. in Arrow K.J., Karlin S. and Suppes P. (eds.) *Mathematical methods in the social sciences*. Stanford University Press, Stanford.

De Marchi, B., Ravetz, J. (2001) *Participatory approaches to environmental policy*. Concerted Action EVE, Policy Research Brief, No. 10.

De Marchi, B., Funtowicz, S.O., Lo Cascio, S., Munda, G. (2000) Combining participative and institutional approaches with multi-criteria evaluation. An empirical study for water issue in Troina, Sicily. *Ecological Economics*, 34(2), pp. 267–282.

Diamond, P., Stiglitz, J.E. (1974) Increases in risk and risk aversion. *Journal of Economic Theory*, 8, pp. 337–360.

Dietz, F.J., van der Straaten, J. (1992) Rethinking environmental economics: missing links between economic theory and environmental policy. *Journal of Economic Issues*, XXVI(1), pp. 27–51.

Dubois, D., Prade, H., Sabbadin, R. (2001) Decision-theoretic foundations of qualitative possibility theory. *European Journal of Operational Research*, 128(3), pp. 459–478.

Duchin, F., Lange, G.M. (1994) *The future of the environment*. Oxford University Press, Oxford.
Dwork, C., Kumar, R., Naor, M., Sivakumar, D. (2001) *Rank aggregation methods for the web*. Proceedings of 10th WWW, pp. 613–622.
Ecological Economics (2000) *Forum: the ecological footprint*. 32(3), pp. 341–394.
Ehrlich, P., Raven, P. (1964) Butterflies and plants: a study in co-evolution. *Evolution*, 18, pp. 586–608.
Espelta, J.M., Retana, J., Habrouk, A. (2003) An economic and ecological multi-criteria evaluation of reforestation methods to recover burned *Pinus Negra* forests in NE Spain. *Forest Ecology and Management*, 180, pp. 185–198.
Esty, D.C., Levy, M., Srebotnjak, T., de Sherbinin, A. (2005) *Environmental sustainability index: benchmarking national environmental stewardship*. New Haven, Conn.: Yale Center for Environmental Law and Policy.
European Commission (1996) *Environmental indicators and green accounting*. Brussels.
European Commission (2001) *European governance. A white paper*. Brussels, 25.7.2001, COM 428 final.
European Environment Agency (2001) *Late lessons from early warnings: the precautionary principle 1896–2000*. Environmental Issue Report, No. 22, Copenhagen.
Falconi, F. (2002) *Economía y desarrollo sostenible*. FLACSO, Quito.
Faucheux, S., O'Connor, M. (eds.) (1998) *Valuation for sustainable development:mMethods and policy indicators*. Edward Elgar, Cheltenham.
Figueira, J., Roy, B. (2002) Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139, pp. 317–326.
Figueira, J., Greco, S., Ehrgott, M. (eds.)(2005) *Multiple-criteria decision analysis. State of the art surveys*. Springer International Series in Operations Research and Management Science, New York.
Fishburn, P.C. (1970) Utility theory with inexact preferences and degrees of preference. *Synthese*, 21, pp. 204–222.
Fishburn, P.C. (1973a) Binary choice probabilities: on the varieties of stochastic transitivity. *Journal of Mathematical psychology*, 10, pp. 327–352.
Fishburn, P.C. (1973b) *The theory of social choice*. Princeton University Press, Princeton.
Fishburn, P.C. (1978) A survey of multiattribute/multicriteria evaluation theories, in Ziontz S. (ed.) Multicriteria problem solving. Springer, Berlin, Heidelberg, New York, pp. 181–224.
Fishburn, P.C. (1982) Monotonicity paradoxes in the theory of elections. *Discrete Applied Mathematics*, 4, pp. 119–134.
Fishburn, P.C. (1984) Discrete mathematics in voting and group choice. *SIAM Journal of Algebraic and Discrete Methods*, 5, pp. 263–275.
Fishburn, P.C. (1991) Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty* 4, 113–134.
Fishburn, P.C., Gehrlein, W.V., Maskin, E. (1979) Condorcet's proportions and Kelly's conjecture. *Discrete Applied Mathematics*, 1, pp. 229–252.
Folke, C. (1991) Socio-economic dependence on the life-supporting environment, in Folke C. and Kaberger T. (eds.) *Linking the natural environment and the economy: essays from the Eco–Eco Group*. Kluwer, Dordrecht, pp. 77–94.
Folke, C., Larsson, J., Swetzer, J. (1996) Renewable resources appropriation by cities, in Costanza R., Segura O. and Martinez-Alier J. (eds.) *Getting down to Earth*. Island Press, Washington, pp. 201–221.
Foster, C.D. (1960) Surplus criteria for investment. *Bulletin of Oxford University*, Institute of Economics and Statistics, 22 November.
French, S. (1986) *Decision theory*. Ellis Horwood, London.
Frey, H. (1999) *Designing the city. Towards a more sustainable urban form*. E & FN SPON, London.

Friedman, M. (1953) On the methodology of positive economics, in *Essays in positive economics*. University of Chicago Press, Chicago.

Funtowicz, S.O., Ravetz, J.R. (1990) *Uncertainty and quality in science for policy*. Kluwer, Dordrecht.

Funtowicz, S.O., Ravetz, J.R. (1991) A new scientific methodology for global environmental issues, in Costanza R. (ed.) *Ecological economics*. New York, Columbia, pp. 137–152.

Funtowicz, S.O., Ravetz, J.R. (1994) The worth of a songbird: ecological economics as a post-normal science. *Ecological Economics*, 10, pp.197–207.

Funtowicz, S.O., Munda, G., Paruccini, M. (1990) The aggregation of environmental data using multicriteria methods. *Environmetrics*, 1(4), pp. 353–368.

Funtowicz, S., Martinez-Alier, J., Munda, G., Ravetz, J. (1999) *Information tools for environmental policy under conditions of complexity*. European Environmental Agency, Experts' Corner, Environmental Issues Series, No. 9.

Funtowicz, S.O., Martinez-Alier, J., Munda, G., Ravetz, J. (2002) Multicriteria-based environmental policy, in Abaza H. and Baranzini A. (eds.) *Implementing sustainable development*. UNEP/Edward Elgar, Cheltenham, pp. 53–77.

Fusco Girard, L. (1986) The complex social value of the architectural heritage. *Icomos Information*, 1, pp. 19–22.

Fusco Girard, L., Nijkamp, P. (1997) *Le valutazioni per lo sviluppo sostenibile della citta' e del territorio*. Franco Angeli, Milan.

Gamboa Jiménez, G. (2003) *Evaluación multicriterio social de scenarios de futuro en la XI° region de Aysen, Chile*. Master Thesis, Doctoral Programme in Environmental Sciences, Universitat Autónoma de Barcelona.

Gamboa Jiménez, G., Munda, G. (2007) – The problem of wind–park location: a social multi-criteria evaluation framework. *Energy Policy*, 35(3), March, pp. 1564–1583.

Gonzalez, A.C. (2003) *Cost Benefit analysis and multicriteria analysis in developing countries: the case of air pollution in Mexico city*. Master Thesis, Doctoral Programme in Environmental Sciences, Universitat Autónoma de Barcelona.

Geach, P. (1967) Good and Evil, in P. Foot (ed.) *Theories of Ethics*. Oxford University Press, Oxford.

Georgescu Roegen, N. (1954) Choice, expectation and measurability. *Quarterly Journal of Economics*, 68, pp. 503–534.

Giampietro, M. (1994). Using hierarchy theory to explore the concept of sustainable development. *Futures*, 26(6), pp. 616–625.

Giampietro, M. (2003) *Multi-scale integrated analysis of agroecosystems*. CRC Press, New York.

Giampietro, M., Mayumi, K. (2000a) Multiple-scale integrated assessment of societal metabolism: introducing the approach. *Population and Environment*, 22(2), pp. 109–154.

Giampietro, M., Mayumi, K. (2000b) Multiple-scale integrated assessment of societal metabolism: integrating biophysical and economic representations across scales. *Population and Environment*, 22(2), pp. 155–210.

Giampietro, M., Mayumi, K., Munda, G. (2006) Integrated assessment and energy analysis: quality assurance in multi-criteria analyses of sustainability. *Energy*, 31(1), pp. 59–86.

Gomes, L.F.A.M., Lima, M.M.P.P. (1992) From modelling individual preferences to multicriteria ranking of discrete alternatives. A look at prospect theory and the additive utility model. *Foundations of Computing and Decision Sciences*, 17, pp. 172–184.

Gowdy, J. (1994) *Coevolutionary economics: the economy, society, and the environment*. Kluwer, Boston.

Gowdy, J.M., O'Hara, S. (1996) *Economic Theory for Environmentalists*.Saint Lucie Press, St. Lucie County.

Gravelle, H., Rees, R. (1992) *Microeconomics*. Longman, London.

Guimarães-Pereira, A., Rinaudo, J.D., Jeffrey, P., Blasuqes, J., Corral-Quintana, S.A., Courtois, N., Funtowicz, S., Petit, V. (2003) ICT tools to support public participation in water resources

governance and planning experiences from the design and testing of a multi-media platform. *Journal of Environmental Assessment Policy and Management*, 5(3), pp. 395–419.
Guimarães-Pereira, A., Corral-Quintana, S.A., Funtowicz, S. (2005) GOUVERNe: new trends in decision support for groundwater governance issues. *Environmental Modelling & Software*, 20, pp. 111–118.
Guimarães-Pereira, A., Guedes, S., Tognetti, S. (eds.) (2006) *Interfaces between science and society*. Greenleaf Publishing, Sheffield.
Guitouni, A., Martel, J.M. (1998) Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, 109, pp. 501–521.
Hanley, N., Spash, C. (1993) *Cost–benefit analysis and the environment*. Edward Elgar, Cheltenham.
Harker, P.T. (1989) The art and science of decision making: the analytic hierarchy process, in Golden B.L., Wasil E.A., and Harker P.T. (eds.) *The analytic hierarchy process*. Springer, Berlin, Heidelberg, New York, pp. 3–36.
Hartigan, J. (1975) *Clustering algorithms*. Wiley, New York.
Hayashi, K. (2000) Multicriteria analysis for agriculture resource management: a critical survey and future perspectives. *European Journal of Operational Research*, 122, pp. 486–500.
Hayek, F.A. (ed.) (1935) *Collectivist Economic Planning*. Reprinted Augustus M. Kelley, New York, 1970.
Helmers, R.L.C.H. (1979) *Project planning and income distribution*. Martinus Nijhoff, Boston.
Hersh, H.M., Caramazza, A. (1976) A fuzzy set approach to modifiers and vagueness in natural languages. *Journal of Experimental Psychology: General*, 105, pp. 254–276.
Hersh, H.M., Caramazza, A., Brownell, H.H. (1979) Effects of context on fuzzy membership functions, in Gupta M.M., Regade R.K. and Yager R.R. (eds.) *Advances in fuzzy set theory and applications*. Amsterdam, North Holland, pp. 389–408.
Hicks, J.R. (1939) The foundations of welfare economics. *Economic Journal* 49.
Hinloopen, E., Nijkamp, P. (1990) Qualitative multiple criteria choice analysis, the dominant regime method. *Quality and Quantity* 24, 37–56.
Hinloopen, E., Nijkamp, P., Rietveld, P. (1983) Qualitative discrete multiple criteria choice models in regional planning. *Regional Science and Urban Economics*, 13, pp. 77–102.
Horwarth, R., Norgaard, R.B. (1990) Intergenerational resource rights, efficiency and social optimality. *Land Economics*, 66, pp. 1–11.
Horwarth, R., Norgaard, R.B. (1992) Environmental valuation under sustainable development. *American Economic Review Papers and Proceedings*, 80, pp. 473–477.
ICLEI (1995) *International Council for Local Environmental Initiatives*.
Jacquet-Lagrèze, E. (1969) L'agrégation des opinions individuelles. *Informatiques et Sciences Humaines*, 4.
Janssen, R. (1992) *Multiobjective decision support for environmental management*. Kluwer, Dordrecht.
Janssen, R., Munda, G. (1999) Multi-criteria methods for quantitative, qualitative and fuzzy evaluation problems, in van den Bergh J. (ed.) *Handbook of environmental and resource economics*. Edward Elgar, Cheltenham, pp. 837–852.
Jasanoff, S. (1994) *Learning from disaster: risk management after Bhopal*. University of Pennsylavania Press, Philadelphia.
Kacprzyk, J., Roubens, M. (eds.) (1988) *Non-conventional preference relations in decision making*. Springer, Berlin, Heidelberg, New York.
Kahneman, D., Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, 47, pp. 263–291.
Kahneman, D., Slovic, P., Tversky, A. (1981) *Judgement under uncertainty – heuristics and biases*. Cambridge University Press, Cambridge.
Kaldor, N. (1939) Welfare comparison of economics and interpersonal comparisons of utility. *Economic Journal*, 49.
Kasemir, B., Gardner, M., Jäger, J., Jaeger, C. (eds.) (2003) *Public participation in sustainability science*. Cambridge University Press, Cambridge.

Keeney, R. (1992) *Value-focused thinking. A path to creative decision making*. Harvard University Press, Cambridge.
Keeney, R., Raiffa, H. (1976) *Decisions with multiple objectives: preferences and value trade-offs*. Wiley, New York.
Keller, W.J., Wansbeek, T. (1983) Multivariate methods for quantitative and qualitative data. *Journal of Econometrics*, 22, pp. 91–111.
Kelsey, D. (1986) Utility and the individual: an analysis of internal conflicts. *Social Choice and Welfare*, 3(2), pp. 77–87.
Kemeny, J. (1959) Mathematics without numbers. *Daedalus*, 88, pp. 571–591.
Kleijnen, J.P.C. (2001) Ethical issues in modelling: some reflections. *European Journal of Operational Research*, 130, pp. 223–230.
Knight, F.H. (1921) *Risk, uncertainty and profit*. Houghton Mifflin Company, The Riverside Press, Cambridge.
Köhler, G. (1978) *Choix multicritère et analyse algébrique des données ordinales*. Ph.D. Thesis, Université Scientifique et Médicale de Grenoble, France.
Koestler, A. (1969) Beyond atomism and holism: the concept of the holon, in Koestler A. and Smythies J.R. (eds.) *Beyond reductionism*. Hutchinson, London, pp. 192–232.
Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A. (1971) *Foundations of measurement, vol. 1, Additive and polynomial representations*. Academic Press, New York.
Kruskal, J.B. (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, pp. 1–27.
Krutila, J.V., Eckstein, O. (1958) *Multiple purpose river development*. Resource for the future, Baltimore.
Kuhn, T.S. (1962) *The structure of scientific revolutions*. University of Chicago Press, Chicago.
Laffont, J.J. (2000) *Incentives and political economy*. Oxford University Press, Oxford.
Laffont, J.J. (2002) Public economics yesterday, today and tomorrow. *Journal of Public Economics*, 86, pp. 327–334.
Lange, O., Taylor, F.M. (1938) *On the economic theory of socialism*. Benjamin E. (ed.) Lippincott, repr. McGraw-Hill, New York, 1966.
Leontief, W. (1947a) Introduction to a theory of the internal structure of functional relationships. *Econometrica*, 15, pp. 361-373.
Leontief, W. (1947b) A note on the interrelation of subsets independent variables of a continuous function with continuous first derivatives. *Bulletin of the American Mathematical Society*, 53, pp. 343-350.
Leung, Y. (1988) *Spatial analysis and planning under imprecision*. North Holland, Amsterdam.
Lichfield, N. (1964) Cost benefit analysis in plan evaluation. *Town Planning Review*, 35(2), pp. 160–169.
Lichfield, N. (1988) *Economics in urban conservation*. Cambridge University Press, Cambridge.
Lichfield, N. (1993) Problems of valuation discussed, in Banister D. and Button K. (eds.) *Transport, the environment and sustainable development*. E & FN Spon, London, pp. 205–211.
Lockwood, M. (1997) Integrated value theory for natural areas. *Ecological Economics*, 20(1), pp. 83–93.
Luce, R.D. (1956) Semiorders and a theory of utility discrimination. *Econometrica*, 24, pp.178–191.
Luce, R.D., Raiffa, H. (1989) *Games and decisions*. Dover, New York.
Machina, M. (1987) Choice under uncertainty: problems solved and unsolved. *Journal of Economic Perspectives*, 1(1), pp. 121–154.
Markowitz, H.M. (1989) *Mean–variance analysis in portfolio choice and capital markets*. Basil-Blackwell, Oxford.
Martel, J.M., Zaras, K. (1995) Stochastic dominance in multicriteria analysis under risk. *Theory and Decision*, 35(1), pp. 31–49.
Martí, N. (2001) *Processos de decisió i instrumentalització de l'avaluació d'actuacions en el territori. Una proposta metodòlogica d'avaluació integrada a l'entorn del Parc Nacional*

d'Aigüestortes i estany de Sant Maurici: el cas de l'estudi DIAFANIS. Master Dissertation, Doctoral Programme in Environmental Science, Universitat Autónoma de Barcelona, Spain.

Martí, N. (2005) *La multidimensionalidad de los sistemas locales de alimentación en los Andes peruanos: los chalayplasa del Valle de Lares (Cusco)*. Ph.D. Dissertation, Doctoral Programme in Environmental Science, Universitat Autónoma de Barcelona, Spain.

Martinez-Alier, J., (2002) *The environmentalism of the poor*. Edward Elgar, Cheltenham.

Martinez-Alier, J., O'Connor, M. (1996) Ecological and economic distribution conflicts, in Costanza, R., Segura O. and Martinez-Alier J. (eds.) *Getting down to Earth: practical applications of ecological economics*. Island Press/ISEE, Washington D.C.

Martinez-Alier, J., Schluepmann, K. (1987) *Ecological economics*. Blackwell, Oxford.

Martinez-Alier, J., Munda, G., O'Neill, J. (1998) Weak comparability of values as a foundation for ecological economics. *Ecological Economics*, 26, pp. 277–286.

Matarazzo, B. (1988) Preference ranking global frequencies in multicriteria analysis (PRAGMA). *European Journal of Operational Research*, 36, pp. 36–49.

Matarazzo, B., Munda, G. (2001) New approaches for the comparison of L–R fuzzy numbers: a theoretical and operational analysis. *Fuzzy Sets and Systems*, 118(3), pp. 407–418.

Maystre, L., Pictet, J., Simos, J. (1994) *Méthodes multicritères ELECTRE – description, conseils pratiques et cas d'application à la gestion environnementale*. Presses Polytechniques et Universitaires Romandes, Lausanne.

Melyn, W., Moesen, W. (1991) Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available. *Public Economics Research Paper* 17, CES, KU Leuven.

Mishan, E.J. (1971) *Cost–Benefit Analysis*. London: Allen and Unwin.

Miyamoto, S. (1990) *Fuzzy sets in information retrieval and cluster analysis*. Kluwer, Dordrecht.

Moreno-Jiménez, J.M., Aguaron, J., Escobar, T., Turon, A. (1999) Multicriteria procedural rationality on SISDEMA. *European Journal of Operational Research*, 119(2), pp. 388–403.

Moser, P.K. (ed.) (1990) *Rationality in action*. Cambridge University Press, Cambridge.

Moulin, H. (1981) The proportional veto principle. *Review of Economic Studies*, 48, pp. 407–416.

Moulin, H. (1985) From social welfare orderings to acyclic aggregation of preferences. *Mathematical Social Sciences*, 9, pp. 1–17.

Moulin, H. (1988) *Axioms of co-operative decision making*. Econometric Society Monographs, Cambridge University Press, Cambridge.

Mueller, D. (1978) Voting by veto. *Journal of Public Economics*, 10, pp. 57–75.

Munda, G. (1993) Multiple-criteria decision aid: some epistemological considerations. *Journal of Multi-Criteria Decision Analysis*, 2, pp. 41–55.

Munda, G. (1995) *Multicriteria evaluation in a fuzzy environment*. Physica-Verlag, Contributions to Economics Series, Heidelberg.

Munda, G. (1996) Cost–benefit analysis in integrated environmental assessment: some methodological issues. *Ecological Economics*, 19(2), pp. 157– 168.

Munda, G. (1997a) Environmental economics, ecological economics and the concept of sustainable development. *Environmental Values*, 6(2), pp. 213–233.

Munda, G. (1997b) Multicriteria evaluation as a multidimensional approach to welfare measurement, in van den Bergh J. and van der Straaten J. (eds.) *Economy and ecosystems in change: analytical and historical approaches*. Edward Elgar, Cheltenham, pp. 96–115.

Munda, G. (1998) A theoretical inquiry on the axiomatic consistency of distributional weights used in cost–benefit analysis. *International Journal of Applied Economics and Econometrics*, 7(1), pp. 41–47.

Munda, G. (2004) "Social multi-criteria evaluation (SMCE)": methodological foundations and operational consequences. *European Journal of Operational Research*, 158(3), pp. 662–677.

Munda, G. (2005a) Multi-criteria Ddcision analysis and sustainable development, in Figueira J., Greco S. and Ehrgott M. (eds.) *Multiple-criteria decision analysis. State of the art surveys*. Springer International Series in Operations Research and Management Science, New York, pp. 953–986.

Munda, G. (2005b)"Measuring sustainability": a multi-criterion framework. *Environment, Development and Sustainability*, 7(1), pp. 117–134.
Munda, G. (2006a) Social multi-criteria evaluation for urban sustainability policies. Land Use Policy, 23(1), January, pp. 86–94.
Munda, G. (2007) Solving the Discrete Multi-Criterion Problem: An Overview of Lessons Learned from Voting Theory. Unpublished manuscript.
Munda, G., Nardo, M. (2003) Mathematical modelling of composite indicators for ranking countries. *Proceedings of the First OECD/JRC Workshop on Composite Indicators of Country Performance*, JRC, Ispra, available at *http://webfarm.jrc.cec.eu.int/uasa/events/oecd_12may03/Background/*.
Munda, G., Nardo, M. (2005) Constructing Consistent Composite Indicators: the Issue of Weights. EUR 21834 EN, Joint Research Centre, Ispra.
Munda, G., Nardo, M. (2007) Non-compensatory/non-linear composite indicators for ranking countries: a defensible setting. Forthcoming in *Applied Economics*.
Munda, G., Nijkamp, P., Rietveld, P. (1994a) Qualitative multicriteria evaluation for environmental management. *Ecological Economics*, 10, 97–112.
Munda, G., Nijkamp, P., Rietveld, P. (1994b) Fuzzy multigroup conflict resolution for environmental management, in J. Weiss (ed.) *The economics of project appraisal and the environment*. Edward Elgar, Cheltenham, pp. 161–183.
Munda, G., Nijkamp, P., Rietveld, P. (1995) Qualitative multi-criteria methods for fuzzy evaluation problems. *European Journal of Operational Research*, 82, pp. 79–97.
Munda, G., Paruccini, M., Rossi, G. (1998) Multicriteria evaluation methods in renewable resource management: the case of integrated water management under drought conditions, in Beinat E. and Nijkamp P. (eds.) *Multicriteria evaluation in land-use management: methodologies and case studies*.Kluwer, Dordrecht, pp. 79–94.
Munda, G., Gamboa, G., Russi, D., Garmendia, E. (2005) *Social Multi-Criteria Evaluation of Renewable Energy Sources: Two Real World Catalan Examples*. Report for the European Union research project "Development and Application of a Multicriteria Decision Analysis software tool for Renewable Energy sources (MCDA-RES)", Contract NNE5-2001-273.
Musgrave, A. (1981) Unreal assumptions in economic theory: the F-twist untwisted. *Kyklos*, 34, pp. 377–387.
Musu, I., Siniscalco, D. (1996) *National accounts and the environment*. Kluwer, Dordrecht.
Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E. (2005) *Handbook on constructing composite indicators: methodology and user guide*. OECD Statistics Working Paper, Paris.
Neurath, O. (1919) *Through war economy to economy in kind*, in Neurath, 1973.
Neurath, O. (1925) *Wirtschaftsplan und naturalrechnung*. Laub, Berlin (in German).
Neurath, O. (1928) *Personal life and class struggle*, in Neurath, 1973.
Neurath, O. (1973) *Empiricism and sociology*. Reidel, Dordrecht.
Nijkamp, P. (1979) *Multidimensional spatial data and decision analysis*. Wiley, New York.
Nijkamp, P., Perrels, A. (1994) *Sustainable cities in Europe*. Earthscan, London.
Nijkamp, P., Rietveld, P., Voogd, H. (1990) *Multicriteria evaluation in physical planning*. Amsterdam: North-Holland.
Norese, M.F. (1991) A multiple criteria approach to complex situations, in Jackson et al. (eds.) *Systems thinking in Europe*. Plenum Press, New York, pp. 361–369.
Norgaard, R.B. (1994) *Development betrayed*. Routledge, London.
Norwich, A.M., Turksen, I.B. (1982) The fundamental measurement of fuzziness, in Yager R.R. (ed.) *Fuzzy set and possibility theory*. Pergamon, New York.
Norwich, A.M., Turksen, I.B. (1984) A model for the measurement of membership and the consequences of its empirical implementation. *Fuzzy Sets and Systems*, 12, pp. 1–25.
Novotny, P. (1994) Popular epidemiology and the struggle for community health: alternative perspectives from the environmental justice movement. *Capitalism, Nature, Socialism*, 5(2), pp. 29–42.
Oates, W.E., Baumol, J. (1996) *The Economics of environmental regulation*. Edward Elgar, Cheltenham.
O'Connor, M., Faucheux, S., Froger, G., Funtowicz, S.O., Munda, G. (1996) Emergent complexity and procedural rationality: post-normal science for sustainability, in Costanza R., Segura O.

and Martinez-Alier J. (eds.) *Getting down to earth: practical applications of ecological economics*. Island Press/ISEE, Washington D.C., pp. 223–248.

OECD (2003) *Composite indicators of country performance: a critical assessment*. DST/IND (2003) 5, Paris.

Olson, M. (1982) *The rise and decline of nations: economic growth, stagflation and social rigidities*. Yale University Press, New Haven, Connecticut.

O'Neill, J. (1993) *Ecology, policy and politics*. Routledge, London.

O'Neill, J. (1995a) In partial praise of a positivist. *Radical Philosophy*, 74, pp. 29–38.

O'Neill, J. (1995b) Policy, economy, neutrality. *Political Studies*, XLIII, pp. 414–431.

O'Neill, J. (1996a) Value pluralism, incommensurability and institutions, in J. Foster (ed.) *Valuing nature? economics, ethics and environment*. Routledge, London.

O'Neill, J. (1996b) Who won the socialist calculation debate? *History of Political Thought*, 17, pp. 431–442.

Ostanello, A. (1990) Action evaluation and action structuring: different decision aid situations reviewed through two actual cases, in Bana e Costa C.A. (ed.) *Readings in multiple criteria decision aid*. Springer, Berlin, Heidelberg, New York, pp. 36–57.

Ozkaynak Ortakoyluoglu, B. (2005) *Indicators and scenarios for urban development and sustainability. A participatory case study of Yalova*. Ph.D. Thesis, Doctoral Programme in Environmental Sciences, Universitat Autónoma de Barcelona, Ciencias Económicas y Empresariales.

Ozturk, M., Tsoukias, A., Vincke, Ph. (2005) Preference modelling, in Figueira J., Greco S. and Ehrgott M. (eds.) *Multiple-criteria decision analysis. State of the art surveys*. Springer International Series in Operations Research and Management Science, New York, pp. 27–71.

Paelinck, J.H.P. (1978) Qualiflex, a flexible multiple criteria method. *Economic Letters*, 3, pp. 193–197.

Pearce, D.W., Atkinson, G.D. (1993) Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability. *Ecological Economics*, 8, pp. 103–108.

Pearce, D.W., Nash, C.A. (1981) *The social appraisal of projects*. MacMillan, London.

Pearce, D., Hamilton, G., Atkinson, G.D. (1996) Measuring sustainable development: progress on indicators. *Environment and Development Economics*, 1, pp. 85–101.

Perny, P., Roy, B. (1992) The use of fuzzy outranking relations in preference modeling. *Fuzzy Sets and Systems*, 49, pp. 33–53.

Podinovskii, V.V. (1994) Criteria importance theory. *Mathematical Social Sciences*, 27, pp. 237-252.

Poincaré, H. (1935) *La valeur de la science*. Flammarion.

Ragade, R.K., Gupta, M.M. (1977) Fuzzy set theory: Introduction, in Gupta M.M., Saridis G.N. and Gaines B.R. (eds.) *Fuzzy automata and decision processes*. North-Holland, Amsterdam, pp. 105–131.

Rajan, S.R. (2002) Disaster, development and governance: reflections on the 'lessons' from Bhopal. *Environmental Values*, 11(3), 369–394.

Rauschmayer, F. (2000) Ethics of multicriteria analysis. *International Journal of Sustainable Development*, 3(1), pp. 16–25.

Ray, A. (1984) *Cost–benefit analysis*. The John Hopkins University Press (published for The World Bank), Baltimore.

Ray, P. (1973) Independence of irrelevant alternatives. *Econometrica*, 41(5), pp. 987–991.

Rietveld, P. (1989) Using ordinal information in decision making under uncertainty. *Systems Analysis, Modeling, Simulation*, 6, pp. 659–672.

Ringius, L., Asbjørn, T., Holtsmark, B. (1998) Can multi-criteria rules fairly distribute climate burdens? OECD results from three burden sharing rules. *Energy Policy*, 26(10), 777–793.

Rìos Insua, D. (1990) *Sensitivity analysis in multi-objective decision making*. Springer, Berlin, Heidelberg, New York.

Roberts, F.S. (1979) *Measurement theory with applications to decision making, utility and the social sciences.* Addison-Wesley, London.

Romero, C., Rehman, T. (1989) *Multiple criteria analysis for agricultural decisions*. Elsevier, Amsterdam.

Rosen, R. (1977) Complexity as a system property. *International Journal of General Systems*, 3, 227-232

Rothschild, M., Stiglitz, J.E. (1970) Increasing risk and risk: 1. A definition. *Journal of Economic Theory*, 2, pp. 225–243.

Rothschild, M., Stiglitz, J.E. (1971) Increasing risk and risk: 2. Its consequences. *Journal of Economic Theory*, 3, pp. 66–84.

Roubens, M., Vincke, Ph. (1985) *Preference modeling.* Springer, Berlin, Heidelberg, New York.

Roy, B. (1978) ELECTRE III: un algorithme de classement fondé sur une répresentation floue des prèferences en prèsence de critères multiples. *Cahiers du Centre d' Etudes de Recherche Opèrationnelle* Paris, 20(1), pp. 3–24.

Roy, B. (1985) *Méthodologie multicritere d' aide à la decision*. Economica, Paris.

Roy, B. (1990) Decision aid and decision making, in Bana e Costa C.A. (ed.) *Readings in multiple criteria decision aid*. Springer, Berlin, Heidelberg, New York, pp. 17–35.

Roy, B. (1996)*Multicriteria methodology for decision analysis*. Kluwer, Dordrecht.

Roy, B., Bertier, P. (1973) La methode ELECTRE II. Une application ou media planning, in Ross M. (ed.) *Operational research ´72*, North Holland, Amsterdam, pp. 291–302.

Roy, B., Damart, S. (2002) L'analyse Coûts-Avantages, outil de concertation et de légitimation? *Metropolis*, 108/109, pp. 7–16.

Russi, D. (2004) *Social multicriteria evaluation of rural electrification*. Master Thesis, Doctoral Programme in Environmental Sciences, Universitat Autónoma de Barcelona.

Saari, D.G. (1989) A dictionary for voting paradoxes. *Journal of Economic Theory*, 48, pp. 443–475.

Saari, D.G. (2000) Mathematical structure of voting paradoxes. 1. Pairwise votes, *Economic Theory*, 15, pp. 1–53.

Saari, D.G. (2006) Which is better: the Condorcet or Borda winner? *Social Choice and Welfare*, 26, pp. 107–129.

Saari, D.G., Merlin, V.R. (2000) A geometric examination of Kemeny's rule. *Social Choice and Welfare*, 17, pp. 403–438.

Saaty, T.L. (1980) *The analytic hierarchy process*. McGraw Hill, New York.

Saisana, M., Tarantola, S. (2002) *State-of-the-art report on current methodologies and practices for composite indicator development*. EUR 20408 EN Report, European Commission, JRC, Ispra, Italy.

Saisana, M., Tarantola, S., Saltelli, A. (2005) Uncertainty and sensitivity techniques as tools for the analysis and validation of composite indicators. *Journal of the Royal Statistical Society A*, 168(2), pp. 307–323.

Salminen, P., Hokkanen, J., Lahdelma, R. (1998) Comparing multicriteria methods in the context of environmental problems. *European Journal of Operational Research*, 104, pp. 485–496.

Saltelli, A., Chan, K., Scott, M. (2000) *Sensitivity analysis*. Probability and Statistics series, John Wiley & Sons, New York.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. (2004) *Sensitivity analysis in practice. A guide to assessing scientific models*. Wiley, New York.

Savage, L. (1954) *The foundations of statistics*. Second edition, Wiley, New York.

Schmidt-Bleek, F. (1994) *Wieviel umwelt braucht der mensch? MIPS Das Mass für ökologisches wirtschaften*. Birkh user, Berlin.

Simon, H.A. (1976) From substantive to procedural rationality, in Latsis J.S. (Ed.) *Methods and appraisal in economics*. Cambridge University Press, Cambridge.

Simon, H.A. (1983) *Reason in human affairs*. Stanford University Press, Standord.

Simpson, P. (1969) – On defining areas of voter choice: Professor Tullock on stable voting, *Quarterly Journal of Economics*, 83, pp. 478–490.

Sittaro, F. (2006) – *Participative multicriteria planning in community development. A case study: the Cuyabeno wildlife reserve, ecuador*, Ph.D. Dissertation, Venice International University, International Ph.D. Programme in "Analysis and Governance for Sustainable Development".

Spash, C., Hanley, N. (1995) Preferences, information, and biodiversity preservation. *Ecological Economics*, 12, pp. 191–208.

Spronk, J. (1981) *Interactive multiple goal programming for capital budgeting and financial planning*. Martinus Nijhoff, Boston.

Stagl, S. (2006) – Multicriteria Evaluation and Public Participation: The Case of UK Energy Policy. *Land Use Policy*, 23(1), 53–62.

Steuer, R.E. (1986) – *Multiple criteria optimization*. Wiley, New York.

Stewart, T.J., Joubert, A. (1998) Conflicts between conservation goals and land use for exotic forest plantations in South Africa, in Beinat E. and Nijkamp P. (eds.) *Multicriteria evaluation in land-use management: methodologies and case studies*. Kluwer, Dordrecht, pp. 17–31.

Stiglitz, J.E. (2002) New perspectives on public finance: recent achievements and future challenges. *Journal of Public Economics*, 86, pp. 341–360.

Todhunter, I. (1949) *A history of mathematical theory of probability*. Chelsea, New York.

Truchon, M. (1995) Voting games and acyclic collective choice rules. *Mathematical Social Sciences*, 25, pp. 165–179.

Truchon, M. (1998a) *Figure skating and the theory of social choice*. Cahier 9814 du Centre de Recherché en Economie et Finance Appliqués (CREFA).

Truchon, M. (1998b) *An extension of the Condorcet criterion and Kemeny orders*. Cahier 9813 du Centre de Recherche en Economie et Finance Appliquées (CREFA).

Turner, R.K., Pearce, D.W., Bateman, I. (1994) *Environmental economics: an elementary introduction*. Harvester Wheatsheaf, London.

Tversky, A. (1969) Intransitivity of preference. *Psychological Review*, 79, pp. 31–48.

Ülengin, B., Ülengin, F., Güvenç, Ü (2001) A multidimensional approach to urban quality of life: the case of Istanbul. *European Journal of Operational Research*, 130, pp. 361–374.

Urban Indicators Programme, *http://www.urbanobservatory.org/indicators/database*

Vanderpooten, D. (1989) The interactive approach in MCDA: a technical framework and some basic conceptions. In *Models and methods in multiple criteria decision making*, Colson G. and De Bruyn Chr. (eds.) Pergamon Press, Oxford, pp. 1213–1220.

Vansnick, J.C. (1986) On the problem of weights in multiple criteria decision making (the non-compensatory approach). *European Journal of Operational Research*, 24, pp. 288–294.

Vansnick, J.C. (1990) Measurement theory and decision aid, in Bana e Costa C.A. (ed.) *Readings in multiple criteria decision aid*. Springer, Berlin, Heidelberg, New York, pp. 81–100.

Vargas Isaza, O.L. (2004) *La evaluación multicriterio social y su potencial en la gestión forestal de Colombia*. Ph.D. Thesis, Doctoral Programme in Environmental Sciences, Universitat Autónoma de Barcelona.

Vidu, L. (2002) Majority cycles in a multi-dimensional setting. *Economic Theory*, 20, pp. 373–386.

Vincke, Ph. (1985) Multiattribute utility theory as a basic approach, in Fandel G., Matarazzo B. and Spronk J. (eds.) *Multiple criteria decision methods and applications*. Springer, Berlin, Heidelberg, New York, pp. 27–40

Vincke, Ph. (1992) *Multicriteria decision aid*. Wiley, New York.

Vincke, Ph. (1994) Recent progresses in multicriteria decision-aid. *Rivista di Matematica per le scienze Economiche e Sociali*, 2, pp. 21–32.

Vitousek, P., Ehrlich, P., Ehrlich, A., Matson, P. (1986) Human appropriation of the products of photosynthesis. *Bioscience*, 34(6), 368–373.

Voogd, H. (1983) *Multicriteria evaluation for urban and regional planning*, London, Pion.

Wackernagel, M., Rees, W.E. (1995) *Our ecological footprint: reducing human impact on the earth*. Gabriola Island, BC and Philadelphia, PA: New Society Publishers.

Wallsten, T.S., Budescu, D.V., Rapoport, A., Zwick, R., Forsyth, B. (1986) Measuring the vague meaning of probability terms. *Journal of Experimental Psychology: General*, 115, pp. 348–365.

Weber, J. (2002) How many voters are needed for paradoxes? *Economic Theory*, 20, pp. 341–355.

Weisbrod, B. (1968) Income redistribution effects and benefits–cost analysis, in Chase S.B. (ed.) *Problems in public expenditure analysis*. Brookings Institution, Washington.

Wierzbicki, A.P. (1982) A mathematical basis for satisficing decision making. *Mathematical Modelling*, 3, pp. 391–405.

Williams, B. (1972) *Morality*. Cambridge University Press, Cambridge.

van Winden, F. (1999) On the economic theory of interest groups: towards a group frame of reference in political economics. *Public Choice*, 100, pp. 1–29.

Winkler, R.L., Hays, W.L. (1975) *Statistics*. Holt, Rinehart and Winston, New York.

World Commission on Environment and Development (1987) *Our common future*. Oxford University Press, Oxford.

Wynne, B. (1992) Uncertainty and environmental learning: reconceiving science in the preventive paradigm. *Global Environmental Change*, 2 (June), pp. 111–127.

Young, H.P. (1978) On permutations and permutation polytopes. *Mathematical Programming Study*, 8, pp. 128-140.

Young, H.P. (1986) Optimal ranking and choice from pair-wise comparisons, in Grofman B. and Owen G. (eds.) *Information pooling and group decision-making*, Greenwuich, CT: JAL.

Young, H.P. (1988) Condorcet's theory of voting. *American Political Science Review*, 82(4), pp. 1231–1244.

Young, H.P. (1995) Optimal voting rules. *Journal of Economic Perspectives*, 9, pp. 51–64.

Young, H.P., Levenglick, A. (1978) A consistent extension of Condorcet's election principle. *SIAM Journal of Applied Mathematics*, 35, pp. 285–300.

Yu, P.L. (1973) A class of solutions for group decision problems. *Management Science*, 19(8), pp. 936–946.

Yu, P.L. (1985) *Multi-criteria decision making: concepts, techniques and extensions*. Plenum Press, New York.

Yu, P.L. (1990) *Forming winning strategies: an integrated theory of habitual domains*. Springer, Berlin, Heidelberg, New York.

Yusuf, J.A., El Serafy, S., Lutz, E. (1989) *Environmental accounting for sustainable development*. A UNEP World Bank Symposium, Washington D.C.

Zadeh, L.A. (1965) Fuzzy sets. *Information and Control*, 8, pp. 338–353.

Zeleny, M. (1974) A concept of compromise solutions and the method of the displaced ideal. *Computers and Operations Research*, vol. 1, N. 4, pp. 479–496.

Zeleny, M. (1982) *Multiple criteria decision making*. McGraw Hill, New York.

Zimmermann, H.J., Zysno, P. (1983) Decisions and evaluations by hierarchical aggregation of information. *Fuzzy Sets and Systems*, 10, pp. 243–260.

Index

M

N

O

P

R

S