Project No 518294 SES6

CASES

Cost Assessment of Sustainable Energy Systems

Instrument: Co-ordination Action

Thematic Priority: Sustainable Energy Systems

DELIVERABLE No D.11.1
[Policy Assessment Guideline]

Due date of deliverable: 31st January 2007
Actual submission date: 15th March 2007

Start date of project: 1st April 2006 Duration: 30 months

Organisation name of lead contractor for this deliverable:

National Technical University of Athens
Danae Diakoulaki and Christos Tourkolias

Revision: FEEM

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

TABLE OF CONTENTS

1. Introduction

Policy and project appraisal usually involves the consideration of alternative courses of action along multiple decision aspects, such as their cost of implementation and their positive and negative effects on the social and natural environment. In order to provide a basis for selecting among the alternatives, all these effects need to be expressed in comparable 'value' terms. Although the term 'value' has different meanings in different disciplines, a simple definition is that value shows *'the difference a good or service makes'* while a more comprehensive one refers to *'the contribution of a good or service to specific goals, objectives or conditions'* (Costanza, 2000). For non-traded goods, such as environmental quality or human health, 'values' do not exist in the marketplace and have to be elicited or uncovered through appropriate value elicitation techniques.

One can distinguish two main valuation approaches, which in turn specify two different assessment frameworks applicable in environmental policy making.

The first valuation approach is based on neo-classical economics and provides the value information needed in Cost-Benefit Analysis (CBA). The underlying assumption is that the market, as guided by individual preferences is the best indicator for environmental values, while social preferences resulting as the sum of individual preferences secure optimal allocation of resources and social welfare. Money is the appropriate vehicle to express these preferences in order to make them commensurate with the values of other goods and to comply with the rules of the market mechanism. Therefore, in the case of non traded environmental and social goods, where stated preferences are missing, one should strive to directly extract social preferences by survey-based techniques such as the Contingent Valuation method (CVM), or to reveal these preferences by analyzing human behavior in relevant or surrogate markets. The values obtained from these methods can be straightforwardly used in CBA by translating each effect in monetary terms. Thus, in CBA policy makers can make their choice on the basis of the total net cost or benefit of the considered alternative actions.

The second valuation approach is part of the Multiple Criteria Decision Analysis (MCDA) framework which is increasingly applied during the last 30 years for aiding decisions involving several actions compared along multiple competing criteria. The underlying assumption is that values do not a-priori exist in human mind –especially for unfamiliar environmental goods- and that there is a need to apply a systematic value elicitation procedure to facilitate stakeholders construct and articulate their preferences. Value elicitation and the subsequent aggregation phase takes usually place with the participation of all parties involved or interested in the decision to be taken. A diversity of multicriteria methods and weighting techniques are in use today for translating the performances of the examined actions, measured in physical or qualitative scales, into a global value or preference index for each alternative reflecting human judgment about the underlying scores.

Following a short description of CBA and MCDA, the present document includes basic guidelines for the proper implementation of the two assessment frameworks and a comparative examination of their strengths and weaknesses.

2. Cost-Benefit Analysis

2.1 General overview

Cost-Benefit Analysis (CBA) has long been used as a powerful appraisal technique for investments in the developing world where markets are seriously distorted and prices do not provide a satisfactory basis for assessing their overall impact on society. However, during the last thirty years, CBA is also widely applied in developed countries for the assessment of public projects and policies, because of the growing concern for distorted market prices and the existence of externalities, both indicating a divergence between private and social costs and an inefficient allocation of resources. In particular, due to an increasing environmental awareness, CBA is now routinely used in environmental policy assessment in order to secure that policy choices are as cost-effective as possible (Schulze, 1994; Virani and Graham, 1998).

CBA aims at aggregating all costs and benefits associated with each policy option, the former defined as reductions and the latter as increases in human wellbeing. Hence, any policy impact affecting social wellbeing should be included in CBA.

Central in the theoretical foundation of CBA is that costs and benefits reflect the preferences of individuals which are measured as Willingness To Pay (WTP) for benefits and Willingness To Accept compensation (WTA) for costs. It is furthermore assumed that individual preferences can be aggregated so that social benefit is simply the sum of all individuals' benefits and social cost is the sum of all individuals' costs (Pearce et al. 2006).

According to the above aggregation rule 'society' is defined as the sum of all individuals without strict geographical boundaries. Although, the national scale is typically used in many practical applications of CBA, especially those concerned with project assessment or national level policies, it is increasingly recognised that in several circumstances, policies affect neighboring nations (e.g. transboundary pollution) or the entire planet (e.g. climate change). Thus, aggregating costs or benefits across different social and income groups implies simplifying assumptions that should be kept in mind in real life policy applications.

For non-traded or intangible goods, for which WTP/WTA measures do not exist in the marketplace, a range of approaches are in use today for deriving such estimates to be integrated in CBA. However, in many cases significant environmental impacts are not easily amenable to defensible monetary valuation and are therefore omitted from CBA.

Given that policies affect societies over a long time horizon, costs and benefits should be also aggregated over time. Here also, there are no hard rules of how far into the future one has to go. Summing costs and benefits over time requires that 'time preferences' of individuals have also to be accounted for in a way that current choices do not disregard the interests of future generations. This practically consists in selecting an appropriate discount rate (Horowitz, 1996).

For a project or policy to be selected benefits must exceed costs, while alternative options are ranked according to their net benefits. The most general formulation of this decision rule is the following:

$$\left[\sum_{i,t} B_{i,t} \cdot (1+r)^{-t} - \sum_{j,t} C_{j,t} \cdot (1+r)^{-t}\right] > 0$$

where $B_{i,t}$ and $C_{j,t}$ are the benefit i and the cost j occurring in time t, while r denotes the discount rate used to calculate the present value of all costs and benefits. The Benefit/Cost (B/C) ratio or the internal rate of return (IRR) that is compared with the usually applicable social discount rates are also commonly used as evaluation indices depending on the specific decision context, such as the existence of budget constraints or the consideration of mutually exclusive options (Pearce et al. 2006).

Long-term policies especially those targeting at environmental benefits involve a high level of uncertainty calling for the use of appropriate uncertainty treatment techniques.

2.2 Cost-Effectiveness Analysis

Cost-Effectiveness Analysis (CEA) aims at determining the least cost option achieving a predefined policy goal. For this reason it is principally implemented in cases where the assessment refers to different actions intended to achieve some predefined standards set by legislation (e.g. health standards) or related with other commitments (e.g. meeting of Kyoto targets).

The aggregation in CEA refers only to the costs imposed by the alternative courses of action which are compared to a single benefit expected to result from policy application. This means that there is no need to assign a monetary value on the desired beneficial effect, so that difficulties and biases associated with monetary valuation can be avoided. The attribute chosen to represent the expected benefit is expressed in physical terms or, in the case of more complicated effects of a composite indicator defined by experts. The resulting ratio relating the effect with the corresponding cost is a meaningful index offering important guidance on which is the most cost-effective action to achieve a policy target. But it does not provide any information if this action is worth undertaking or not, i.e. if the resulting benefits are greater than the respective costs (Snell, 1997; Pearce et al., 2006).

In the case of policies targeting at a multiplicity of environmental benefits, CBA is the recommended appraisal tool.

3. Multiple Criteria Decision Analysis

3.1 General overview

Multiple Criteria Decision Analysis (MCDA) is aiming at providing a formal approach helping decision makers to effectively handle complex decision situations in which the level of conflict between criteria is such that intuitive solutions can not be satisfactory. MCDA is particularly suited if, in addition to the conflict between criteria, there is significant ambiguity in measuring performances and/or in articulating preferences. Finally, MCDA can help in resolving disagreement if stakeholders have different views on the relative importance of the considered criteria. It is important to stress, that MCDA is not a tool providing the right solution in a decision problem, simply because it is admitted that such a universally accepted right solution does not exist. Instead, it is an aid to decision making that helps decision makers organize the available information, think on the consequences, explore their own wishes and tolerances and minimize the possibility for a post-decision disappointment (Belton and Stewart, 2002).

The strength of MCDA is better reflected in problems of a strategic nature encountered in many different fields of economic activity. These problems refer to non-repeated decision situations with a medium-to long-term planning horizon and usually have more serious and often non-reversible consequences. Similar types of problems are the design of policies and the formulation of action plans, technological choices, siting decisions, project evaluation and approval etc. In addition, this type of problems are more complicated as including several interrelated systems and sub-systems, while uncertainties are also higher because of the variability of these systems and the limited scientific knowledge about the driving forces behind relevant changes.

The decision type sought when implementing a MCDA method can be one of the following:

- ***Choice:*** selecting one action among two or several alternative options.
- ***Ranking:*** placing alternatives in a preference order for selecting those ranked at the higher places.
- ***Sorting:*** grouping alternatives into broad hierarchical categories, each one including a number of non-distinctive alternatives.
- ***Portfolio:*** identifying the best combination of alternative actions by taking into account not only the alternatives' individual characteristics but also their interactions and synergies.

Finally, a completely different type of decision is required in cases where alternatives are not a priori defined but result as combinations of continuous or binary decision variables by optimizing a given system with respect to specific objectives and by taking into account a number of operational or policy constraints. Both, constraints and objectives are expressed as functions –usually linear- of the considered decision variables and the resulting problem is solved by means of Multiple Objective Programming (MOP) models. The outcome of the optimization procedure is a set of non-dominated solutions among which the decision maker(s) is called to make the final choice.

The analysis hereafter is restricted to a choice or ranking type of problem dealing with discrete alternatives which are a-priori defined and evaluated with respect to a number of evaluation criteria.

3.2 Classification of MCDA methods

MCDA methods are classified in two broad categories that are shortly described in the following paragraphs. Among other methodological and computational differences, these two categories use a completely different approach in preference elicitation and in the aggregation of criteria information. The first category (MAVT methods) use a compensatory approach assuming that it is possible to completely offset a disadvantage on one evaluation dimension by an advantage in another dimension, while the second one (outranking methods) follow a non-compensatory approach denying such possibility.

3.2.1 Multi-Attribute Value Theory methods

Multi-Attribute Value Theory (MAVT) methods associate with each alternative option a unique number (‘value’) providing an overall evaluation of this option if all criteria are taken into account according to the decision maker’s preferences. MAUT methods based on Utility Theory constitute an extension of MAVT methods integrating uncertainty in the measurement of performances of alternatives (Keeney and Raiffa, 1976). The brief description hereafter will be restricted to the more general notions of value theory assuming that performances are deterministically defined.

The starting point is the definition of partial value (or utility) functions in each single criterion. A value function translates the performances of the alternatives in this criterion into a [0-1] interval, with 0 denoting the worst and 1 the best performance, respectively. Worst and best performances refer to either the considered set of alternatives (local scale) or to potentially achievable scores (global scale).

It is important to note that for the construction of partial value functions human preferences are not necessarily related in a linear way with the performances measured on a ‘natural’ or ‘objective’ scale. Furthermore, in the case of a criterion where such natural measurement scales do not exist, it is the decision maker who has to construct a scale by assigning values to the examined alternatives according to his/her own view about their non quantified scores. Hence, each alternative a is associated with a value $v_i(a)$, translating its performance in criterion i in terms of the particular decision maker’s preference system. The basic property of these partial value functions is that alternative a is strictly preferred to alternative b if $v_i(a)>v_i(b)$, while indifference between the two alternatives holds only if $v_i(a)=v_i(b)$.

The transition from partial to global value functions (taking into account the whole set of criteria) implies the use of an aggregation formula together with the inter-criterion preferences provided by the decision maker. The simplest and most commonly used aggregation model is the additive one:

$$V(a) = \sum_i w_i \bullet v_i(a) \qquad \text{with } \Sigma w_i = 1$$

V(a) is the total value associated with each alternative *a*, and w_i is the weight reflecting the relative importance attached to each criterion *i* by the decision maker.

Weights in MAVT methods play the role of scaling factors in the sense that they relate scores in one criterion, to the scores of all other criteria. This means that by assigning weights of relative importance, decision makers implicitly determine how much units in one criterion they are willing to give up, in order to improve the performance of another criterion by one unit. So, if the weight of criterion *i* is double the weight of criterion *j*, then the decision maker values 10 units on criterion *i*, the same as 20 units on criterion *j* (Belton and Stewart, 2002). In order for the decision makers to more clearly realize their preferences in terms of the necessary trade-offs between criteria, weights are defined on the initial natural scales and by taking into account the absolute level of performances and absolute differences in scores. Different weights elicitation methods that have been developed for helping decision makers in articulating weights in a systematic and –more or less- consistent way will be described in more detail in Section 3.3.

The preference and indifference conditions defined at the level of single criteria apply also to total value functions, which can thus be exploited for constructing a complete preorder of the examined alternatives.

3.2.2 Outranking methods

Outranking methods proceed to a pair-wise comparison of alternatives in each single criterion in order to first determine partial binary preference functions which are subsequently synthesized over all criteria to provide total preference relations for each pair of alternatives denoting the evidence that *'an alternative a is at least as good as alternative b'*. The total preference relations can be exploited for providing a partial or complete preorder of the examined alternatives.

For deriving partial preference functions, the comparison of performances is made on the initial scale of each single criterion, either a natural cardinal scale or a qualitative ordinal one. Thus, it is not the performances of alternatives but the strength of preference between pairs of alternatives that are determined on a normalized scale, from 0 to 1, the former denoting no preference (including indifference) and the latter strict preference. A common characteristic in most outranking methods is the extension of the usual notion of criterion by 'pseudo'-criteria associated with indifference and/or preference thresholds. These thresholds denote that in front of pair-wise comparisons, human preferences do not abruptly pass from the state of indifference (*p(a,b)*=0) to the state of preference (*p(a,b)*=1). In other words, by considering a certain criterion, the decision maker can be indifferent between two alternatives a and b presenting only a small difference in their performance, or weakly prefer one alternative over the other ($0<p(a,b)<1$) if their difference is larger but not large enough to justify strict preference.

Once partial preference functions have been calculated for all pairs of alternatives in each criterion, one has to proceed to their aggregation by taking into account the weight of relative importance associated to each criterion. Total preference functions $P(a,b)$ result as the weighted average of partial ones:

$$P(a,b) = \sum_i w_i \bullet p_i(a,b) \qquad \text{with } \Sigma w_i = 1$$

At a final step total preference functions are exploited -following a different technique in each method- in order to construct outranking relations and establish preorders of the examined alternatives. Due to the imprecise nature of the preferential information, the resulting preorders are in all cases non-complete, meaning that some of the alternatives might appear as incomparable to each other. This result is in many occasions very useful since it indicates strong conflicts in the criteria in which incomparable alternatives show the one high and the other a low performance and forces decision makers to think harder on their preferences. However, in most cases a further refinement of the initial preorders is possible in order to obtain a complete preorder i.e. to remove incomparabilities.

The two most known outranking approaches are the ELECTRE family of methods developed by Roy and his collaborators (Roy, 1996) in the Paris Dauphine University (with ELECTRE I being published already in 1968 and ELECTRE III being the most widely used), and PROMETHEE method developed by Brans and Vincke (1985) in the Free University of Brussels.

It should be noted that in outranking methods weights of importance have a different meaning than in MAVT/MAUT methods. They do not represent trade-offs between criteria scores, since they are used to combine preference relations and not scores assigned to individual alternatives. Therefore, they should be interpreted rather as measures of the degree (the voting strength) each criterion influences a final statement of whether or not *'alternative a is equal or preferred to b'*. It is clear that if this statement is valid in the most important criteria then there are more arguments to accept the overall validity of such an assertion.

3.3 Weighting techniques

Weighting techniques are one of the several preference elicitation approaches used to derive in quantitative or qualitative terms the importance attached by people to various criteria considered in a decision or a policy making procedure. These methods include surveys trying to approximate public values in monetary terms, indirect economic approaches where public values are obtained indirectly from consumer behavior at the market place, and focus groups or workshops where experts or stakeholders construct their preferences through a deliberative approach (Keeney et al., 1990).

Weighting techniques are developed in the framework of MCDA. As a consequence they are also divided in two broad categories following the classification of MCDA methods: Compensatory weighting techniques are used for aggregating partial values in MAVT methods, while non-compensatory ones are used for aggregating partial preferences in outranking methods. The former assume strong compensation between criteria and are used as scaling factors, while the latter reject this assumption and are used as importance coefficients in the respective aggregation formula.

3.3.1 Compensatory weighting techniques

All compensatory weighting methods aim at highlighting the hidden dilemmas behind a number of mutually exclusive options evaluated across multiple criteria for making stakeholders become aware of the potential gains and losses implied by their choice in the specific decision context. Thus, it is meaningful to elicit them by taking into consideration the impact range in each single criterion (Keeney, 1992). In this sense, the derived weights have no absolute meaning and do not reflect general values in life but only preferences and priorities in front of the considered alternatives. The most widely used compensatory weighting methods are the following:

Trade-off method (Keeney and Raiffa, 1976): The Trade-off method is grounded on a rigid theoretical background and is revealing the dilemmas posed to the stakeholders through a pair-wise examination of criteria. Each pair comprises two hypothetical alternatives, differing only in two criteria. The first alternative has the best performance in one criterion and the worst in the other criterion, while opposite performances are assumed for the second alternative. Stakeholders are asked to choose one of the two alternatives, their choice indicating the most important criterion. In a second step, they define how much they are willing to give up in this criterion in order to improve the other at its best level. Their answer reveals directly the trade-off between the criteria. The minimum number of comparisons is n-1, while a self-consistency check necessitates the examination of all possible combinations of criteria, thus of n·(n-1) comparisons.

Swing method (von Winterfeldt and Edwards, 1986): The Swing method is first constructing two extreme hypothetical Scenarios W and B, the former presenting the worst performance in all criteria (worst score of the examined alternative options) and the latter the corresponding best performance. It is assumed that the current state for the stakeholder is Scenario W. The preference elicitation procedure consists in asking stakeholders to carefully look at the potential gains if moving from W to B and then to decide which of the criteria they want to first shift to Scenario B. Assuming that this first swing is valued with 100 units on a hypothetical value scale, the stakeholders are asked to assign a value (<100) to the second criterion moved to B, then to the third and so on until the last criterion is moved to Scenario B.

SMART method (Edwards and Barron, 1994): the method is a simplified multi-attribute rating approach including also a procedure of attribute weighting which is performed in two steps. First, the decision maker is asked to rank the importance of changes in the attributes from the worst attribute levels to the best levels. Next, he is asked to make ratio estimates of the relative importance of each attribute relative to the one ranked lowest in importance. The resulting weights are normalized to sum to one.

MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) (Bana e Costa and Vansnick, 1994): MACBETH is a MAVT method integrating the weights elicitation procedure as an essential part of the overall evaluation approach. The method comprises "swing" and trade-off elements while comparisons include hypothetical reference scores characterized as good and neutral. It furthermore provides the necessary consistency tests for the coherency of the procedure.

Conjoint method: This method is a typical example of a holistic procedure, requiring the stakeholders to rank or rate the examined alternatives by taking into account their scores in the whole set of criteria. These holistic preferences are then disaggregated by using regression analysis in order to derive the single value functions and the corresponding weights. The method is very convenient for stakeholders because of the simple preferential information required. However, people have the tendency to ignore or misinterpret many attributes, still important ones, when ranking alternatives, while treating them in an inconsistent way. In addition, they are not forced to deliberate and become conscious of their own preferences, so that the derived weights cannot be considered as defensible and balanced (Hobbs and Meier, 2000).

3.3.2 Non-compensatory weighting techniques

Contrary to compensatory approaches, non-compensatory ones reflect principally global values about the relative importance of criteria, and do not pay particular attention to the impact range of the specific decision context, which is assumed to be taken into account in the construction of partial preference relations. The most widely used compensatory weighting methods are the following:

Direct point allocation or fixed point scoring techniques: it is the simplest among all weighting techniques. The decision maker or stakeholder is asked to distribute a fixed number of points among the criteria. Usually they are expressed as percentages where 100 points are allocated among the criteria. The attribute with the highest score is the most important one.

Ratio weighting techniques: they are similar to direct point allocation differing only in that preferences are elicited in two stages. First, the decision maker is required to rank the criteria and then he is asked to rate them according to their relevant importance. For example the least important attribute can be assigned with a value of 10 and all the others can be rated as multiplies of 10, or alternatively the most important criterion can be assigned a value of 100 and all the others may be expressed in proportion to it. The resulting weights are normalized to sum to one.

Resistance to change (Rogers and Bruen, 1998): This method is based on the Personal Construct Theory and is recently adapted to estimate the relative importance of environmental criteria in outranking MCDA methods. It has elements from the "swing" method but is considered as reflecting rather relative importance coefficients than scaling factors. Each criterion is assumed to have two different poles of performance, the desirable (best) and the undesirable (worst) side. Assuming that all criteria are at the desirable side, stakeholders are asked to compare all criteria between each other in pairs and to choose which one they are willing to move from the desirable side to the undesirable side. The relative importance of each criterion is obtained from the number of times it was resistant to change.

The Analytic Hierarchy Process (AHP) (Saaty, 1980) is an integrated MCDA method including elements from both categories, MAVT and outranking methods. As far as the attribute weighting is concerned, the method requires the DM to examine pairs of criteria and rate their relative importance on a nine-point scale, varying from 1 (equally important) to 9 (extremely more important). The weights are derived by the principal eigenvector of the comparison matrix and are subsequently normalized. The number of the pair-wise comparisons that should be executed in order to have all the criteria compared to each other is $N = c\,(c-1)/2$. Hence, the procedure might be long and lead to inconsistencies which have to be removed.

4. Guidelines for implementing CBA and MCDA in policy assessment

4.1 Overview of the assessment procedure

A policy appraisal procedure –independently of the methodological approach used- can be divided in three distinct phases as shown in Fig.1:

Phase A) Pre-assessment phase, in which the problem faced is defined and its structural elements are identified and analysed. The approach to be followed for the problem solution is selected and the necessary preparatory actions are undertaken.

Phase B) Assessment phase, in which the valuation procedure is implemented and the distinct decision perspectives are synthesized. This aggregation procedure provides an ordinal or cardinal ranking of the examined alternative options, while an uncertainty analysis is performed before concluding on the most advantageous policy action(s).

Phase C) Post-assessment phase, in which the outcome of the assessment procedure is subject to the examination of different actors or policy centres for arriving, through deliberation and negotiations, at a commonly agreed solution to be disseminated to a wider public and successfully implemented.

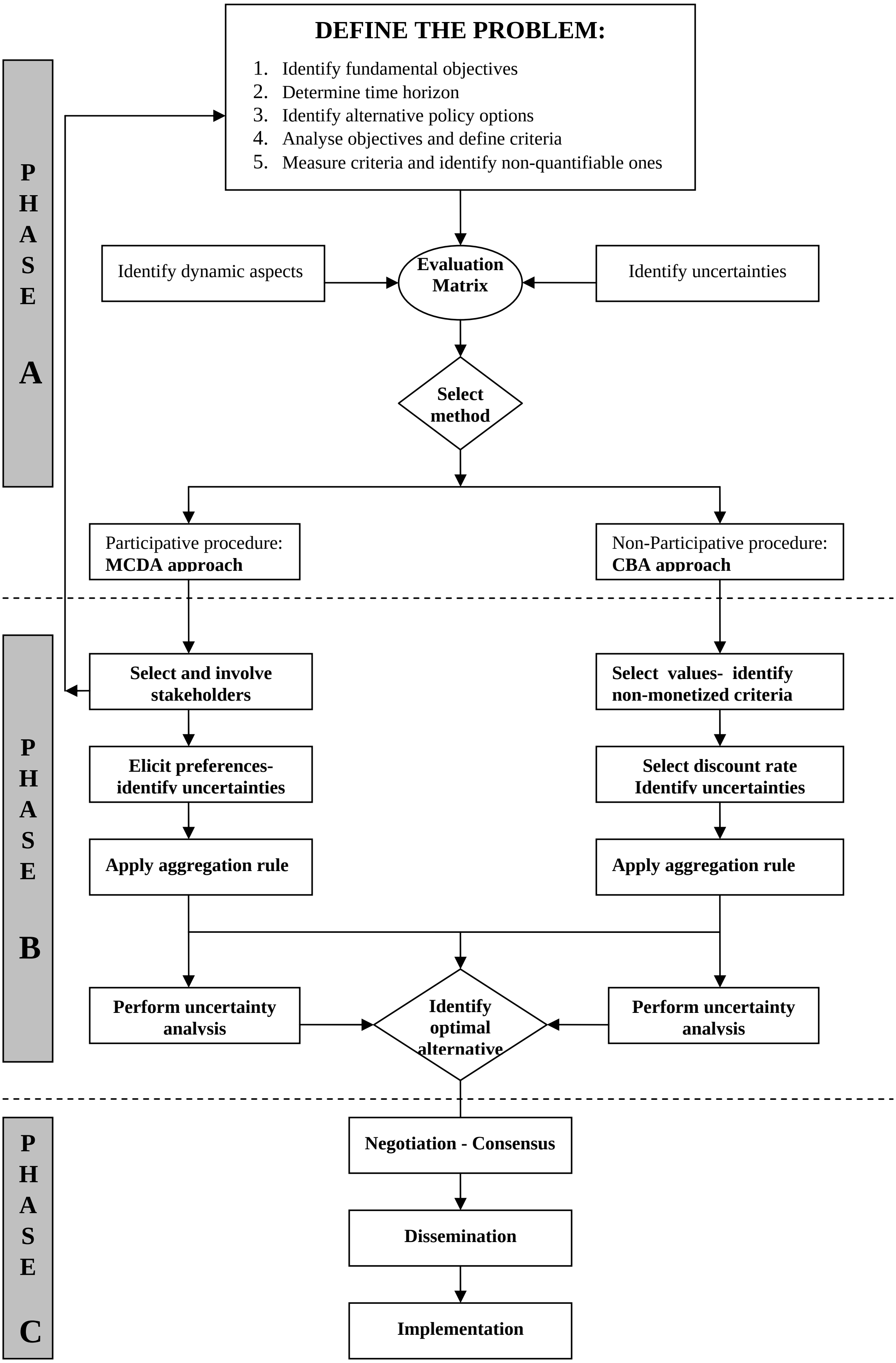

Figure 1: Phases and stages of the assessment procedure

Each phase comprises a number of steps corresponding to different tasks of the team charged with the responsibility of making the choice among alternative policy options and eventually of the involved stakeholders. In order to underline the association between the two major approaches the step-wise procedure can be depicted in a unique flow-diagram as shown in Fig. 1.

These procedural steps are briefly described below by emphasizing the main assumptions and prerequisites of each methodology.

4.2 Pre-assessment phase

The pre-assessment phase has the same function in both CBA and MCDA assessment frameworks and consists practically in defining the policy problem under consideration and selecting the methodological tool used to synthesize all its structural elements and assist the policy making procedure.

4.2.1 Problem definition

Problem definition and structuring is an integral stage in most MCDA applications, especially those dealing with unstructured, one-off problems. On the contrary, CBA is usually applied in more or less well-structured problems (e.g. evaluation of alternative projects). However, here also the first step in any CBA application is to specify the policy issue at hand and the policy options to be assessed.

In fact, policy assessment is an intricate task because the problems addressed are usually encountered within a complex environment including several interrelated systems defined upon a multiplicity of parameters. Moreover, policy makers usually strive for satisfying a multiplicity of goals or, for meeting some policy objectives without generating serious adverse effects in other areas of public concern. Besides complexity, uncertainty is another inherent characteristic in policy assessment, stemming from the lack of information for different aspects under consideration, the variability of systems' parameters, the limited scientific knowledge about physical phenomena, and the hesitations of the decision makers about their aspirations. Thus, policy problems have to be gradually structured through a systematic process of eliciting and exchanging ideas and information.

'A well structured problem is a problem half solved' is an often quoted statement emphasising the significance of looking in depth on the problem faced, in order to define the scope and the focus of the analysis, to identify its main structural elements and to recognise external constraints such as physical or legislative environment, time and resources available etc.

The structural elements to be identified during the problem structuring process are the following:

Define fundamental objectives: fundamental objectives constitute the main areas of concern that have to be pursued in policy formulation. Although several policy making schemes start by considering a given set of alternative policy options among which one has to make his choice, there are arguments for reversing this order. The rationale behind the so-called 'value-focused thinking' is that by exploring and increasing awareness about fundamental objectives might lead to the creation of new interesting policy options that were not visible at first glance (Keeney, 1992).

Identify the time horizon: the time horizon specifies the period over which the policy is expected –or assumed- to affect the socio-economic environment. In this sense, it sets the time boundaries for the subsequent analysis of objectives and the specification of alternative policy options. Furthermore, the time preference of individuals has to be taken into account in the measurement of some economic criteria (e.g. annualized investment cost), in the discounting of future costs or benefits in CBA, as well as in determining dynamic aspects of the decision context.

Identify alternative policy options: the existence of alternatives is the essence of any decision or policy making procedure, even if one considers only one option: here also the problem involves two alternatives, either adoption or rejection of the considered option. Sometimes, there is a multiplicity of options that are explicitly or implicitly presented to policy makers. In such case, it may be advisable to reduce the number of alternatives into a manageable set including only prevailing options, or characteristic options representing the whole range of possible courses of action.

Analyse objectives and define criteria: the analysis here aims at identifying a hierarchical structure into the initial set of fundamental objectives. This means that a fundamental objective can be divided into specific sub-objectives, which in turn can be also split in lower-level objectives as to arrive at a consistent set of evaluation criteria. This top-down approach helps in clarifying a vague area of concern by producing more detailed concepts that are relevant to the examined alternatives, measurable and amenable to value judgments. Measurability implies the possibility to gauge performances of alternative actions in quantitative or qualitative terms, while value-relevance implies the possibility to link these measures with human preferences. In such a procedure one has to avoid including redundant criteria or double counting of the same concept expressed in different terms.

Measure criteria and identify non-quantified ones: the lower level criteria resulting from the hierarchy of objectives are usually associated with a commonly understood measurement scale. However, it may happen that some criteria can not be measured in such an objective quantitative scale, although expressing a clear concept that is not possible to be further analysed. In this case, it is necessary to construct a subjective ordered scale (e.g. 0-1, or 1-7) on which to assign the score characterizing each alternative option. The measurement of criteria provides the performances of the examined alternatives either in absolute terms, or relative to some baseline, which is usually set as the performance in a specific point in time, or the performance of a Do-Nothing scenario.

4.2.2 The evaluation matrix

The direct outcome of the above structuring process is an evaluation matrix comprising:

- A set *A* of *n* alternative –mutually exclusive- options: $A= (A_1,A_2,\ldots,A_n)$ and
- A set *C* of *m* evaluation criteria: $C= (C_1,C_2,\ldots,C_m)$ of which k criteria can be measured in an objective physical scale, while *m-k* criteria can only be ranked on a subjective ordered scale.

Alternative A_i can be described in terms of each of the evaluation criteria so that the elements of the evaluation matrix are the performances g_{ij} of alternative A_i in the criterion C_j. In this sense, the evaluation matrix represents a simple model of the considered policy problem integrating the basic information needed to proceed to the assessment of the examined options.

However, it is often necessary to investigate two additional issues that might influence the structure and/or the mathematical representation of the evaluation matrix:

Uncertainty: in the pre-assessment stage uncertainty relates mainly to the measurement of performances and is very common in policy issues. This is because the systems concerned are large and complicated and the time horizon is extended so that it is difficult to accurately estimate the consequences of the examined actions on the natural and social environment. Such uncertainty could emerge as a consequence of missing data or of incomplete understanding of the interactions between system parameters. Obviously, uncertainty is larger in non-quantified criteria There are different possibilities to integrate uncertainty in the assessment procedure. Probability theory and fuzzy sets theory are two of the most effective ways to handle uncertainties provided there are data about the distribution of uncertain parameters in the former case, or the range and the most plausible values in the latter. In some cases in which uncertainty refers to a single or a few evaluation criteria, one can introduce additional criteria representing the risk to deviate from the estimated performances.

Dynamic aspects: dynamic aspects constitute another type of uncertainty related with future developments. More specifically, the effects of policies are expected to vary with time, as long as the systems change or the policy itself evolves according to these changes. Thus, one should focus rather on the time dimension influencing some critical parameters than on detailed information about the whole system. A scenario approach is the most suitable way to deal with dynamic aspects. One should stress that these scenarios are external to the assessment procedure, since they are out of the policy makers' control and are used to describe plausible future developments that may influence strategic decisions (Van der Heidjen, 1996). The number of external scenarios must be limited, while each one has to be a meaningful representation of plausible future developments and secure internal consistency.

It is clear that uncertainty and dynamic aspects add two additional dimensions in the evaluation matrix. This creates further difficulties in the assessment procedure and necessitates the adjustment of the evaluation method used and of the corresponding modeling procedure.

4.2.3 Method selection

It is already emphasized throughout this document that the evaluation of discrete policy options with respect to multiple criteria can be accomplished with two alternative assessment frameworks: Cost-Benefit Analysis (CBA) and Multiple Criteria Decision Analysis (MCDA). Besides the differences in their theoretical background that are already described in section 2 and 3, the choice among the two methods implies the consideration of the particular decision context. Specifically, one has to take into account the degree to which public involvement and/or experts' opinion is desired in the assessment procedure and the time this involvement is better to come about. Based on this consideration the selection of the assessment method can be seen as the selection between a participative and a non-participative assessment procedure.

MCDA is usually implemented in a participative framework involving experts and stakeholders interacting with each other and facilitating a deeper understanding of the policy issue at hand. Participation and interaction constitute two essential elements of the policy making process which is transformed into a learning procedure that motivates stakeholders to think harder about the conflicts addressed and eventually to arrive at a better and commonly accepted solution (Martinez-Alier et al., 1998; Lahdelma et al. 2000). Moreover, the involved actors might influence the structure of the problem in the sense that they can suggest inclusion of additional alternatives or criteria or rejection of some of the proposed ones, provide novel insights about performances and relevant uncertainties etc.

CBA is generally conceived as a non-participative procedure capable of including in an objective way public opinion, as revealed in the marketplace or derived from relevant surveys. Objectivity is recognised as an important advantage of CBA together with the commensurability of the examined effects in a familiar valuation scale. Nevertheless, exactly these advantages can be the source of dispute and opposition, especially if the values used are not adequately explained or relate to a completely different decision context (Munda, 1996). CBA can be seen as a valuable tool for providing policy advice and its outcome can be confirmed and enriched in an open deliberative procedure (e.g. workshop or public forum) in which additional non-monetized effects are possible to be taken into account through MCDA or by means of an informal evaluation approach.

In case a participative assessment framework is selected, it is necessary to specify the particular MCDA approach (MAVT or outranking) and the particular method that will be used in the assessment phase. Similarly, in case the preference is for a non-participative assessment framework, one has to choose among CBA and CEA depending on the particular policy context and data availability.

4.3 Assessment phase

The stages comprised in the assessment procedure differ according to the selected assessment framework, although in both methodologies the objective is the same: to synthesize the multidimensional problem in order to obtain on a single value scale an index reflecting human preferences and denoting the relative dominance of the examined alternatives.

4.3.1 Assessment in a MCDA framework

The steps included in a typical MCDA assessment procedure are the following:

Specify the stakeholders to be involved: the organization or entity responsible for making the policy choice has to identify a comprehensive list of stakeholders encompassing individuals or groups of people with an interest or concern for the examined policy issue. The selection should ensure that all different views are represented in the list, although in practice such an all encompassing list it not realistic. Attention should be also paid to the propensity of the selected stakeholders to influence other stakeholders as well as to the potential of making coalitions (Belton and Stewart, 2002). It is important to note that the term stakeholder(s) is usually distinguished from the term decision maker(s), the latter being those assigned with the responsibility to take the final decision, eventually but not necessarily with the active involvement of stakeholders. On the other side, experts are those having a particular experience on aspects of the policy context and can provide valuable insights in the nature of the problem without necessarily participating to the assessment procedure, especially if their experience is limited to a single aspect and could provide biased judgments.

Elicit preferences: in this step the aim is to capture the stakeholders' preferences in front of the particular problem as defined in the specific decision context. The difficulty here is that preferences can not be considered as definite and a priori stored within human mind, but they are modeled during the assessment procedure. One has to distinguish two types of preferential information:

Intra-criterion preferences: judgements refer to the value attached to different levels of performances and to differences between them. In MAVT methods, this kind of preferential information is derived through a scoring procedure aiming at defining partial value functions in each particular criterion (see 2.4.1). Since this type of preferential information is usually difficult to be derived, linear value functions are often assumed. The linearity assumption is found to affect decisions less than variations in weights whose elicitation merits particular attention (von Winterfeldt and Edwards, 1986; Hobbs and Horn, 1997). In outranking methods intra-criterion preferences are expressed by defining indifference, preference or veto thresholds (see 2.4.2).

Inter-criterion preferences: judgements refer to the relative importance attached to the information carried by each single criterion. This kind of preferential information is derived by using weighting techniques (see 2.5) enabling decision makers to realise their aspirations. As already explained, in MAVT methods, weights represent scaling factors relating scores and their differences in one criterion to scores in another criterion, while in outranking methods they are simply denoting the influence that each criterion has in building up the total preference relation. In each case weights have a different meaning and are derived through different techniques.

Apply aggregation rule: the aim in this step is to combine alternatives scores and preferential information in order to arrive at a final solution that takes into account all evaluation criteria. The aggregation procedure and its theoretical background differ in the two MCDA categories as described in 2.4.1 and 2.4.2.

4.3.2 Assessment in a CBA framework

A typical CBA procedure is normally performed by the group of analysts, consultants or researchers assigned with this responsibility. The tasks accomplished are the following:

Select monetary values: The monetary values assigned to the quantified criteria play the role of the weights used in MAVT aggregation. They namely express a compensatory preference model in which losses in one aspect can be compensated by gains in another aspect. In the case of traded goods monetary values are defined on the basis of existing market prices, assuming they are not seriously distorted. For non-traded goods where prices are missing, monetary estimates can be drawn through a wide variety of economic approaches determining stated or revealed preferences of individuals or simply transferring relevant estimates by carefully selecting the original source studies and applying a systematic adjustment procedure, if necessary (Pearce et al., 2006).

Impacts that are not amenable to quantification or for which existing monetary estimates are not reliable, should be clearly indicated in order to be taken into account in an ex-post assessment of results or for inclusion in MCDA.

Select discount rate: Discounting practically means assigning a lower weight on future costs and benefits. The more distant in the future a certain cost or benefit occurs, the lower its weight. Moreover, the higher the discount rate the lower the present value of the respective item. Despite the many critics and disputes about the 'fair' discount rate, in many practical applications the social discount rate varies between 3-4%.

Apply aggregation rule: similarly to MCDA approaches, aggregation implies the summation over all evaluation criteria, here expressed in monetary terms as costs and benefits. The NPV criterion (benefits greater than costs) is applicable in cases of independent projects or policies, implying that any option with NPV> 0 is acceptable. In the case of choice among mutually exclusive policies, alternative options are usually ranked on the basis of the B/C ratio, both components expressed in present values. Another way is to calculate the IRR of the time series of net costs or benefits. The advantage here is that one does not have to select a discount rate but only to compare the resulting IRR with those applicable in practice. However, if the compared policies differ largely in the implementation cost, the IRR can provide misleading results in favour of the lower cost option.

4.3.3 Ranking and uncertainty analysis

The ranking obtained by applying the aggregation rule of either the selected MCDA approach or of CBA highlights the most preferred policy option as the one ranked first. Outranking approaches might provide a partial ranking in which two or more actions are found to outrank other candidate options although being incomparable between each other. This result can motivate a further meditation on the pros and cons of each option and can offer valuable insights to the policy process.

In any case an uncertainty analysis is necessary in both CBA and MCDA in order to check the robustness of the revealed solution. The problem structuring phase usually involves many uncertainties related mainly with the measurement of criteria performances. Other sources of uncertainty are: in MCDA the elicitation of preferences, and in CBA the monetization of physical impacts and the selection of discount rate.

A multiplicity of uncertainty treatment techniques can be used, from a simple sensitivity analysis focusing on variations of the most crucial parameters, to more sophisticated methods such as probabilistic approaches, Monte Carlo analysis or use of fuzzy set theory.

4.4 Post-assessment phase

During this phase the results of the appraisal procedure have to be scrutinized in order for them to be widely accepted and be transformed in a workable operational plan for policy implementation. To this end the following steps should be undertaken:

Negotiation and consensus: in the framework of MCDA this stage is usually the last stage of the assessment procedure. This is because in the presence of multiple stakeholders it is very common the case of diverging results indicating different value systems and policy priorities. However, this disagreement should not lead to a dead-end of the whole procedure but instead, give rise to a constructive dialogue where each stakeholder tries to justify her/his opinion by at the same time being motivated from the arguments of other stakeholders to recognise other points of view or eventually realise false expression of his/her own values. Such a structured discussion is often supported by specific computational techniques aiming at discovering the main sources of divergence and highlighting common views and possible compromises.

In the case of CBA, where the solution results from an ‘objective’ assessment procedure, an ex-post involvement of stakeholders is also highly recommended. This involvement can take the form of a workshop in which the working team introduces the policy issue under consideration, describes the analytical procedure followed, gives the necessary details for underlying values and relevant assumptions and presents the obtained results. The stakeholders are called to express and justify their opinion for the overall task and eventually, participate to a final decision stage by including non-monetized impacts through qualitative statements or through a formal MCDA approach.

In both cases, during this negotiation process it is possible to gradually arrive at a commonly accepted solution that has more chances to be adopted by all interested parties and be successfully implemented

Dissemination and implementation: the aim in this final stage is to communicate the whole policy problem and the agreed solution to a wider public in order to achieve the maximum possible social awareness and minimize the risks of hindering, altering or delaying its practical implementation.

References

Bana e Costa, C. and Vansnick, J.C. (1994). MACBETH: An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research*, 1 (4), 489-500.

Belton, V. and Stewart, T.J. (2002). Multiple Criteria Decision Analysis: An integrated approach. Kluwer Academic Publishers, Dordrecht.

Brans, J.P. and Vincke, P. (1985). A preference ranking organization method: the PROMETHEE method for multiple criteria decision-making. Management Science 31, 647-656.

Costanza, R. (2000). Social goods and the valuation of ecosystems sevices. Ecosystems, 3, 4-10.

Edwards, W. and Barron, F.H. (1994). SMARTS and SMARTER: improved simple methods for multiattribute utility measurement. Organizational Behavior and Human Decision Processes 60, 306-325.

Hobbs, B. and Horn, G. (1997). Building Public confidence in energy planning: a multimethod MCDM approach to demand - side planning at BC gas. Energy Policy, 25, 357-375.

Hobbs, B. and Meier, P. (2000). Energy Decisions and the Environment: A guide to the use of Multicriteria Methods. Kluwer Academic Publishers, Dordrecht.

Horowitz, J.K. (1996). Environmental Policy under a non-market discount rate. Ecological Economics 16, 73-78.

Keeney, R. and Raiffa, H. (1976). Decisions with Multiple Objectives. J. Wiley & Sons, New York.

Keeney, R., von Winterfeldt, D. and Eppel, T. (1990). Eliciting public values for complex policy decisions. Management Science, 36, 1011-1030.

Keeney, R. (1992). Value - Focused Thinking: A path to Creative Decision Making, Harvard University Press. Cambridge, Massachusetts.

Lahdelma, R., Salminen, P. and Hokkanen, J. (2000). Using Multicriteria Methods in Environmental Planning and Management. Environmental Management, 26, 595-605.

Martinez–Alier, J., Munda, G. and O'Neill, J. (1998). Weak comparability of values as a foundation for ecological economics. Ecological Economics, 26, 277-286.

Munda, G. (1996). Cost – Benefit Analysis in integrated environmental assessment: some methodological issues. *Ecological Economics*, 19, 157-168.

Pearce, D., Atkinson, G. and Murato, S. (2006). Cost-Benefit Analysis and the Environment: Recent Developments. OECD Publ. Paris.

Rogers, M. and Bruen, M. (1998). A new system for weighting environmental criteria for use within ELECTRE III. *European Journal of Operational Research*, 107 (3), 552-563.

Roy, B. (1996). Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers, Dordrecht.

Saaty, T.L. (1980). The Analytic Hierarchy Process. McGraw-Hill. New York.

Schulze, P.C. (1994). Cost-Benefit Analyses and Environmental Policy. Ecological Economics 9, 197-199.

Snell, M. (1997). Cost-benefit analysis for engineers and planners. Thomas Telford Ed. London.

Van der Heijden, K. (1996). Scenarios: The art of strategic conversation. John Wiley & Sons

Virani, S. and Graham, S. (1998). Economic Evaluation of Environmental Policies and Legislation. Final Report prepared for the Eur. Commission (DG III) by Risk and Policy Analysts Ltd. Norfolk, UK.

von Winterfeldt, D. and Edwards, W. (1986). Decision analysis and Behavioural research. Cambridge University Press, New York