Quality Assurance and Safety of Crops & Foods

Quality Assurance and Safety of Crops & Foods 2010, **2**, 182–196

REVIEW

Recent developments in multi-criteria evaluation of regulations

F. Bartolini and D. Viaggi

Department of Agricultural Economics and Engineering, University of Bologna, Bologna, Italy

Keywords
Multi-Criteria Analysis; policy evaluations; decision support; indicators; multiple impact.

Correspondence:
Fabio Bartolini, Department of Agricultural Economics and Engineering, University of Bologna, Bologna, Italy.
Email: fabio.bartolini@unibo.it

Received 31 May 2010; revised 29 July 2010; accepted 5 August 2010.

doi:10.1111/j.1757-837X.2010.00076.x

Abstract

Introduction Agricultural activity is affected by a wide set of regulations and prescriptions due to agricultural and non-agricultural policies. In agricultural policy literature, growing attention is being paid to the issue of policy evaluation. *Objectives* The objective of this paper is to provide a literature review of the Multi-Criteria Analysis applied in the evaluation of policy and regulations. *Methodology* The paper proceeds by firstly identifying the main issues of the policy evaluation process and then discussing the potential contribution of Multi-Criteria Analysis as a decision support tool in a policy-making context. *Results* Multi-Criteria Analysis was created as a decision support tool for the selection of alternative projects when decisions are effected by trade-offs between different criteria. However, existing applications to policy evaluations still show significant room for improvement. In particular, an improved ability to support decision-making processes can be expected if Multi-Criteria Analysis is included in a consistent process of monitoring and gathering of data, and both Multi-Criteria Analysis and monitoring are developed in such a way as to provide a structured interaction with the decision maker in a participative approach. *Conclusion* A cautious integration of Multi-Criteria comparison techniques and policy-analysis concepts should be sought in order to better contribute to policy evaluation processes.

Bartolini F & Viaggi D (2010). Recent developments in multi-criteria evaluation of regulations. *Quality Assurance and Safety of Crops & Foods*, **2**, 182–196.

Introduction

Agricultural and agri-food activities are affected by a wide range of regulations and prescriptions due to agricultural and non-agricultural policies. Generally multiple institutions representing different geographical levels [municipality, province, region, country, European Union (EU), international] participate in the policy-making process. The decision maker's (DM) actions affect and impacts on the stakes and interests of a variety of stakeholders, including: consumers, farmers, food retailers and food processors. In addition, the DM's actions (or non-actions[1]) have impacts on democracy, ethics, transparency, responsibility and accountability (Funtowicz & Ravetz, 1994; Munda, Nijkamp & Rietveld, 1994; Munda, 2004; Van den Hove, 2006).

The complexity of human interactions, the uncertain context and the value conflicts make it difficult to foresee and estimate the effects/impacts of the DM's actions (Munda, 2004). In recent years, as described by Gibbons and Georghiou (1987) and Georghiou and Roessner (2000), the demand for policy evaluations by public administrations has increased significantly, mainly with the aim of improving the quality of policy making, and the policy-implementation process. In fact, the importance of public policy evaluation has increased significantly in the last 20 years, and it is now rather common to have *ex ante* or *ex post* evaluation exercises attached to policy-making procedures, undertaken either by the public institutions themselves, by consultants or by scientists. In fact,

[1] Dye (1972) defined public policy as 'whatever governments choose to do or not to do'.

it is increasingly common to find an allocation of funds to cover the cost of internal/peer/independent evaluation procedures in programme budgets. In the last 10 years the EU has made increased efforts to evaluate its policies. In this context the Directorates General have increased both research and consultancy aimed at providing *ex ante* evaluations of agricultural policy, and in particular those belonging to the first pillar (SFP, direct market support, milk quotas). Also, since the inception of the 2000–2006 programme a compulsory evaluation (*ex ante*, mid-term and *ex post*) for decentralized policies such as those belonging to the second pillar has been established. In particular, the *ex post* and mid-term analyses of Rural Development Policy are recognized as being particularly important for the entire Rural Development Policy cycle. Such importance has also been highlighted by the evaluation exercise fund allocations.

In such a complex and uncertain framework DM have increased their interactions with scientists, and the use of advice instruments, projects, tenders, etc., which support their activities and allow for a more robust analysis of policy impacts (Munda, 2004). In spite of this, the policy-evaluation results provided by the available tools are not completely satisfactory (see e.g. Finn *et al*., 2009 for an exemplary discussion concerning agri-environmental schemes). This is the case for at least three reasons. Firstly, the representation of complex systems often results in significant simplifications, and the consideration of only partial components of systems (Funtowicz *et al*., 1999; Munda, 2000). Secondly, in general, normative exercises such as evaluations imply an operational definition of 'value' that is a representation of the importance placed on different elements by different social actors (Munda, 2004). Finally, the limited knowledge about future states of nature/conditions and the temporal distance between policy implementation and policy-impact results in evaluation exercises being characterized by a high degree of uncertainty (Stirling, 1998).

A wide body of literature exists on the assessments of agricultural and food policies, most of which are evaluated through a quantification of efficiency or effectiveness of a whole programme or a single measure. The evaluation tools are mainly based on either a monetization of impact [Cost–Benefit Analysis (CBA)] or a quantification of utility based on physical, economic and social indicators [Multi-Criteria Analysis (MCA)]. The first analysis is based on a monetization of both private and public costs associated with the adoption of a policy. Through a comparison of all costs and benefits it is possible to identify the net economic impact of each policy alternative. The MCA, for its part, is based on the aggregation of several criteria expressed by monetary and non-monetary indicators, and produces a set of preferred or dominant policies, as the expression of those options associated with higher utility for DMs/stakeholders.

In recent years several review papers have been published on the use of MCA in various fields: Hajkowicz and Collins (2007) reviewed over 110 studies on the application of MCA to water planning; Pohekar and Ramachandran (2004) reviewed over 90 papers on MCA studies in the field of energy management; Hayashi (2000) examined the application of MCA to agriculture with a comparison of 80 studies concerning agricultural systems, but no specific review was found on the assessment of policy evaluations.

Application of MCA concerning regulations or prescriptions in the fields of quality assurance and food and crop safety is quite common in literature, due to the fact that decisions in this field affect several actors (consumers, society, retailers, processors, farmers, the feed and food industries, etc.) and takes into account multiple dimensions, such as profitability, safety, public health, environmental quality (Krieger, Schiefer & da Silva, 2007; Fazil *et al*., 2008; Ruzante *et al*., 2010). In fact, examples of applied works in this field can be found in either technical reports of projects or in peer reviewed papers often published in multi-disciplinary journals. Generally, in these works an evaluation of alternative regulations or prescription impacts were compared in order to identify the best strategy or a set of dominating alternatives to adopt. Alternatives may be represented by different quality systems or different crop or food protection practices or indeed by the sustainability of different crops or food products.

The objective of this paper is to provide a literature review of MCA as applied in the context of policy and regulation evaluations, with some specific references to food quality and safety. This paper highlights the main characteristics of the use of MCA as applied to public policy or regulations, and identifies crucial issues in the empirical use of MCA in order to improve the quality and robustness of results. The paper also seeks to highlight the main open issues and to identify some of the most promising areas for further research in the application of MCA to the evaluation of food quality and safety policies.

The remainder of the paper is organized as follows: a brief description of the policy-evaluation framework is provided in the second section. The third section describes the MCA approach, the fourth section illustrates the use of the MCA in the context of policy evaluations. The fifth section provides a discussion, while the concluding remarks are provided in the sixth section.

Policy evaluation

As described in the introduction, policy evaluation is a growing topic in social sciences, as well as in economics and agricultural economics literature. Following the definition of Nachmias (1979) policy evaluation can be defined as the 'objective, systematic, empirical examinations of the effects ongoing policies and public programs have on their targets in terms of the goals they are meant to achieve'. Policy evaluation can be seen as the final phase of a general policy cycle.[2]

Four different purposes for which to evaluate a programme or policy are identified: planning/efficiency, accountability, implementation and institutional strengthening (Evalsed, 2009). Economists, including agricultural economists, have focused their research mainly on planning/efficiency and the implementation of the policies/programmes (Gallerani, 2008).

Generally, the main objective of a given policy evaluation is to assess its effectiveness, through the quantification of the net benefit provided (additionality of the programme) expressed as the degree to which the objectives are achieved.

The challenges in the evaluation process can roughly be divided in two groups: (a) those related to the measurement of programme outputs, impacts or results; and (b) those related to the identification and quantification of objectives (Finn *et al*., 2009). With regard to (a), the measurement of programme impacts generally face a number of practical problems (such as cost and time), which prevent accurately measuring the benefits generated. Consequently, evaluations tend to focus on programme outputs, rather than impacts (Primdahl *et al*., 2003; Viaggi *et al*., 2010). On the other hand, with respect to (b), a proper evaluation of the effectiveness of matching results and objectives cannot be performed in most cases due to the lack of clearly defined programme objectives. The first difficulty encountered is that, in the majority of cases, the policy contemporaneously addresses several specific objectives. The statement of these objectives differs across programmes. The most common situation is that programmes provide a statement of their objectives, but lack both a clear specification of the target level to be achieved for each objective, and of the ultimate importance of each objective.

With regards to the timing of an evaluation in an ideal policy cycle, the DM can undertake three different kinds of evaluations: *ex ante*, ongoing/mid-term and *ex post* (Figure 1).

Ex ante refers to a stage when the policy is still to be implemented and the evaluation is performed in order to define/choose policy parameters (policy design). At this stage, many hypothetical alternatives may be devised and their outcomes have to be computed using some more or less sophisticated tool.

Ex post refers to a stage where the policy has already taken place. The purpose is to evaluate its outcome in order to gain information that is useful to revise policy design. The main issues are which references and methodologies are used and which data sources are available for measuring indicators.

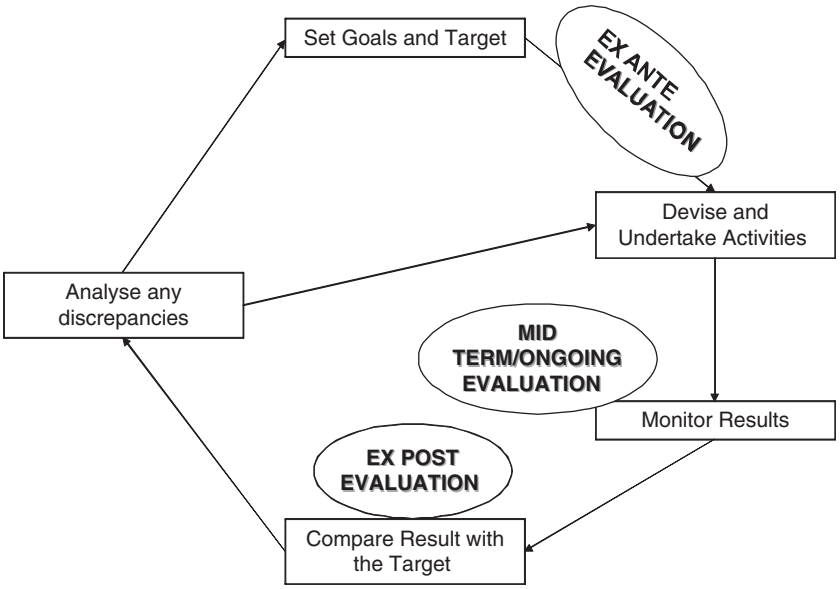
A mid-term/interim evaluation is placed between the *ex ante* and *ex post* evaluation and represents an initial verification of preliminary results and the achievement of policy objectives, or can be used to collect information about the feasibility of the *ex post* evaluation (e.g. measurability of specific indicators).

Ex ante, mid-term and *ex post* evaluations are interconnected, and should be comparable with respect to criteria and objectives. An example of the chronology of an evaluation of three different policies is presented in Figure 2.

Following the definition of evaluation provided in the first part of this section, public policy can be articulated in terms of needs, objectives, inputs and effects (outputs, results and impacts), that follow different policy cycle steps (see European Commission, 2004; Evalsed, 2009 for a definition). The structure of the evaluation criteria is presented in Figure 3.

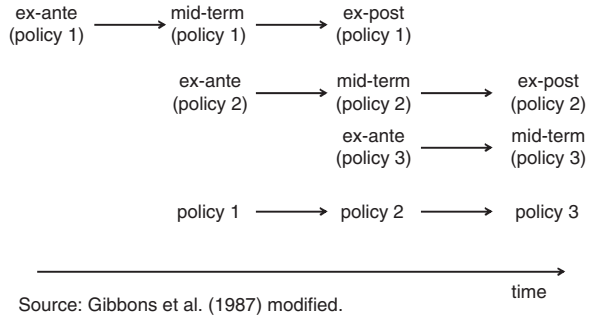
It is possible to identify five different criteria for the evaluation of policy or programme performance: relevance, efficiency, effectiveness, utility and sustainability. The two criteria most commonly used to measure the effect of a

[2]The description of the policy cycle is beyond the scope of this paper. Howlett & Ramesh (2003) have identified the ideal policy cycle as consequentiality of five phases: agenda setting, policy formulation, decision process, policy implementation and policy evaluation. Ideally the policy is realised in response to a need, a problem or an issue arising in society (policy objective). When the objective has been set the DM makes an operative definition of the allocated budget, the target, the policy instrument adopted by the agents involved, the timing for interventions required for policy implementation (policy formulation, decision-making process and policy implementations). Finally, when the programme has completed its effect the DM makes a comparison with the results of the programme/policy with respect to the target in order to analyse the policy performance and the impact of the policy, and if the initial need/issues has been solved by the programme.



Source: Bartolini et al., 2005 (modified).

Figure 1 *Ex ante*, mid-term/ongoing and *ex post* evaluations in the ideal policy loop.



Source: Gibbons et al. (1987) modified.

Figure 2 Evaluation chronology.

programme are effectiveness and efficiency. Effectiveness is the extent to which a policy achieves its objectives, including threshold levels, targets, etc. Efficiency refers to the extent to which a policy may achieve these objectives at minimum cost. In other words, effectiveness measures the achievement while the efficiency measures the achievement/cost ratio.

The criteria of utility and sustainability are two different additional concepts (Evalsed, 2009). They represent, respectively, a judgement on the degree to which the impacts obtained by the programme are in relation to broader societal and economic needs (utility), and whether they are durable over time (sustainability). The relevance criterion measures the coherence and pertinence of the policy in relation to societal needs and problems, as well as *vis-à-vis* economic or environmental needs. For a more thorough explanation of the criteria and the policy questions answered by these criteria see Evalsed (2009).

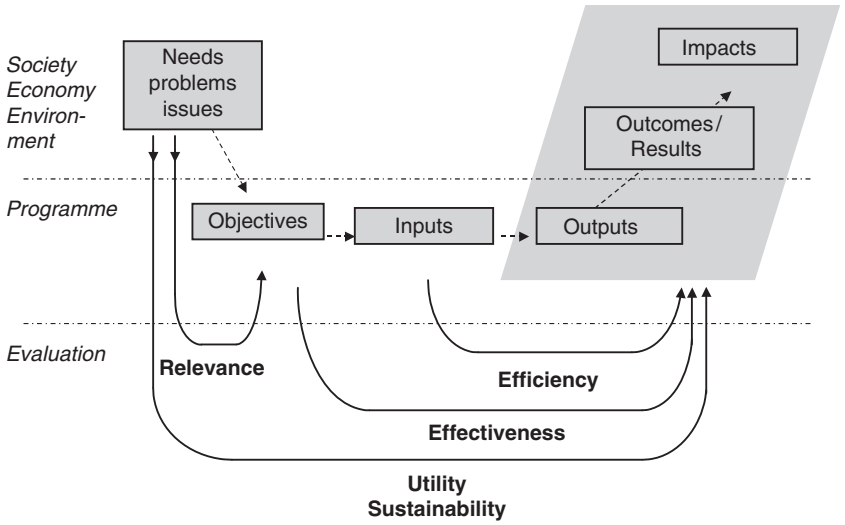
As explained in the beginning of this section each evaluation typology (*ex ante*, *ex post* or mid-term evaluation) is undertaken in order to answer different evaluation questions, which should be expressed by different evaluation criteria (Table 1).

Generally the quantification of policy/programme effectiveness and efficiency are undertaken in all evaluations (*ex ante*, mid-term and *ex post*), while the quantification of relevance is specific to *ex ante* and mid-term evaluations and the measurement of sustainability and utility is most frequently undertaken in *ex post* evaluations. In addition, the quantification of effectiveness and efficiency requires the application of differentiate approaches/methodologies. Such methodologies are mainly differentiated by the measurement and typology of indicators used to express changes, and by the approach used to isolate the programme/policy additionality (see Evalsed, 2009 for a review of the methodologies adopted in the different stages of the evaluation).

Multi-Criteria Analysis

The basic feature of MCA is that it evaluates (compares) items (alternatives) on the basis of more than one criteria or objective. This makes MCA a powerful tool for multi-dimensional problems such as policies with a wide set of economic, environmental and social impacts, often involving different actors/stakeholders (French, 1993; Nijkamp & Vindigni, 1998).

Originally MCA was developed to support project selection and management, and has matured significantly since



Source: Evalsed 2009.

Figure 3 Evaluation criteria.

Table 1 Evaluation purposes

	Evaluation typology		
Criteria	*Ex ante*	Mid-term	*Ex post*
Relevance	X	X	
Effectiveness	X	X	X
Efficiency	X	X	X
Sustainability			X
Utility			X

Source. European Commission (2004), modified.

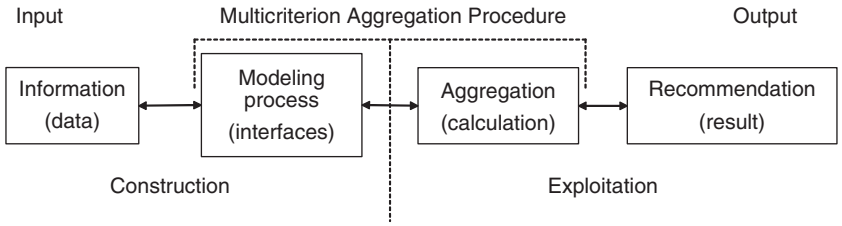
its inception. The initial application to policy evaluation was done by way of the multi-criteria decision-making (MCDM) approach, in which a structure was sought for the decision problem. The goal of MCDM is to provide guidance to DM in the determination of the most desirable solutions – generally those which lead to the achievement of the long-term goals identified by the DM (Stewart, 1992). The MCDM results identify the alternatives with the highest DM utility function values (Munda, 2004). The MCDM application is based on an assumption of additive utility functions, and on the existence of a DM that is able to incorporate all societal preferences. In Europe the concept of MCDM was developed for Multi-Criteria Decision Aid (MCDA), emphasizing the role of the analysts in providing support to DMs' choices rather than in selecting the best alternative/actions. In fact, MCDA is less ambitious than the MCDM, and the results are a set of alternatives that are dominating, allowing greater 'freedom of speech' for the DM (Roy, 1985). Finally, a more recent development was to increase the level of public participation in the decision-making process by way of a more participative and transparent evaluation process. This was done in order to take into account the policy/political constraints and responsibility (see Banville et al., 1998, for the developed of Stakeholder MCDA, and Munda, 2004 for the social multi-criteria evaluation).

The assumptions inherent in the MCA approach are the following: (a) many criteria have a role in guiding the evolution of a system; (b) such criteria are, at least locally, in conflict with each other; (c) the criteria tend to require a compromise or a choice (*arbitrage*) (Roy, 1985).

The theoretical basis for solving multi-criteria problems can be found in utility theory, where an action or an alternative is preferred if its expected utility is higher than the expected utility of the alternatives to which it is compared (Keeney & Raiffa, 1976). However, comparing alternatives highlights a rather delicate issue. To start with, by comparing two generic alternatives three different relationships between alternatives can be identified:

- preference (strict, weak);
- indifference;
- incomparability.

The 'preference' relationship can be divided into strict or weak preferences (see Roy, 1985 for a better explanation of the preference relationship). An 'indifference' relationship exists when the expected utility of two actions is considered to be equal. Finally, there is 'incomparability' when no choice may be made between two alternatives because of opposite preferences or a lack of information (Roy, 1985; Nijkamp & Vindigni, 1998).



Source: Guitouni and Martel (1998), modified.

Figure 4 Multi-Criteria Decision Aid scheme.

MCA can be summarized in two consequential phases: model construction and model exploitation (Guitouni & Martel, 1997). In Figure 4 the process of MCA for supporting decision making is represented.

The first phase begins with the analysis of the data and information available, and with the modelling process in which the evaluator seeks to represent the DM's evaluation structure. This phase, in particular the structuring of decision making, is one of the most complicated phases of the MCA due to the amount of information and knowledge required and to the interaction with the DM (Roy, 1985; Bouyssou, 1996). The second phase (model exploitation) involves the aggregation process, representing the calculation of the utility provided by each alternative, and which will be completed with the production of the recommendations.

The characteristics of decision-making process modelling and aggregation distinguish the several MCA models. These two parts (together called Multi-Criteria Aggregation procedure) are the most developed in the literature (see Table 2 for a taxonomy of the aggregation methods). In this literature, and in particular under the operational research discipline, a very large number of algorithms have been proposed in order to synthesize DM preferences into a single judgement.

Three main families of methods are provided in the literature:

(1) Elementary methods, in which a simple choice criterion is used (e.g. maxmin method; weighted sum).
(2) Single synthesizing criterion: the preference is based on the choice of the alternative that provides the higher value of DM utility function. Such values correspond to the sum/product of the utility provided by all criteria, or by the lower distance with respect to an ideal situation. This method allows for a preference in which, for each alternative, it is possible to state a situation of preference or a situation of indifference. Such methods are fully compensatory, i.e. very bad performance and good performance compensate among indicators.
(3) Outranking methods: introduce aggregation procedures based on concordance or discordance concepts (Hayashi, 2000). With respect to the previous families, these methods enable the broadening of judgement preferences to include strict or weak preferences, identifying thresholds for a concordance or discordance index on which one alternative is at least as good as another one. Such methods are partially or non-compensatory (see Roy, 1985 for a review of this concept).

Table 2 A taxonomy of Multi-Criteria Aggregation Methods

Elementary methods	Weighted sum
	Lexicographic method
	Conjunctive method
	Disjunctive method
	Maxmin methods
	Maxmax methods
Single synthesizing criteria	Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)
	Multi-Attribute Value Theory (MAVT)
	Utility Theory Additive (UTA)
	Simple Multi-Attribute Rating Technique (SMART)
	SMARTS
	SMART Exploiting Ranks (SMARTER)
	Multi-Attribute Utility Theory (MAUT)
	Analytic Hierarchy Process (AHP)
	EVAMIX
	Multi-criteria robust interactive decision analysis (MCRID)
	Fuzzy weighted sum
	Fuzzy maximin
	Fuzzy maximax
Outranking methods	Elimination Et Choix Traduisant la Réalité (ELECTRE) and further development
	Preference Ranking Organisation Method for Enrichment Evaluations (PROMETHEE) and further development
	MELCHIOR
	ORESTE
	Regime methods
	Novel Approach to Imprecise assessment and Decision Environmental (NAIDE)

Source. Guitouni and Martel (1997), modified.

Among the numerous methods proposed, some of the most commonly used MCA methods are (Bartolini *et al.*,

2005): (a) Weighted sum; (b) Elimination Et Choix Traduisant la Réalité (ELECTRE); (c) Preference Ranking Organisation METHod for Enrichment Evaluations (PROMETHEE); (d) Analytic Hierarchy Process (AHP); (e) Regime analysis; (f) Fuzzy techniques.

The weighted sum methods are the most simple multi-criteria tools used for solving problems related to the ranking of alternatives, and are perhaps the most commonly used. The weighted sum allows a ranking of the alternatives based on the sum of the utility provided by each indicator. This method allows for the maximum consideration of trade-offs among indicators (i.e. it is fully compensatory).

The ELECTRE family of methods are developed and described by Maystre, Pictet and Simos (1994); Roy and Bouyssou (1993); Vincke (1992). They are based upon a comparison of alternatives through quantitative parameters calculated mostly through concordance or non-discordance indexes. This kind of index represents the degree to which it is possible to concord or to non-oppose to the relation that 'an alternative a is at least as good as alternative b'. The concordance index expresses the aggregation of the criteria that support the assertion that 'an alternative a is at least as good as alternative b'. Otherwise the non-discordance index expresses the aggregation of the criteria for which the relation 'an alternative a is at least as good as alternative b' is violated. The computation of the index and the generation of an outranking is strongly differentiated among the different methods of the ELECTRE family (see Figueira, Mousseau & Roy, 2005 for the description of the different methods and Figueira, Greco & Roy, 2009 for the extensions of the method).

The PROMETHEE method was created by Brans (1982) and was developed by Brans and Vincke (1985) and Brans and Mareschal (1992). This method uses an outranking principle, based on the identification of positive and negative preference flows that represent, respectively, how much an alternative is dominating the others, and to what extent an alternative is dominated by the others (Brans & Mareschal, 1992). An order of the alternatives is obtained by comparing the positive and negative flows (see Brans & Mareschal, 1992, for a review of the methods and Chenayah & Takeda, 2008 and Rao & Patel, 2010, for the extensions of the method).

The AHP is an MCA procedure based on different (hierarchic) levels of aggregation (Saaty, 1980). The idea behind AHP is to structure the decision problem in sub-problems for which it is easier to express a judgement. Such a judgement is structured as a pairwise comparison between all elements placed in the same hierarchical level, and express the importance with respect to an element placed at the upper hierarchical level (using a nine step verbal scale). Various methods are used to derive priorities/rankings from the pairwise comparison matrix: eigenvector (Saaty, 1980); arithmetical average or geometrical average (Gass & Rapcsak, 1998); goal programming (Bryson, 1995) regression (Laininen & Hamalainen, 2003); or finally the separable representations derived from the field of mathematical psychology (Bernasconi, Choirat & Seri, 2010).

The regime method is designed to use qualitative and quantitative information in MCA (Hinloopen & Nijkamp, 1990; Munda, Nijkamp & Rietveld, 1994). The regime method is a discrete multi-assessment method which is appropriate for analyzing projects or policies (Nijkamp *et al.*, 2007). This method is based on the pairwise comparison of each alternative for all criteria and when such pairwise comparison is ultimate, on the quantification of a value that represent the probability of dominance of an alternative with respects the other. Such operation is performed for each pair of alternatives. Finally, the ranking of alternatives is obtained by aggregating the probability of dominance for each alternative (Hinloopen & Nijkamp, 1990).

During the 1990s, using the concept of fuzzy sets as proposed by Zadeh (1965), the fuzzy MCA techniques were developed, often applied to previously existing MCA approaches (Cornelissen *et al.*, 2002). Such an approach allows for the inclusion of uncertainty in the treatment of information (indicators, alternative values, etc.). In fact, the fuzzy concept is based on the extension of the notion of probability, using a 'multi-value logic' that helps assign a degree of membership in a specified set to each value (Kaufmann, 1975; Dubois & Prade, 1980). Generally speaking, the fuzzy approach has been extended to MCA methods which are structured following five consequential steps: (a) definition of the linguistic variable; (b) construction of the membership functions; (c) computation of the degree of membership (fuzzyfication); (d) determination of the fuzzy conclusion (fuzzy inference); (e) achievement of the fuzzy conclusion (defuzzyfication). See Cornelissen *et al.* (2002) for a review of fuzzy approaches, and see Fernandez *et al.* (2009) and Sevkli (2010) for examples of the application of fuzzy sets with ELECTRE methods; see Li and Li (2010) and Zhang *et al.* (2009) for a recent application of fuzzy sets with the PROMETHEE method, and finally see Aly and Vrana (2008); Kahraman and Kaya (2010) for an example of the application of a fuzzy set with the AHP method.

Adaptation of MCA to policy evaluation context

The comparison between CBA and MCA has been a central point in the debate about project and policy evaluation in recent decades, often using the same arguments either in favour or against one of the two. On the one hand, CBA is more able to provide a judgement, which takes into account the preferences of the entire population with respect to MCA. In fact, MCA bases the judgement on a more 'technocratic approach' on which the decision expresses the priorities of only a few decision makers/stakeholders (Munda, 2009). On the other hand, MCA mimics the decision-making process better than CBA, and can hence be used as a direct support to decision making and possibly as a tool to involve different stakeholders explicitly in the decision-making process, including negotiations or the search for compromises. Altogether, the use of aggregation based on non-monetary indicators can be considered more suitable in the context of policy evaluation for at least three reasons: (i) the multiple nature of the policy impacts which requires the use of different indicators measured through different scales; (ii) the difficulty in the monetization of social and environmental benefits due to the absence of a market value for environmental goods (environmental effects are often examples of externalities); and (iii) finally, the possibility of expressing the differences in importance of the policy objectives using the concept of weights.

However, due also to the criticisms highlighted above, the application of MCA to the policy and regulation evaluation context is strongly dependent on the accuracy of the process, paying particular attention (in addition to the methods adopted as explained in the previous paragraph) to the following crucial points: (a) setting the problem; (b) identification and quantification of evaluation criteria; (c) weighting; (d) results interpretation and analysis.

Setting the problem

The first step-in of an MCA is to set the problem in terms of the objectives of the evaluation: What is going to be evaluated? Why? How? In the majority of cases, the MCA applied to policy analysis is used to support decision making about alternative project/policy options based on a comparison of the performance of a policy expressed in terms of efficiency or effectiveness or cost-effectiveness. In this case, the analysis can be carried out: (a) *ex ante*, and the problem is mostly to compare/rank project alternatives in order to decide which one to implement; (b) *ex post*, with the aim of (b1) comparing results among cases/areas, or (b2) classifying cases or (b3) comparing real outcomes with expected/counterfactual/optimal outcomes, or (b4) comparing the cases at different points in time. The four options (b) may also be combined. In both *ex ante* and *ex post* evaluations a crucial role is the interaction between the evaluator and the DM, in order to identify and adapt the right methodology/methods while taking into account the role of DM in the evaluation (e.g. interactive methods versus non-interactive methods), the level of complexity and transparency required in the process and finally the information available, its quality and its robustness.

Most MCA methods are based on a comparison of different alternatives, and the objects of comparison may differ if the evaluation is performed *ex post* or *ex ante*. Such evaluation exercises imply the identification of the variables that define each alternative, in such a way as to make alternatives comparable.

Application to policy or regulations connected with the implementation of quality assurance or crop safety systems are unique in that the evaluation problems are mainly *ex ante* and the alternatives represent the implementation at several levels (farm, firm, Europe, a production sector, etc.) of quality systems or the adoption of technological innovation with several purposes. Recent examples of the use of MCA in this context include several issues: the comparison of the implementation of several quality systems (ISO 9000, EuropeGAP, both or neither) for extra-EU firms that trade in Europe (Krieger, Schiefer & da Silva, 2007); the comparison of different interventions aimed to reduce pathogen risks or increase the protection of food or crop production (see, Themelin *et al.*, 1997; Fazil *et al.*, 2008; Mouron *et al.*, 2010; Ruzante *et al.*, 2010); the comparison of different food and nutrition policy options (see Gonzalez-Zapata *et al.*, 2008, 2009 for those policies related to the reduction of obesity) and the sustainability impact assessment of food or crop production practices in a given territory or in an area (Janikowski, Kucharski & Sas-Nowosielska, 2000; De Lange *et al.*, 2009; Siciliano, 2009; Witters *et al.*, 2009; Laudien, Pofagi & Roehrig, 2010; Turner, Morse-Jones & Fisher, 2010).

An evaluation of a regulation requires some representation of the implementation process and of the causal relationships between policy/prescription design or implementation and outcomes. In fact, an institutional analysis seeks to understand and formulate the explanation hypothesis (e.g. causal framework), to identify the main stakeholders involved in the policy process and the actors affected by the policy, and finally to generate the alternative options required (Munda, 2000; O'Connor, 2000).

Identification and quantification of the evaluation criteria

The evaluation criteria are aimed at quantifying the consequences of the alternative actions. For this reason the criteria used in the evaluation should in principle meet three specific characteristics (Maystre, Pictet & Simos, 1994):

- exhaustiveness: implies that all relevant criteria covering the relevant economic, environmental, and social factors must be considered;
- consistency: implies that the overall judgement of the DM must be coherent with the preferences structure of each criteria. This implies that, if the DM is indifferent between two alternative for all criteria minus 1, and then alternative A is preferred with respect to one criterion, and/or alternative B degrades with respect to one criterion then alternative A is preferred to alternative B. Otherwise the criteria considered are inconsistent with respect to the decision-maker's preferences (Yoe, 2002);
- non-redundancy: implies avoiding duplication and overlapping of the criteria.

Criteria may be measured through indicators, which may also form the basic information for an MCA. Onate *et al.* (2000) and Primdahl *et al.* (2003) highlight some major distinctions for the selection of (environmental) indicators for policy-evaluation purposes, with a distinction between those that measure policy impacts and those that measure policy outcomes/results. In addition, the family of policy-output indicators could be identified and measured (coherently with Figure 3).

Because of the high specificity of the MCA applied to quality assurance or crop safety policy, the selection of indicator definitions is often undertaken through participative approaches such as focus groups or brainstorming sessions involving several expertises (Fazil *et al.*, 2008; Kuzma *et al.*, 2008). Fazil *et al.* (2008) have identified four main properties necessary for the selection of criteria, which are important in the MCA applied in this field. Such criteria are (a) weight of evidence, (b) effectiveness, (c) cost and (d) practicability. While the first set of indicators captures the scientific evidence of the policy-interventions proven by the specific literature, the second and the third are related to the expected results and their costs. Finally, the fourth set of indicators will express their ability to be implemented within the specific context of the policy.

The choice of the evaluation criteria and their nature (qualitative, quantitative, stated, measured, etc.) may be conditioned on the possibility to collect reliable data to measure them, in particular when an *ex post* evaluation is required on an issue for which a clear project for data collection aimed at the evaluation was not set up from the beginning. On the other hand, the specific data requirements will vary according to the choice of indicators. When an *ex ante* evaluation is carried out, or counterfactual alternatives need to be produced, data collection is not sufficient, but outcomes need to be generated for the counterfactual situation. Different techniques may be used, ranging from simple intuitive computation, to expert interviews, to economic or physical/biological models. Measurement methodologies that are homogeneous across alternatives (e.g. an expert opinion or a focus group) must be also characterized by the same features, (e.g. the same experts) for all the different alternatives, in order to increase the actual comparability.

Weighting

Weights represent the relative importance of each criterion in determining the social welfare associated with each alternative (Roy & Bouyssou, 1993), and allow for the determination of explicit social preferences about different objectives. Nijkamp and Vindigni (1998) classify the methods for the quantification of weights into four categories: (a) ranking criteria, (b) rating method, (c) verbal statement, (d) paired comparison. The difficulties in deriving weights are numerous. A key issue is the identification of relevant stakeholders participating in the evaluation/weight elicitation procedure. A second problem rests in the difficulty for individuals to express importance in numerical terms, in particular if the numbers do not represent a measure with which the DM has experience (e.g. money). Furthermore, when produced through interviews or Delphi, strategic answers may strongly affect the results. Also, the psychological state or the specific interest of the person being interviewed, may affect the results and their reliability. Weights may be also criticized for their background mathematical implications/assumptions. In particular, they are normally assumed to incorporate all relevant preferences, to be independent from changing external conditions, and to be linear (constant) in all indicators.

Difficulties with weights highlighted in the literature (Hayashi, 2000) are mostly in the field of preferential independence among criteria, which may not always hold. In addition, range sensitivity (weights should vary as a function of the range of the level of each criteria) may be important to consider. Finally, the way in which weighting is performed may affect the weights obtained, and should be

consistent with the selected aggregation procedure. For example, when a set of indicators have a given hierarchical structure the method used to elicit weights must maintain that structure (e.g. through AHP). Finn *et al.* (2009) provide an example in which a hierarchical weighting procedure is adopted by copying the hierarchical structure of the criteria.

Of the issues inherent in weight elicitation, the main problem remains who to ask about weights. Most MCA procedures rely on DMs, i.e. the private actors in charge of the decision or the policy makers that should take the decision. However, in the case of policy evaluation, policy makers are the target respondents as long as they may be expected to interpret the collective preferences. This may be questioned on various different grounds, and generally depends on the fact that the primary policy-maker's objective is to maintain the consensus (Howlett & Ramesh, 2003), which may not necessarily reflect the preference structure of the collectivity. Another approach is to realize interviews in a contingent valuation style (in such a way as to reproduce the willingness to pay in a CBA). This increases the costs and requires a strong refinement of methodologies. Finally, individual representatives of different stakeholders may be involved and interviewed. In this case, the problem is how to aggregate the values expressed by different groups. The synthesis of the opinions when different groups are involved in weight elicitation can be obtained using two different approaches: group-based or individual-based (Tsiporkova & Boeva, 2006). The former approach is oriented towards obtaining a consensus within a group, mainly via discussion or negotiations. On the contrary, the latter approach is oriented towards using each individual response as a unit of decision making, and applying methods to synthesize judgements. In this approach, one option is for different weight vectors representing different stakeholder groups to be calculated and impacts quantified separately for each stakeholder group. In this case the comparison of the alternatives is judged from the point of view of the different stakeholder groups, which helps to identify possible compromises across actors. In some cases, using MCA as a support for participatory project/policy evaluation, the discussion may focus on the trade-offs between objectives relevant to different actors (see Bartolini *et al.*, 2010 for an application). The result could be some compromise solution derived without formal weight estimation.

Unweighted procedures are possible, and have the advantage of skipping a difficult part of the procedure. Yet they omit an important issue. In many cases, unweighted procedures are nothing more than hidden weighted procedures. For example, if no weights are used in a weighted sum, the result is the same as giving to each one out of n criteria a weight equal to $1/n$. As a result, the weight of different indicators is simply determined by the number of indicators proposed by each evaluation dimension.

The weighting phase in the MCA applied to policies and regulations addressing the field of quality assurance and food or crop safety is very dependent on the multi-criteria method used for the aggregation procedure. However, in this field, and in the analysis performed using emerging technologies, weights are often elicited using participative approaches which involve several representatives of the various actors/stakeholders involved in the calculation (Kuzma *et al.*, 2008).

Interpretation and analysis of results

The results of MCA require interpretation. Interpretation means answering such questions as: (a) altogether are the results stable and trustable?; (b) are differences in scoring/ranking relevant, and if so; (c) to what extent?; and (d) what are the determinants of the results? The latter issue is particularly relevant if it helps gain an understanding of the results and hints for the improvement of the design of alternatives. This may be done both by classifying alternatives according to design/results, or by analyzing the components of the final score/ranking in order to understand more about its determinants.

A sensitivity analysis constitutes a common method for the verification of the 'trustability' and credibility of the results. An analysis of this kind is aimed at evaluating how results would change as a result of changes in assumptions or parameters. A sensitivity analysis is particularly important when data are uncertain, for example when estimation procedures are not completely reliable or when there are strong assumptions in the computation procedure. A good sensitivity analysis may be very useful in ruling out some expectations or excluding some alternatives 'whatever the external conditions may be'. In a strict sense, sensitivity analysis applies to test results variation towards a change in one single parameter. It can be substituted by a scenario analysis, where consistent combinations of external parameters are devised and fed into the analysis. Results are then compared across scenarios. This helps simplify the analysis when single possible parameters on which to perform sensitivity are too numerous. In any case, a relevant issue is how to generate the new measure of indicators when the changing parameters are those determining the value of indicators. In this case, the procedure should return to the measurement step for each scenario, hence requiring tools to estimate the change in the indicator values that are sensitive enough to

allow for a simulation of meaningful differences in indicator values as a function of changes in the scenario parameters.

Discussion and open issues

MCA was developed as a decision support tool for the choice of alternative projects when the decision is effected by trade-offs between criteria. In recent years it has been extended to both *ex ante* and *ex post* policy analysis. However, applications to policy evaluation still demonstrate significant room for improvement. This is partly due to the difficulties in bringing a methodology from the well-defined setting of a project, to a wider and less stringent set of decisions represented by a policy. This problem is generally common to CBA; in fact, CBA also appears to suffer more difficulties when applied to wider issues (e.g. recent applications to climate change policies, and to the full cost recovery for water services).

Because of the fact that it is reliant on bringing together information from multiple sources, MCA has developed into a 'catch-all' of different mathematical algorithms. Consequently, it offers a huge set of alternative methods from which to draw. For this reason, application to policy analysis does not seem to require particular advances in the direction of finding further alternative algorithms for the calculation of evaluation parameters. On the contrary, greater attention should be paid to the suitability of such algorithms for the specific decision-making problems at stake, and to avoiding the confusion that arises for researchers/DMs as a result of an excessive number of methods.

In addition to these general issues related to application of MCA to policy evaluation, specific areas for improvement emerge when MCA is confronted with specific policy areas. We discuss some of the relevant open issues for quality policy evaluation more in detail in the following, focusing on three main areas of concern:

- selection of basic parameters (alternatives, objectives, indicators);
- a better incorporation of preferences, beyond the use of weights;
- comparison of monetary costs and multi-criteria effects.

The first stage is devoted to the structuring of the problem. This stage requires a very careful understanding of the aims of the intervention to be evaluated, and of its expected causal effects in the system. This is typically a difficult issue in cases like quality and safety regulations and systems in which the measures taken usually involve marginal changes in very complex systems, potentially affecting a number of different dimensions. The bias caused by an inappropriate selection of relevant objectives/indicators can result in the whole analysis being totally misleading. The use of systematic checklists of potential areas of impact, causal connections and indicators is the main approach used to overcome this problem. Relevant examples are provided by Krieger *et al.* (2007), but greater efforts in this direction should be sought.

The selection of appropriate indicators in this case goes far beyond the selection of thematic issues, but has rather to do with the adoption of measures consistent with the policy evaluation framework. This concerns, in particular, the use of indicators accounting explicitly for the additional effects of the measures taken and the use of a consistent baseline across the different dimensions of policy evaluation. The issue of additionality, and hence, implicitly, the identification of differential effects compared with some counterfactual, is one of the most practical unresolved issues in policy evaluation. The main approach to deal with this issue, however, is the use of a highly transparent approach allowing the users of the evaluation to appropriately appreciate the information content of the results. Typical potential errors are threefold. First is the confusion between benefits resulting from different levels of product quality or safety, and the differential benefits arising from a change in quality levels exclusively due to regulations. This error may stem from overlooking the fact that there may be other drivers of quality/safety improvement, and attributing all of them to one specific regulation during the evaluation process. A second source of error results from disregarding the heterogeneity of the conditions of firms with regard to quality/safety standards when a given measure is introduced. The evaluator is often tempted to assume that the relevant agents are totally non-compliant, and hence to overestimate the effect of some policy. On the contrary, when standards and regulations are introduced, firms generally range from already being close to compliance (in this case no additional effect would arise), to being totally incompliant (which is the case that generates higher benefits, but also higher costs). Third, when the regulation/policy concerns the application of procedures, rather than standards, such as in the case of quality systems, the additional problem arises of evaluating the degree to which the procedure will change the actual flows or benefits. Analogous problems arise when information measures are taken. As these systems are likely low cost compared with the economic dimension of the process to which they are applied, optimistic assumptions about their effectiveness very easily lead to an overestimation of their effectiveness.

The incorporation of preferences through public participation and direct stakeholder MCA-assisted negotiation appears increasingly to be the desired path to follow, as a complementary and more preference-rich method compared with weights elicitation. Once again, the feasibility of such an approach is strictly connected to the procedural design of policies. An appealing field of research seems to be in the direction of using MCA approaches together with the rich amount of information already available from studies on consumer behaviour. In particular, consumer studies explicitly based on the evaluation of a range of product/process attributes seem to offer interesting opportunities for integration.

In addition, even in the participative MCA, and even when all actors' preferences are correctly identified, no truly convincing solution has been applied to the problem of how to aggregate or take into account the conflicts between the utility perceived by different actors. However, this problem is solved by providing the DM with a description of impacts across all stakeholders, and by giving the DM the ability to state preferences based on a description of policy impacts across actors. This is very relevant for quality/safety policies as they typically involve different points of view, ranging from the firm level to the public, as well as the food chain management level (Krieger, Schiefer & da Silva, 2007).

Finally, in a policy evaluation, a particularly tempting and useful approach is to compare budgetary and monetary information with MCA results. A relevant issue here is to distinguish the budgetary costs of policies from the actual economic costs. This is particularly relevant for policies supporting quality and safety standards as budget costs are typically low, resulting in policies being perceived as low cost from the administrative point of view, whereas economic costs for implementation and compliance may in fact be high. As this may be very important to understand the actual incentive compatibility for economic agents (i.e. are firms encouraged or discouraged to comply based on private revenue-cost ratios), it is essential that this information be considered, even in MCA exercises. In both cases, including costs as an additional indicator can be misleading in the analysis, as their importance could be underestimated. This happens frequently in MCA due simply to the fact that cost information tends to be easily collapsed into one indicator only, whilst benefits are often represented through a number of different fields of improvement. A different and more realistic approach is the use of utility/cost ratios (where costs may be either budgetary or economic costs) as indicators of policy efficiency. The use of utility/cost ratios may be reasonably performed only when specific conditions are met; in particular (Krieger, Schiefer & da Silva, 2007): (a) if the decision alternatives involve similar costs, and (b) if the utilities build on a similar composition of benefits (at least a similar distribution between monetary and non-monetary elements).

Concluding remarks

MCA is often suggested as a support mechanism to increase the robustness of evaluation exercises. However, such robustness is strongly linked to the quality of the MCA implementation process, and the quality of the participation of the DM and the stakeholders involved in the evaluation exercise. This is particularly relevant when MCA is applied to the fields of quality assurance and crop and food safety, where the amount of very specific technical information needed is high, and the analysis requires the varied expertise of different categories of scientists (e.g. economists, chemists, physicists). In addition, the quality of the evaluation is highly dependent on the involvement of the various societal actors who will be affected by the different potential effects of such policies (from individuals to public).

While MCA provides a wide range of tools and methods of very high interest for policy analysis, this field of application nonetheless requires improvement. Given the specific characteristics of the policy process, and MCA, the most interesting strategy in this context is to further develop the consistency between MCA techniques and the deployment of decisions, as well as the proper understanding of causality mechanisms in their downstream effects. The stream of literature related to social MCA seems to be a good example in this direction. Improving the structured interaction with DMs and stakeholders through a participative approach is also key. In addition, a better foundation of MCA in the administrative process and in economic conceptualization seems necessary. In particular, an improved ability to support the decision-making process can be expected if the MCA is included in a consistent process of monitoring and data collecting, and both MCA and monitoring are developed in such a way as to provide a cautious integration between MC comparison techniques and policy analysis concepts (e.g. additionality).

References

Aly S., Vranal I. (2008) Evaluating the knowledge, relevance and experience of expert decision makers utilizing the Fuzzy-AHP. *Agricultural Economics-Zemedelska Ekonomika*, **54**, 529–535.

Banville C., Landry M., Martel J.-M., Boulaire C. (1998) A stakeholder approach to MCDA. *System Research and Behavioral Science*, **15**, 15–32.

Bartolini F., Gallerani V., Raggi M., Viaggi D. (2010) Water management and irrigated agriculture in Italy: multicriteria analysis of alternative policy scenarios. *Water Policy*, **12**, 135–147.

Bartolini F., Gallerani V., Samoggia A., Viaggi D. (2005) Methodology for multicriteria analysis of Agri-Environmental Schemes. *ITAES Project*, DEIAgra, Bologna, 126pp.

Bernasconi M., Choirat C., Seri R. (2010) The analytic hierarchy process and the theory of measurement. *Manage Science 2010 April 1*, **56**, 699–711.

Bouyssou D. (1996) Outranking relations: do they have special properties? *Journal Multi-Criteria Decision Analysis*, **5**, 99–111.

Brans J.P. (1982) L'ingéniérie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE. In: *L'aide à la décision: Nature, instruments et prospectives d'avenir* eds Nadeau R., Laundry M. pp. 183–214, Presses de l'Université Laval, Québec, Canada.

Brans J.P., Mareschal B. (1992) Promethee-V. – Mcdm Problems with Segmentation Constraints. *INFOR journal*, **30**, 85–96.

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

Bryson N. (1995) A goal programming method for generating priority vectors. *Journal of Operational Research Society*, **46**, 641–648.

Cheneyah S., Takeda E. (2008) Exploitation Procedeure Based on Eigenvector Revisited: The Concept of Weighted Preference Flows in Multicriteria Outranking Analysis. *Cybernetics and Systems: An International Journal*, **39**, 61–78.

Cornelissen A.M.G., Van den Berg J., Koops W.J., Kaymak U. (2002) Eliciting Expert Knowledge for Fuzzy Evaluation of Agricultural Production Systems. Erasmus Research Institute of Management, Erasmus University. ERIM Report Series Reference No. ERS-2002-108-LIS.

De Lange H.J., Lahr J., Van der Pol J.J.C., Wessels Y., Faber J.H. (2009) Ecological vulnerability in wildlife: an expert judgment and multicriteria analysis tool using ecological traits to assess relative impact of pollutants. *Environmental Toxicology and Chemistry*, **28**, 2233–2240.

Dubois D., Prade H. (1980) *Fuzzy Sets and Systems: Theory and Applications*. Didier Dubois, Henri Prade Academic Press, New York.

Dye T.R. (1972) *Understanding public policy*. Prentice-Hall: Englewood Cliffs, New Jersey.

European Commission. (2004) *Evaluating EU Activities – A Practical Guide for the Commission Services. (Ed Office for Official Publications of the European Communities*. European Commission, Brussels, p. 114.

Evalsed. (2009) *The Evaluation of Socio-Economic Development – the Guide (Ed Policy D-GfR)*. Evalsed, Brussels, p. 176.

Fazil A., Rajic A., Sanchez J., McEwen S. (2008) Choices, choices: the application of multi-criteria decision analysis to a food safety decision-making problem. *Journal of Food Protection*, **71**, 2323–2333.

Fernandez E., Navarro J., Bernal S. (2009) Multicriteria sorting using a valued indifference relation under a preference disaggregation paradigm. *European Journal of Operational Research*, **198**, 602–609.

Figueira J., Mousseau V., Roy B. (2005) ELECTRE methods. In: *Multiple Criteria Decision Analysis: State of the art surveys* eds Figueria J., Greco S., Ehrgott M. Springer, New York: pp. 133–162.

Figueira J.R., Greco S., Roy B. (2009) ELECTRE methods with interaction between criteria: an extension of the concordance index. *European Journal of Operational Research*, **199**, 478–495.

Finn J.A., Bartolini F., Bourke D., Kurz I., Viaggi D. (2009) Ex post environmental evaluation of agri-environment schemes using experts' judgements and multicriteria analysis. *Journal of Environmental Planning and Management*, **52**, 717–137.

French S. (1993) A manifesto for the new MCDA era. *Journal of Multi-Criteria Decision Analysis*, **2**, 125–127.

Funtowicz S., Martinez-Alier J., Munda G., Ravetz J. (1999) Information tools for environmental policy under conditions of complexity. In: *Environmental Issues Series* eds European Environmental Agency, 34 pp., Copenhagen, Denmark.

Funtowicz S.O., Ravetz J.R. (1994) The worth of a songbird: ecological economics as a post-normal science. *Ecological Economics*, **10**, 197–207.

Gallerani V. (2008) Nuove problematiche per la valutazione delle politiche agricole. In: *Riforma della PAC, evoluzioni tecnologiche e trasformazioni ambientali: aspetti economici, estimativi, giuridici e urbanistici* eds Marone E. University Press, Firenze: pp. 37–49.

Gass S.I., Rapcsak T. (1998) A note on synthesizing group decisions. *Decision Support System*, **22**, 59–63.

Georghiou L., Roessner D. (2000) Evaluating technology programs: tools and methods. *Research Policy*, **29**, 657–678.

Gibbons M., Georghiou L. (1987). *Evaluation of Research: A Selection of Current Practices. Organisation for Economic Co-operation and Development*. OECD Publications and Information Centre, Paris.

Gonzalez-Zapata L.I., Alvarez-Dardet C., Ortiz-Moncada R., Clemente V., Millstone E., Holdsworth M., Sarri K., Tarlao G., Horvath Z., Lobstein T., Savva S. (2009) Policy options for obesity in Europe: a comparison of public health specialists with other stakeholders. *Public Health Nutrition*, **12**, 896–908.

Gonzalez-Zapata L.I., Diaz C.A.D., Clemente V., Davo M.C., Ortiz-Moncada R. (2008) Assessment criteria for public policies on obesity: the view of Spanish stakeholders. *Gaceta Sanitaria*, **22**, 309–320.

Guitouni A., Martel J.M. (1997) Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, **109**, 501–521.

Hajkowicz S., Collins K. (2007) A review of multiple criteria analysis for water resource planning and management. *Water Resource Management*, **21**, 1553–1566.

Hayashi K. (2000) Multicriteria analysis for agricultural resource management: a critical survey and future perspectives. *European Journal of Operational Research*, **122**, 486–500.

Hinloopen E., Nijkamp P. (1990) Qualitative multiple criteria choice analysis - the dominant regime method. *Quality and Quantity*, **24**, 37–56.

Howlett M., Ramesh M. (2003) *Studying Public Policy: Policy Cycles and Policy Subsystems*. Oxford University Press, New York.

Janikowski R., Kucharski R., Sas-Nowosielska A. (2000) Multi-criteria and multi-perspective analysis of contaminated land management methods. *Environmental Monitoring and Assessment*, **60**, 89–102.

Kahraman C., Kaya I. (2010) A fuzzy multicriteria methodology for selection among energy alternatives. *Expert Systems with Application*, **37**, 6270–6281.

Kaufmann A. (1975) *Introduction to the Theory of Fuzzy Subsets*. Academic Press, New York.

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

Krieger S., Schiefer G., da Silva C.A. (2007) *Costs and Benefits in Food Quality Systems: Concepts and a Multi-Criteria Evaluation Approach. (Ed Agricultural Management MaFSA, Division RIaA-I)*. FAO (Food and Agriculture Organization), Rome, p. 56.

Kuzma J., Paradise J., Ramachandran G., Kim J.A., Kokotovich A., Wolf S.M. (2008) An integrated approach to oversight assessment for emerging technologies. *Risk Analysis*, **28**, 1197–1219.

Laininen P., Hamalainen R.P. (2003) Analyzing AHP-matrices by regression. *European Journal of Operational Research*, **148**, 514–524.

Laudien R., Pofagi M., Roehrig J. (2010) Development and implementation of an interactive Spatial Decision Support System for decision makers in Benin to evaluate agricultural land resources-Case study: AGROLAND. *International Journal of Applied Earth Observation and Geoinformation*, **12**, S38–S44.

Li W.X., Li B.Y. (2010) An extension of the Promethee II method based on generalized fuzzy numbers. *Expert Systems with Application*, **37**, 5314–5319.

Maystre L.Y., Pictet J., Simos J. (1994) *Méthodes multicritères ELECTRE*. Politechniques et Universitaires Romandes, Lausanne, Switzerland.

Mouron P., Aubert U., Heijne B., Naef A., Strassemeyer J., Hayer F., Gaillard G., Mack G., Hernandez J., Avilla J., Solé J., Saupanor B., Alaphilippe A., Patocchi A., Samietz J., Höhn H., Bravin E., Lavigne C., Bohanec M., Bigler F. (2010) A multicriteria decision method assessing the overall sustainability of new crop protection strategies: the case of apple growing in Europe. In *9th European IFSA Symposium*, Wien, Austria, 2010, pp. 1–11.

Munda G. (2000) *Conceptualising and Responding to Complexity*. Environmental Valuation in Europe and Cambridge Research for the Environment.

Munda G. (2004) Social multi-criteria evaluation: methodological foundations and operational consequences. *European Journal of Operational Research*, **158**, 662–677.

Munda G., Nijkamp P., Rietveld P. (1994) Qualitative multicriteria evaluation for environmental management. *Ecological Economics*, **10**, 97–112.

Munda G. (2009) A conflict analysis approach for illuminating distributional issues in sustainability policy. *European Journal of Operational Research*, **194**, 307–322.

Nachmias D. (1979). *Public Policy Evaluation: Approaches and Methods*. St. Martin's Press, New York.

Nijkamp P., Borzacchiello M.T., Ciuffo B., Torrieri F. (2007) Sustainable urban land use and transportation planning: a cognitive decision support system for the Naples metropolitan area. *International Journal of Sustainable Transportation*, **1**, 91–114.

Nijkamp P., Vindigni G. (1998) Integrated multicriteria analysis for sustainable agricultural policy evaluation. *Rivista di Economia Agraria*, **LIII**, 9–40.

O'Connor M. (2000) The VALSE project - an introduction. *Ecological Economics*, **34**, 165–174.

Onate J.J., Andersen E., Peco B., Primdahl J. (2000) Agri-environmental schemes and the European agricultural landscapes: the role of indicators as valuing tools for evaluation. *Landscape Ecology*, **15**, 271–280.

Pohekar S.D., Ramachandran M. (2004) Application of multi-criteria decision making to sustainable energy planning - A review. *Renewable and Sustainable Energy Review*, **8**, 365–381.

Primdahl J., Peco B., Schramek J., Andersen E., Onate J.J. (2003) Environmental effects of agri-environmental schemes in Western Europe. *Journal of Environmental Management*, **67**, 129–138.

Rao R.V., Patel B.K. (2010) Decision making in the manufacturing environment using an improved PROMETHEE method. *International Journal of Production Research*, **48**, 4665–4682.

Roy B. (1985). *Méthodologie multicritère d'aide a la décision*. Economica, Paris.

Roy B., Bouyssou D. (1993) *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris.

Ruzante J.M., Davidson V.J., Caswell J., Fazil A., Cranfield J.A.L., Henson S.J., Anders S.M., Schmidt C., Farber J.M. (2010) A multifactorial risk prioritization framework for foodborne pathogens. *Risk Analysis*, **30**, 724–742.

Saaty T.L. (1980) *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. McGraw-Hill, New York.

Sevkli M. (2010) An application of the fuzzy ELECTRE method for supplier selection. *International Journal of Production Research*, **48**, 3393–3405.

Siciliano G. (2009) Social multicriteria evaluation of farming practices in the presence of soil degradation A case study in

Southern Tuscany, Italy. *Environmental Development and Sustainnens*, **11**, 1107–1133.

Stewart T.J. (1992) A critical survey on the status of multiple criteria decision-making theory and practice. *OMEGA-International Journal of Management Science*, **20**, 569–586.

Stirling A. (1998) Risk at a turning point? *Journal of Risk Research*, **1**, 97–109.

Themelin A., RaoultWack A.L., Lebert A., Danzart M. (1997) Multicriteria optimization of food processing combining soaking prior to air drying. *Dry Technology*, **15**, 2263–2279.

Tsiporkova E., Boeva V. (2006) Multi-step ranking of alternatives in a multi-criteria and multi-expert decision making environment. *Information Science*, **176**, 2673–2697.

Turner R.K., Morse-Jones S., Fisher B. (2010) Ecosystem valuation a sequential decision support system and quality assessment issues. *Ecological Economics Review*, **1185**, 79–101.

Van den Hove S. (2006) Between consensus and compromise: acknowledging the negotiation dimension in participatory approaches. *Land Use Policy*, **23**, 10–17.

Viaggi D., Finn J., Kurz I., Bartolini F. (2010) Using information from mid term evaluations of rural development programmes for the multicriteria analysis of agri-environmental schemes. *Agricultural Economics Review*; in press.

Vincke P. (1992) *Multicriteria Decision-Aid*. John Willey & Sons, New York, USA.

Witters N., Van Slycken S., Ruttens A., Adriaensen K., Meers E., Meiresonne L., Tack F.M.G., Thewys T., Laes E., Vangronsveld J. (2009) Short-rotation coppice of willow for phytoremediation of a metal-contaminated Agricultural area: a sustainability assessment. *Bioenergy Research*, **2**, 144–152.

Yoe C. Trade-off analysis planning and procedures guidebook. Prepared for Institute for Water Resources, U.S. Army Corps of Engineers, Alexandria, VA, 2002, p. 155.

Zadeh L.A. (1965) Fuzzy sets. *Information Control*, **8**, 338–353.

Zhang K.J., Kluck C., Achari G. (2009) A comparative approach for ranking contaminated sites based on the risk assessment paradigm using fuzzy PROMETHEE. *Environmental Management*, **44**, 952–467.