

Multi-criteria Environmental Performance Assessment with an Additive Model

Luís C. Dias ^{(a) (b) (*)}

Sandra Silva ^{(a) (c)}

Luís Alçada-Almeida ^{(a) (b)}

(*) Corresponding author. E-mail: lmcdias@fe.uc.pt

^(a) *Institute for Systems Engineering and Computers, Rua Antero de Quental, N°199, 3000 - 033 Coimbra, Portugal.*

^(b) *Faculty of Economics, University of Coimbra, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal*

^(c) *College of Agriculture, Polytechnic Institute of Viana do Castelo, Praça General Barbosa 4900-34, Viana do Castelo, Portugal*

Abstract:

Multi-criteria decision analysis (MCDA) is an approach to evaluate alternatives (projects, plans, candidates, courses of action, etc.) that explicitly accepts more than one evaluation dimension, instead of using a single indicator. This chapter discusses the use of MCDA for assessing environmental performance taking into account multiple evaluation criteria. Additive models are commonly used for their apparent simplicity, but they require a careful interpretation of the parameters involved, namely concerning normalization and weighting. Moreover, it is debatable whether environmental performance assessment should reflect the subjective values of a single decision maker. Therefore, a partial information multi-attribute value model is used that does not require setting exact values for criteria weights. As a case study, this chapter evaluates the environmental performance of a set of Portuguese dairy farms. Five environmental criteria are considered, which have been defined by experts from the problem domain. The VIP Analysis software is used to obtain results for different partial information assumptions. These results may support recommendations in the context of selection, ranking, or classification of problems.

1. Introduction

Environmental studies are often performed with the purpose of informing decision making. Using decision aiding methodologies (Bouyssou et al. 2000; Roy 1996) can therefore be an important part of such studies. Decision aiding consists in helping decision makers to communicate and reason about a decision, to increase their insight into the problem, and to identify which grounds justify their decision in an auditable way. Thus, it should not be regarded as a tool that would dictate decision makers what is the (supposedly objective and optimal) decision they should make.

Multi-criteria decision analysis (MCDA) (Belton and Stewart 2002; Bouyssou et al. 2006; Ishizaka and Nemery 2013; Roy 1996) is an approach to evaluate alternatives (projects candidates, courses of action, etc.) that explicitly accepts more than one evaluation dimension. It is a methodological approach that recognizes that most decisions involve the need to compromise between conflicting objectives (Bouyssou 1993). MCDA addresses this conflict by acknowledging multiple evaluation criteria explicitly. This is more transparent than translating all aspects in monetary currency or following some other, single criterion to compare alternatives, considering any other concerns as constraints that define which alternatives are feasible. For instance, instead of seeking to minimize costs subject to an environmental impact constraint, or instead of seeking to minimize an environmental impact subject to a budget constraint, these two dimensions are considered as objectives to be balanced.

MCDA can be useful to support decision makers in different types of problems, or *problématiques* (Roy 1996) that require evaluating alternatives:

- Selection (or choice) problems aim at identifying the best alternative (or the best k alternatives) among a set of possibilities. Examples are the choice of the best environmental rehabilitation project (to award a prize, or to implement), the selection of the best technology for some purpose, or the choice of location for a waste treatment facility.
- Ranking (or prioritization) problems aim at obtaining a preference order among the alternatives, from best to worst. Examples are a prioritization of environmental rehabilitation projects (defining the order by which they should be implemented), or establishing a ranking of the environmental performance of countries.
- Sorting (or classification) problems aim at assigning alternatives to classes, which are typically defined a priori and ordered. Examples are sorting environmental rehabilitation projects among the categories “not urgent”, “urgent” and “very urgent”, accept/reject decisions for projects classifying the alternatives into categories “no”, “maybe, if...” and “yes”, classifying plots of land according to their adequacy for installing solar farms, or even performing an environmental rating of a product in a qualitative scale (A+, A, B, etc.).

An MCDA study usually encompasses three stages. The first stage is structuring the problem. This is arguably the most important stage because it is the basis for all analyses that ensue. Structuring entails defining what the problem is, what the alternatives and their consequences are, and what criteria should be used to evaluate alternatives. Problem Structuring Methods (Rosenhead 1996) can be helpful at this stage. Such methods are qualitative approaches to make sense of a problematic situation, not necessarily linked with MCDA, and therefore will not be detailed here. For an example in energy efficiency, the reader may see (Neves et al. 2009).

A second stage in an MCDA study is to assess the performance of each alternative on each evaluation criterion. This implies that a set of criteria has been defined to measure how well an alternative performs on the objectives defined. This measurement can be made on a quantitative or a qualitative scale. Criteria such as costs, emissions, populations or areas and volumes are prone to be measured quantitatively. When a direct indicator is not available for the objective in question, it is possible to use an indirect indicator, in the same way an IQ test is an indirect (*i.e.*, imperfect) indicator of intelligence. For instance, acres of forest lost can be an indirect indicator for the impact on a particular bird species. On the other hand, criteria such as degree of opposition of the population, or aesthetic perception of the landscape, will usually be assessed on a qualitative scale (*e.g.*, negligible, significant, etc.). When using qualitative scales, it is important to attach a precise description (a *descriptor table*) for the meaning of each level, to avoid different interpretations of the same words. For instance, (Keeney and Sicherman 1983) defined, in the context of evaluating potential power plant sites, that qualitative level “Support” means that “No groups are opposed to the facility and at least one group has organized support for the facility”. Although these authors use a code of “1” for this level, it cannot be interpreted as a number.

The third stage in an MCDA study consists in deriving a recommendation based on the outputs of the previous stages. In general this involves some type of aggregation of the performances assessed on the individual criteria. The different ways this stage can be conducted are the subject of Section 2 of this chapter, which presents a brief overview of MCDA methods.

This section is followed by a more detailed look at one of the mainstream MCDA methods: the additive multi-attribute value model.

It is important to note that the three stages sequence is not necessarily linear, and often the discussion and learning that occurs within one stage motivates a return to a previous stage in a cyclical improvement process. For instance, if it is difficult to assess performance for one criterion in a way that is independent of other criteria, this may require restructuring the criteria set. New alternatives can also be created based on the insights gained while assessing other alternatives.

The methodology presented is illustrated for a case study with real-world data (although we will consider a fictitious decision maker). This application, introduced in the fourth section of this chapter, concerns an assessment of the environmental performance of dairy farms located in a Portuguese region. A fifth section presents and discusses results, followed by a concluding section.

2. MCDA: A brief overview

This chapter provides an overview of a variety of approaches in MCDA. It is restricted to methods that deal with a discrete and finite set of alternatives that have been listed beforehand. It does not cover another very different type of methods in the broader MCDA field, which are multiobjective optimization methods based on mathematical programming formulations (for the latter, see *e.g.* Ehrgott 2005).

Environmental problems have been a fertile ground for the application of MCDA approaches. The interested reader may find several books and reviews with applications, *e.g.*, (Ehrgott and Stewart 2010; Huang et al. 2011; Linkov and Moberg 2012; Pohekar and Ramachandran 2004; Wang et al. 2009). A few examples are also cited in the following presentation.

2.1. A taxonomy of MCDA approaches

Let A denote the set of alternatives being evaluated. We present a taxonomy based on two dimensions (Table 1): one dimension refers to whether the evaluation of one alternative depends on the other alternatives in A ; the other dimension refers to the type of underlying approach used to synthesize the performances assessed on the individual criteria.

		Does the evaluation of one alternative depend on other alternatives belonging to A?	
		<i>No (evaluation independent of other alternatives)</i>	<i>Yes (evaluation relative to other alternatives)</i>
Underlying approach	<i>Value</i>	I. Global value aggregating individual performances	II. Global value synthesizing comparisons of alternatives in A
	<i>Distance</i>	III. Distance to an externally defined reference	IV. Distance to a reference defined from A
	<i>Binary relations</i>	V. Binary relation between alternative and external references	VI. Binary relation on the alternatives in A

Table 1. A taxonomy of MCDA methods

An intuitive and convenient approach in MCDA consists in deriving a global performance value for each alternative. With such values it is possible to identify the most preferred alternative (the one with highest value), or to rank the alternatives (from highest value to lowest value), or to perform an ordinal classification (each category corresponding to an interval of values).

A global value for an alternative can be the result of an aggregation of the performances of this alternative in a way that is independent of the evaluation of the remaining alternatives being considered (Type I on table 1). A common approach to achieve this is the weighted sum method (Yoon and Hwang 1995), whereby the performances of the alternatives on each criterion are multiplied by the weight of the respective criterion, and then these products are summed. For instance, (Roth et al. 2009) use this method to evaluate environmental, social and economic dimensions of sustainability of electricity supply options for a major Swiss utility. The weighted sum method requires that all the performances are measured on commensurable scale, usually by means of normalization (*e.g.* Roth et al. (2009) apply an affine transformation that assigns, for each criterion, a score of 100 to the alternative which does best on that criterion and a score of 0 to the one which does least well). One must however note that if such scales result from a normalization operation, for instance based on the maximum and minimum performances observed in A , then the evaluation of each alternative will no longer be independent of the other alternatives (*e.g.*, see (Dias and Domingues 2014)). A more rigorous approach to obtain comparable scales is to aggregate value or utility functions, as detailed in Section 3, using multi-attribute value (or utility) theory (MAVT/MAUT) (Keeney 2006). For instance, (Keeney 2004) overviews seven applications of MAVT/MAUT in policy decisions with environmental implications, including managing nuclear waste, air pollution control, energy policy, and wastewater treatment. Other approaches of Type I include the Ordered Weighted Average (Yager 1988) and fuzzy integrals (Grabisch and Labreuche 2010).

A global value for an alternative can be a synthesis of how this alternative compares with other alternatives in A (Type II on Table 1). A popular approach in this category is the Analytic Hierarchy Process / Analytic Network Process (Saaty 2008), which is based on pairwise comparisons of alternatives, one criterion at a time, and pairwise comparisons of the criteria, all to be directly performed by decision makers. As an example, (Ramanathan 2001) discusses the use of AHP in environmental impact assessment, using as a case study a multiple stakeholder assessment of socio-economic impacts for a proposed LPG recovery plant. Albeit very popular, this approach has been subject to much debate in academia about the scale that should be used and about the theoretical properties of the method used to derive priorities (see, *e.g.*, Ishizaka and Labib 2011). Another popular approach is PROMETHEE II, which computes how much is each alternative preferred to each other one, and then aggregates these pairwise preference values (Brans et al. 1986). An example can be found in (Oberschmidt et al. 2010), concerning the use of PROMETHEE to rank heat and power supply alternatives for a municipality in Germany. Ordinal voting functions also fall in this category, *e.g.* the Borda score as well as its cardinal extensions (Bouyssou et al. 2000; Marchant 1996).

Other MCDA methods also derive a performance value for each alternative, but this value is based on a distance to one (or more) points in the multidimensional performance space. Approaches of Type III (Table 1) compute a distance to an externally defined reference. Typically, the decision maker defines the performances desired for each criterion (the “aspiration levels”) and then a distance-based method will choose the closest feasible alternative according to some metric. Any metric can be used: the Manhattan distance (L_1 norm), Euclidean distance (L_2 norm) or Chebyshev distance (L_∞ norm) are usually chosen. Weights can be attached to the different criteria to compute weighted distances. The shorter the distance is, the better the alternative.

Some distance-based methods use a reference point that is based on the set A , and thus belong to Type IV. A well-known example is the TOPSIS method (Yoon and Hwang 1995), which is based on weighted distances to the ideal point (a fictitious alternative with the best performances in all criteria taking into account the performances observed in A) to be minimized, and the distance to the anti-ideal point (a fictitious alternative with the worst performances in all criteria), to be maximized. As an example, (Huang et al. 2010) use TOPSIS to evaluate materials for personal computer cases based on an environmental life-cycle impact assessment.

Finally, there are MCDA methods that do not compute a global value for each alternative, but instead build one or several binary relations. Usually, for each pair of alternatives (a, b), the method will conclude that one is preferred to the other, or they are indifferent (equally

preferable) or they are incomparable (if a and b are too different for the method to assert a preference or indifference relation holds). Such methods are usually known as outranking methods (Roy 1991). It is thus possible to have a relation such that alternative a is preferred to alternatives b and c , but b and c are incomparable, meaning that their strengths and weaknesses are so different that one cannot conclude that b is better than c (or vice-versa), but it is equally unwarranted to conclude that they are indifferent (*i.e.*, similarly preferred).

Outranking methods belong usually to Type VI (Table I), since the binary relation is built on the set A . The simplest binary relation is dominance: an alternative a dominates b if it is better on some criteria and is not worse in any other criterion. It does not require any subjective parameters such as criteria weights, but the relation is usually poor (*i.e.*, it applies to few pairs of alternatives). Outranking methods use additional inputs to enrich this relation. Some methods such as ORESTE, QUALIFLEX and REGIME also use ordinal information (Martel and Matarazzo 2005). Better known approaches use cardinal information. This includes ELECTRE methods (Figueira et al. 2005), among which ELECTRE I is used to isolate a subset of candidates to become the best alternative, whereas ELECTRE II-IV methods derive partial preorders. For instance, (Bollinger and Pictet 2008) applied ELECTRE III to obtain a partial preorder of existing and potential treatment technologies for solid urban waste incineration residues, on behalf of a Swiss federal agency. PROMETHEE I (Brans et al. 1986) and NAIADE (Munda 1995) are other popular methods to build a partial preorder on the alternatives set.

The ELECTRE TRI method builds an outranking relation between an alternative to be classified and a set of reference profiles defining the possible categories; hence, the classification of an alternative does not depend on the other alternatives being evaluated and it belongs to Type V on Table I. ELECTRE TRI has been recently applied, for instance, by (Covas et al. 2013), to classify potential locations for a data center in Portugal taking into account technical, social, economic, and environmental dimensions.

The distinction between a method for which the evaluation of one alternative is independent of other alternatives and a method for which the evaluation is relative to other alternatives can be important. In a method of the former type, adding or removing (or even changing) one alternative may change the relative standings of the remaining alternatives. For instance, it may occur a rank reversal phenomenon, where alternative a is better than alternative b if alternative c is present, but b becomes better than a if c is not considered (Dias and Domingues 2014; Forman and Gass 2001). While some users and academics consider this a flaw, other users and academics sustain that the information brought by a comparison with c is relevant to decide between a and b . In practice, rank reversals will not be a problem if the inclusion of new alternatives is not possible, because of legal reasons (*e.g.*, evaluation of proposals that had to be submitted until a certain deadline), or because the set of alternatives is exhaustive. Even in other situations, for some methods (*e.g.*, weighted sum and TOPSIS) a rank reversal can only occur when adding or removing an alternative that has maximum or minimum value for some criterion.

2.2. How the MCDA parameters are set

MCDA is a tool to make evaluations based on judgment, acknowledging the evaluation is partially subjective, starting with the choice of the alternatives to evaluate and the choice of the evaluation criteria, and ending in the definition of the method's parameters. Depending on the MCDA method, its parameters may define the importance of each criterion (*e.g.*, criteria weights), may define how performances are compared when measurement is not exact (*e.g.*, indifference thresholds), may define how strong must arguments be to support a result, etc.

Traditionally, MCDA methods assume that a decision maker (possibly an entity or a group) sets the values of subjective parameters. In the business world, this may be one or more managers from the client organization that owns the problem. In formal evaluation processes, such as calls for tenders or awarding prizes, a jury or panel of evaluators will act as decision makers. In environmental studies, however, the identification of a decision maker in this sense tends to be avoided.

Environmental MCDA studies are performed not only by companies, but also by academia, non-profit organizations, public institutions and government agencies, which often feel they

should not define the evaluation parameters on their own. Therefore, common alternatives to using the subjectivity of a decision-making entity are using the inputs from experts or from the stakeholders identified for these studies. Expert and stakeholder judgment is sometimes elicited using written questionnaires or interviews, *e.g.*, (Myllyviita et al. 2012), but promoting workshops gathering multiple parties to define a model's parameters is also practiced, *e.g.*, (Brucker et al. 2013).

Social and participatory MCDA approaches (*e.g.*, (Marttunen and Suomalainen 2005; Munda 2004; Mustajoki, Hämäläinen, and Marttunen 2004)) bring together the perspectives of multiple stakeholders into an MCDA model, or several variants of a model if consensus is not achieved. The participants are not necessarily experts, being usually representatives of the persons potentially affected by a decision.

Some environmental MCDA studies address the discomfort of subjectivity by considering not one, but many vectors of parameter values. For instance, a study may use vectors representing different profiles, *e.g.*, (Roth et al. 2009). Another possibility consists in considering a continuous region (with an infinite number of vectors) in the parameter space, possibly constrained by widely accepted bounds, *e.g.*, no criterion has more weight than all other criteria, or using a ranking of criteria weights rather than specific numerical weights (an example appears later in this chapter). Then, it is possible to derive robust conclusions about parts of the result that hold true for all these accepted parameter values. Another strategy consists in computing statistics about the results that may occur when the parameters are randomly drawn from a stochastic distribution, as in SMAA methods (Lahdelma and Salminen 2001; Lahdelma et al. 2002).

When subjective statements need to be elicited, one can also make a distinction between direct and indirect approaches. A direct elicitation will focus the dialogue with the decision maker (or experts or stakeholders) on the value of the parameters. An indirect elicitation will focus the dialogue on the results the method ought to produce. In such approaches, the decision makers make judgments about some alternatives used as examples (*e.g.*, stating that *a* is better than *b*, or stating that *a* belongs to category *C*). Then the method infers parameters that reproduce these partial results as close as possible. This MCDA area has evolved a lot in recent years. As examples of Type I (Table I), UTA approaches (Beuthe and Scannella 2001; Jacquet-Lagrèze and Siskos 2001) ask the decision maker to rank a small set of alternatives. Then, an additive value function model is inferred so that it reproduces as well as possible the ranking provided, allowing for some deviations. An example of Type V is IRIS (Dias and Mousseau 2003), which asks the decision maker to indicate a range of possible categories for some alternatives. Then, it infers an ELECTRE TRI model that fully respects these examples or, if this is not possible, it provides information about which of the examples provided contradict each other.

In the next section of this chapter, one of the MCDA approaches of Type I will be presented in more detail. For details concerning other methods the reader may consult the cited references or a more general book on MCDA, *e.g.* (Belton and Stewart 2002; Bouyssou et al. 2006; Ishizaka and Nemery 2013; Roy 1996).

3. The additive multi-attribute value model

Multi-attribute value theory (MAVT) methods (for details see (Keeney 2006; Keeney and Raiffa 1993)) break down the evaluation of alternatives in two stages, based on a rigorous axiomatic foundation (an analogous foundation – MAUT – exists for problems under probabilistic uncertainty). First, the value of the alternatives is assessed on one criterion at a time by means of a value function. Then, these single-criteria values are aggregated into a global performance value (the higher it is, the better is the alternative). This evaluation is independent of the evaluation of the remaining alternatives being considered (Type I on Table 1).

Before entering the details of this approach, let us introduce some notation:

- $A = \{a_1, \dots, a_m\}$ is a set of m alternatives being evaluated,

- $X = \{x_1, \dots, x_n\}$ is a set of n evaluation criteria (also known as attributes in the context of MAVT).
- x_{ij} denotes the performance of alternative $a_i \in A$ with respect to criterion $x_j \in X$, on its measurement scale (whether it is a direct indicator, an indirect indicator, or a qualitative descriptor).
- $v_j(\cdot)$ is a value function for criterion $x_j \in X$.
- $v_j(x_{ij})$ is the value of the performance of alternative $a_i \in A$ on criterion $x_j \in X$.

3.1. Constructing a value function for one criterion

The first step of MAVT is to build a value function $v_j(\cdot)$ for each criterion x_j . Such function should increase with preference:

- $v_j(x_{aj}) > v_j(x_{bj})$ if and only if performance x_{aj} is considered better than performance x_{bj} on criterion x_j (x_{aj} is preferred to x_{bj});
- $v_j(x_{aj}) = v_j(x_{bj})$ if and only if performance x_{aj} is considered as good as performance x_{bj} on criterion x_j (x_{aj} is indifferent to x_{bj}).

Furthermore, a value function also measures differences of value between performances. The higher the difference, the higher the strength of preference is and the more the decision maker cares about it:

- $v_j(x_{aj}) - v_j(x_{bj}) > v_j(x_{cj}) - v_j(x_{dj})$ if and only if the strength of preference for x_{aj} over x_{bj} is higher than the strength of preference of x_{cj} over x_{dj} on criterion x_j ;
- $v_j(x_{aj}) - v_j(x_{bj}) = v_j(x_{cj}) - v_j(x_{dj})$ if and only if the strength of preference for x_{aj} over x_{bj} is equal to the strength of preference of x_{cj} over x_{dj} on criterion x_j ;

Figure 1 illustrates a few typical examples of value functions. These are increasing value functions, meaning that the higher the performance, the higher the value. If the criterion represents a cost or an undesirable impact such as a pollutant, then it would require a decreasing function. The situation depicted in Figure 1a) (affine function) indicates that a given increase in x_j will always be valued the same, regardless of the initial performance, i.e., an increase from 5 to 10 adds the same value to the decision maker as an increase of, say, 55 to 60. The situation depicted in Figure 1b) (convex function) indicates increasing marginal value: the same increase in performance is more appreciated if performance is higher (an increase from 5 to 10 adds less value than an increase from 55 to 60). Figure 1c) (concave function) indicates decreasing marginal value: the more performance one has, the less one cares about its increase (an increase from 55 to 60 adds less value than an increase from 5 to 10). Figure 1d) illustrates a typical situation when there is a target t : the same increase near this target is valued more than when performances are far from this target.

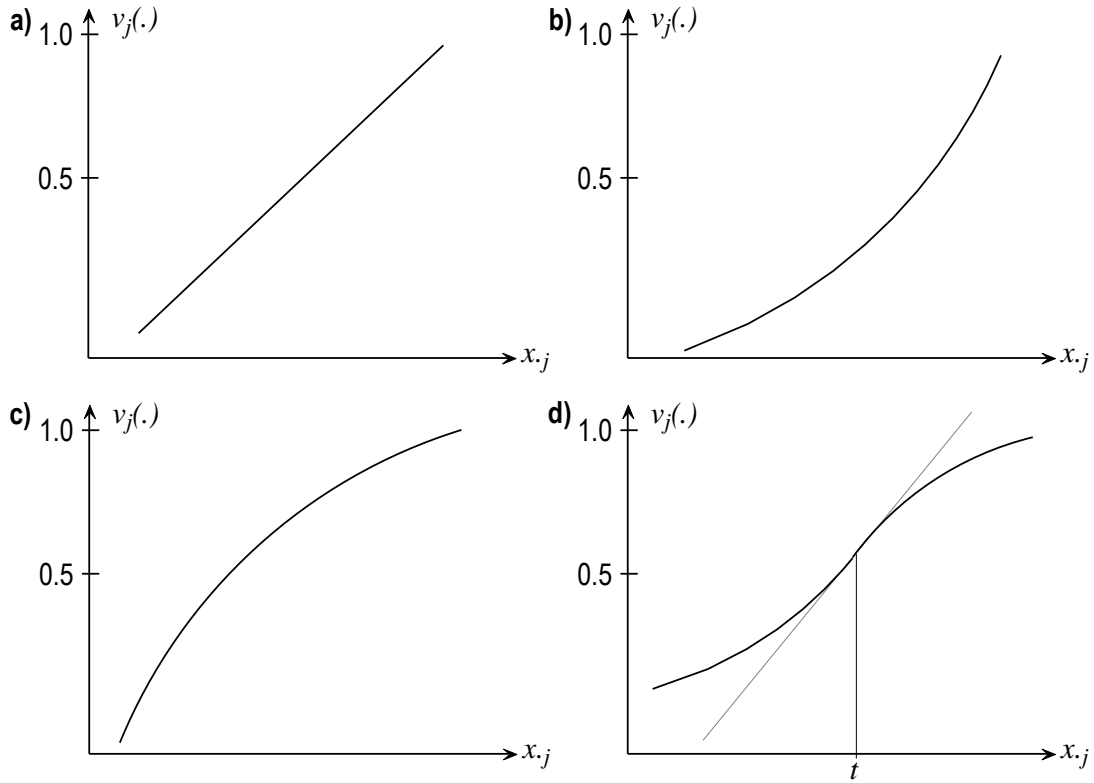


Figure 1. Examples of increasing value functions: a) affine, b) convex, c) concave, d) sigmoidal

The key to elicit a value function is the concept of equal difference in terms of strength of preference (preference intensity). The bisection method (Goodwin and Wright 2014), for instance, starts with two points x_{0j} and x_{1j} for which we can arbitrarily define that $v_j(x_{0j})=0$ and $v_j(x_{1j})=1$. Then the dialogue with the decision maker aims at finding a point x_{cj} such that $v_j(x_{cj})=0.5$, *i.e.*, a central point bisecting the initial interval in two. This dialogue usually proceeds by trial and error.

Suppose $x_{0j}=100$, $x_{1j}=500$, and we tentatively try $x_{cj}=300$. To find out whether $v_j(x_{cj})-v_j(x_{0j}) = v_j(x_{1j})-v_j(x_{cj})$, one can ask the decision maker what sacrifice would he or she be willing to make to increase the performance from 100 to 300. This sacrifice might be a cost, or might be a loss on some other criterion. Then, one would ask the decision maker to assume the performance was 300 (the previous sacrifice was not needed) and ask what sacrifice would he or she be willing to make to increase the performance to 500. If the answer is the same, then the x_{cj} can be considered equal to 300: an increase of 200 is equally valued whether the starting point is 100 or 300, suggesting an affine function (Figure 1a). If, on the contrary, the decision maker would now be less willing to make the same sacrifice this would mean that $v_j(x_{cj})-v_j(x_{0j}) > v_j(x_{1j})-v_j(x_{cj})$. Therefore, one might tentatively try a higher central value, *e.g.*, $x_{cj}=350$. Once a performance x_{cj} such that $v_j(x_{cj})-v_j(x_{0j}) = v_j(x_{1j})-v_j(x_{cj})$ has been found, one can deduce $v_j(x_{cj})=0.5$. Afterwards, one may proceed in a similar way to bisect the value interval $[0, 0.5]$ (searching for a performance with value 0.25) and to bisect the value interval $[0.5, 1]$ (searching for a performance with value 0.75).

This type of elicitation aims at building value functions that the decision maker perceives as adequate to represent his or her judgment about strengths of preference. A precise value measurement is often illusory in practice since there are not stable numerical answers stored in the decision maker's brain. The MACBETH (Bana e Costa et al. 2012) approach allows deriving value functions in an interactive way from qualitative assessments of preference intensities.

3.2. Aggregation of value functions

The second step of MAVT is to compute a global value $v(a_i)$ for each alternative $a_i \in A$ by aggregating the value this alternative has on the multiple value functions:

$$v(a_i) = f(v_j(x_{i1}), v_j(x_{i2}), \dots, v_j(x_{in})) \quad (1)$$

There are multiple formulas to perform this aggregation (Keeney and Raiffa 1993), but the one which is easiest to understand and most used in practice is the additive aggregation model:

$$v(a_i) = k_1 v_j(x_{i1}) + k_2 v_j(x_{i2}) + \dots + k_n v_j(x_{in}) \quad (2)$$

where k_1, \dots, k_n are scaling constants (or scaling weights, or simply weights) that reflect the tradeoffs between the different value functions. All these weights are positive and by convention typically $k_1 + k_2 + \dots + k_n = 1$.

The use of an additive value function assumes a strong independence among the different criteria. This additive independence condition means that if the decision maker compares two alternatives that have equal performance on some of the criteria, then the common performance levels can be ignored. In other words, the preference intensity for one alternative over some other depends only on the criteria in which their performances differ. This may not always be the case, namely if there are positive or negative synergies among the criteria. When the independence condition does not hold, it is advisable to return to the structuring phase and rethink the set of criteria (Belton and Stewart 2002), *e.g.*, in a habitat preservation problem the amount of prey and the amount of predators could hardly be two different criteria, but can be replaced by a criterion assessing the predator/prey balance.

Eliciting values for the scaling weights k_1, \dots, k_n is a crucial and often misunderstood step in this method. The most important aspect to keep in mind is that these parameters do not represent *directly* the importance of the criteria. If for instance $k_p = k_q$ it is not possible to conclude that criterion x_p (*e.g.*, cost) is as important as criterion x_q (*e.g.*, number of trees lost). One can only say that an increase of δ in value function v_p is considered as good as a similar increase in value function v_q , since:

$$k_p = k_q \Leftrightarrow k_p[v_p(x_p) + \delta] + k_q v_q(x_q) = k_p v_p(x_p) + k_q[v_q(x_q) + \delta] \quad (3)$$

Possibly, if an increase of δ in v_p represents one billion dollars and an increase of δ in v_q represents 10 trees then $k_p > k_q$, but if an increase of δ in v_p represents one thousand dollars and an increase of δ in v_q represents 1000 trees then probably $k_p < k_q$.

A ranking of the scaling weights can be obtained by asking a decision maker to compare equal improvements in all value functions: if he or she could make an improvement of δ in a value function, which one would be chosen? And if he or she could improve a second value function, which one would it be? Etc. The relative values for k_1, \dots, k_n can then be obtained by asking the decision maker to attribute a quantitative rating to each one of these improvements. This elicitation approach is known as the swings method (Goodwin and Wright 2014).

The weights can also be elicited by means of trade-off questions. Each question is an equation of the type:

*Having a performance of p on x_p and a performance of q on x_q
is as good as*

Having a performance of p' on x_p and a performance of q' on x_q

where p' is a value different than p , and q' is a value asked to the decision maker, to assess the acceptable sacrifice (or gain) on x_p that would compensate an improvement (or a loss, respectively) on x_q . From this answer it is possible to derive a trade-off ratio k_p/k_q :

$$k_p[v_p(x_p) + \delta] + k_q v_q(x_q) = k_p v_p(x_p) + k_q[v_q(x_q) + \delta'] \Leftrightarrow k_p/k_q = \delta'/\delta \quad (4)$$

By asking the decision maker to make other trade-offs a system of equations for k_1, \dots, k_n can be solved.

The elicitation of “weights” by means of trade-off questions highlights the compensatory nature of this MCDA method. Some authors (Munda 2005) note that it may not be legitimate to

have a very good performance on one criterion hiding a very bad performance on some other criterion. Therefore, to be cautious, alternatives with unacceptable performance on any criterion should not be considered in A . This method can also be criticized for the amount of information it requires and the possibly difficult questions that are posed to the decision maker, but this can also be seen as an advantage as it invites the decision makers to think hard about their preferences (Keeney and Raiffa 1993).

An aspect to keep in mind is that the role of an analyst is to build a model together with the decision maker so that he or she feels comfortable about it. Approximate answers are acceptable and usually the result is quite robust to small imprecisions, including violations of additive independence (Stewart 1996).

The VIP Analysis approach (Dias and Clímaco 2000), which will be illustrated in this chapter, is based on the additive MAVT model, but it is able to work with partial information about the scaling weights. Thus it allows an analysis to be performed without asking for specific figures for these parameters. The partial information can be defined as intervals for the weights, bounds for a ratio of weights, a ranking of weights, or any other information that can be translated as a set of linear constraints. These constraints define implicitly a region K in the space of parameter values. With this information, VIP Analysis computes bounds for the value of alternatives and for the differences of value between alternatives.

4. Presentation of the case study

4.1. Context

To illustrate the additive aggregation method for value functions, introduced in the previous section, a case study based on real-world data is presented next. Although data about the alternatives being evaluated are real, this illustration considers hypothetical preferences of a potential decision maker. This decision maker intends to evaluate the environmental performance of dairy farms in the Entre Douro e Minho region of Portugal, an important facet for their sustainability. The reader is also referred to (van Calster et al. 2006), who also applied MAUT to assess the sustainability of dairy farming systems, including economic, social and ecological aspects, considering four experimental dairy farms as alternatives.

Data from the survey developed during the implementation of the Entre Douro e Minho's Regional Development Plan for Dairy Farming (DRAEDM 2007) are used. Although this survey gathered data from over 1700 farms, for this case study only one parish is included: the parish of Cambeses, located in the municipality of Barcelos.

The environmental performance of dairy farms deals with its impacts on soil, water and air, but also encompasses other issues, namely animal well-being. The present case study is not based on a complete life-cycle inventory due to lack of data, but used indirect indicators to assess potential negative impacts. These indicators are related in particular to the conditions each farm has relatively to the number of animals it has. The set of criteria is described in the following paragraphs and is summarized on Table 2.

4.2. Evaluation criteria (indicators)

The indicators selected for this illustration were defined by experts from the problem domain. The first criterion concerns the farm's capacity to store manure produced by its animals. The chosen indicator is based on the total storage capacity of effluents (m^3) and the number of standard livestock units (LU). The use of the LU standard allows aggregating animals of different types, taking into account the animal species, age, weight and productive vocation, when accounting for food needs and production of livestock effluent. According to the regulations for licensing bovine farming in Portugal, the required total storage capacity of effluent is:

- $6 m^3$ for LU, for farms with separation system (whether mechanical, chemical or physical) of solid and liquid fraction.
- $7 m^3$ for LU, for farms without such separation.

Therefore, indicator x_1 (Table 2) is equal to 1.0 if the farm meets this requirement exactly. A higher ratio means the capacity exceeds the requirement and a lower ratio means the capacity is below the requirement (*e.g.*, $x_1=0.8$ means the farm only has 80% of the capacity it should have and thus it is not self-sufficient and it is more likely to have undesirable impact on the environment).

The second criterion concerns the farm's capacity to incorporate the nitrogen (N) produced in the soil, taking into accounts its Land Area Under Agricultural Production (LAUAP). The Wastewater Management Plan for the region of this case study indicates there should be no more than 5.6 LU per ha of LAUAP (or 2.8 LU per ha in case if the farm is located in a vulnerable area, which is not the case for the Cambeses parish). Indicator x_2 (Table 2) is equal to 1.0 if the farm meets this requirement exactly. For this indicator lower values are better: a value higher than 1.0 means there is an excessive number of LU for the farm's capacity to incorporate organic nitrogen in the soil by applying the resulting effluent on suitable land.

The third criterion addresses the risk of contaminating water lines. The chosen indicator, x_3 (Table 2), consists of the ratio between the number of storage points less than 25 m away from the water lines and the total number of existing storages on the farm. Ideally it should be zero (the less, the better). As this ratio increases, it becomes more likely that the farm will contaminate water.

The fourth criterion addresses the use of water. The chosen indicator, x_4 (Table 2), consists of the proportion of the farms' storage facilities that have individualized collection of rainwater (the higher the better). Rainwater collection and drainage in such facilities increases the effectiveness of storage capacity, provides fertilizers of better quality, reuses water, and thereby reduces the water burden.

The fifth criterion concerns animal well-being, assessed by considering the ratio between the covered area (stall) and the number of cows on the farm (x_5 , Table 2). The lower this indicator is, the greater the number of animals per square meter in the stable. If there is an excessive concentration of animals in a given area, the welfare is at stake and milk production may even be lower.

The reader may notice the absence of criteria reflecting other very important concerns about the performance of dairy farming, such as greenhouse gas (GHG) emissions. This case study is focused on the dairy farming activity, and GHG emissions can be considered proportional to the number of LU. A large farm with many LU will thus have twice the emissions compared with a farm with half of the LU. But this does not mean that a larger dairy farm has worse environmental performance than a smaller one. An indicator taking GHG emissions into account would be quite relevant, however, if dairy farms were being compared with other types of farms (*e.g.*, poultry farming).

<i>Criterion ID</i>	<i>Criterion calculation</i>
x_1 : Storage capacity of manure	$\begin{cases} \left(\frac{\text{Storage Capacity (m}^3\text{)}}{LU \times 7(\text{m}^3)} \right), & \text{if there is no separation system} \\ \left(\frac{\text{Storage Capacity (m}^3\text{)}}{LU \times 6(\text{m}^3)} \right), & \text{if there is a separation system} \end{cases}$
x_2 : Number of livestock units per area of manure application	$\frac{LU}{5.6 (\text{ha}^{-1}) \times LAUAP (\text{ha})}$
x_3 : Storage structures near water lines	$\frac{\text{Number of storages at less than 25m of water lines}}{\text{Number of storages}} \times 100\%$
x_4 : Individualized collection of rainwater	$\frac{\text{Number of storages with the Rainwater Separation}}{\text{Number of storages}} \times 100\%$
x_5 : Animal well-being	$\frac{\text{Covered Stall area (m}^2\text{)}}{\text{number of cows}}$

Table 2. Criteria (environmental indicators)
(LU: Livestock Units; LAUAP: Land Area Under Agricultural Production)

4.3. Value functions

Value functions, as described in Section 3.1, indicate how performances in the criteria x_1, \dots, x_5 are translated into value units, respecting the concept of strength of preference. For criteria x_1, x_4 and x_5 , the value function is increasing (the higher the indicator, the better). For criteria x_2 and x_3 , the value function is decreasing (lower values are preferred). The shape of the value functions depends on the assessment of experts or a decision maker. For this case study, plausible, yet hypothetical, functions are presented.

The value functions are depicted in Figure 2 (a)-(e). In these value functions there are two value levels with a common meaning to allow comparability of improvements. The value level 1.0 indicates what is considered a good performance for the indicator, compatible with considering the farm as environmentally responsible. The value level 0.5 indicates what is considered a mediocre value for the indicator. Thus a value level below 0.5 reflects an indicator level of poor environmental performance.

The value function for *Storage capacity of manure* (x_1) considers that performance is good if the regulations for licensing bovine farming in Portugal are met (6 m^3 or 7 m^3 of storage per LU, for farms with separation system or without, respectively). Having more capacity available still adds some value. Below $x_1=1$ value decreases rapidly. If the farm has less than half of the capacity it should have than its performance is considered poor .

The value function for *LU per LAUAP* (x_2) considers that performance is good when there are 5.6 or less LU per ha of LAUAP. If the number of LU is less ($x_2 < 1$) value increases but with a lower slope. Performance is poor if there are 7 or more LU per ha of LAUAP (25% in excess relatively to the limit). From 11.2 LU per ha onwards the slope is lower because it is already evident that the farm will be incapable of applying most of its manure to the soil.

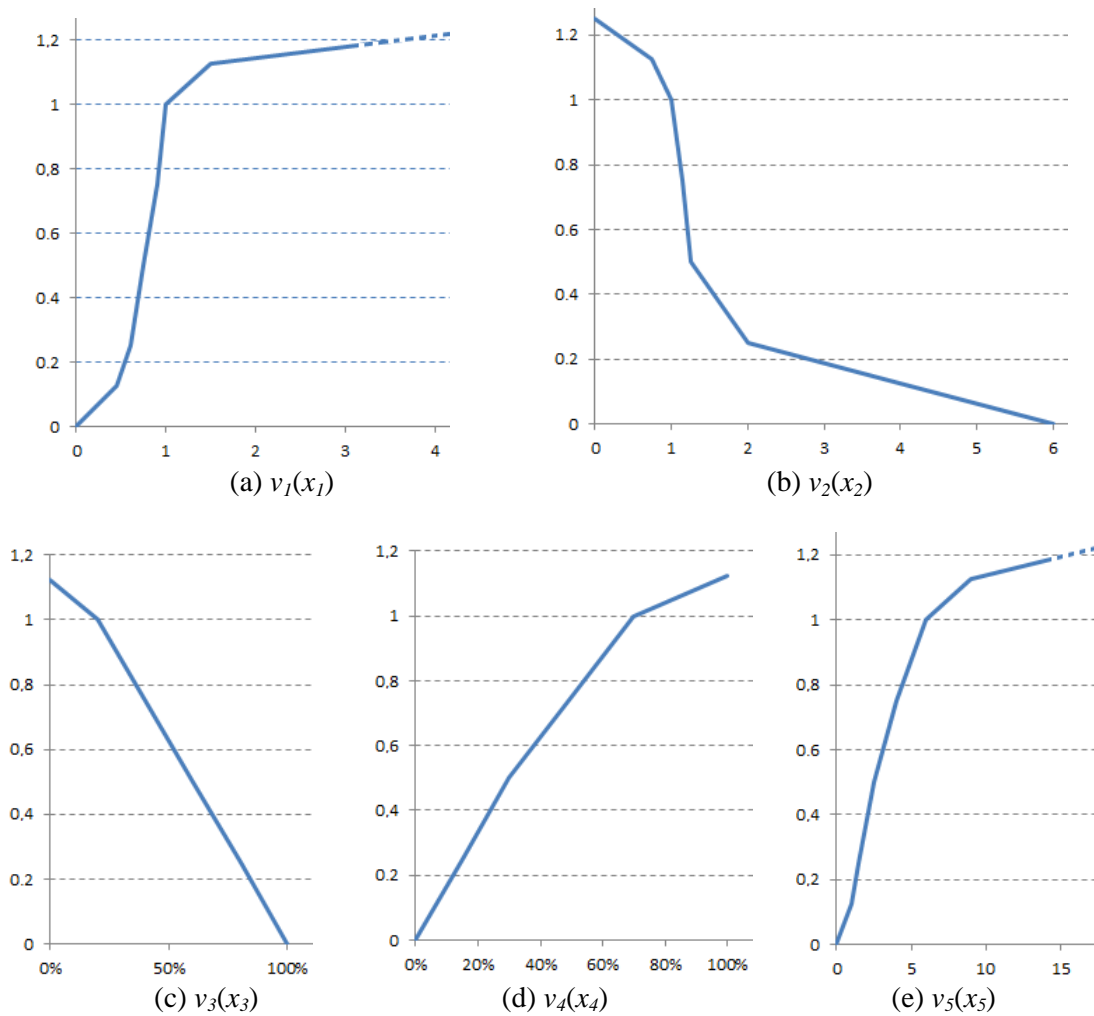


Figure 2. Value functions for the five indicators

The rationale for the remaining value functions is similar. The value function for *Storage structures near water lines* (x_3) considers that value 1.0 is achieved if the proportion is lower than 20%, assuming that these storages would be used only when the remaining storages would have no further capacity, and value 0.5 is achieved if 60% of the storages are near water lines. The value function for *Individualized collection of rainwater* (x_4) considers that value 1.0 is achieved if 70% of the storages have rainwater separation, whereas value 0.5 corresponds to only 30% of the storages having rainwater separation. The value function for *Animal well-being* (x_5) considers that value 1.0 is achieved if the covered stall area reaches 7 m² per cow. Much higher areas do not add much value, but as the area decreases below this threshold, value decreases rapidly. Value 0.5 corresponds to a covered stall area of only 2.5 m² per cow.

4.4. Farms under evaluation

This case study analyses the ten dairy farms located in the parish of Cambeses. This parish belongs to the municipality of Barcelos, the municipality with the highest number of dairy farms in the North of Portugal. This particular parish was selected due to the diversity of farms it contains, but any other set of farms can be analyzed using the approach described in this chapter.

Using data from the Entre Douro e Minho's Regional Development Plan for Dairy Farming (DRAEDM 2007) survey it is possible to compute the selected indicators, as presented in the left columns of Table 3. The indicator values can then be converted into value units according to the value functions described in the previous section (rightmost columns of Table 3).

<i>Farm</i>	$x_1 \uparrow$	$x_2 \downarrow$	$x_3 \downarrow$	$x_4 \uparrow$	$x_5 \uparrow$	$v_1(x_1)$	$v_2(x_2)$	$v_3(x_3)$	$v_4(x_4)$	$v_5(x_5)$
a_1	0.81	1.04	0%	100%	2.13	0.60	0.93	1.13	1.13	0.41
a_2	0.38	1.28	0%	83%	10.00	0.10	0.49	1.13	1.06	1.14
a_3	1.11	1.35	0%	100%	6.43	1.03	0.47	1.13	1.13	1.02
a_4	0.91	0.76	33%	100%	7.69	0.77	1.12	0.83	1.13	1.07
a_5	1.66	0.83	0%	100%	9.13	1.13	1.08	1.13	1.13	1.13
a_6	0.69	1.32	0%	100%	5.45	0.39	0.48	1.13	1.13	0.93
a_7	0.38	5.89	0%	100%	5.56	0.10	0.01	1.13	1.13	0.94
a_8	1.98	0.83	100%	100%	5.83	1.14	1.08	0.00	1.13	0.98
a_9	0.69	0.85	0%	50%	7.86	0.39	1.08	1.13	0.75	1.08
a_{10}	1.61	0.71	100%	100%	6.82	1.13	1.13	0.00	1.13	1.03

Table 3. Indicators and single-criterion value for the 10 dairy farms in Cambeses.
 \uparrow : Indicator to maximize (the higher the better); \downarrow : Indicator to minimize

5. Results and discussion

Given the (single-criterion) values on the rightmost columns of Table 3, the additive model presented in Section 3.2 yields a global value by performing the weighted sum aggregation – Equation (2). This requires setting the scaling weights k_1, \dots, k_5 associated with the value functions. In this case, if two value functions have the same scaling constant, *e.g.*, $k_1 = k_2$, this would mean that an increase from 0.5 to 1.0 in the first value function would be valued as much as an increase from 0.5 to 1.0 in the second value function. As explained in Section 3.2, the weights could be elicited from a decision maker by asking her or him to compare of 0.5→1.0 swings in different criteria. Alternatively, trade-off questions could be asked. If an increase of 0.25 value units in the first value function is valued as much as an increase 0.5 value units in the second value function, this would mean that $k_1 = 2k_2$.

It is also possible to look for conclusions that are robust for a wide range of scaling weights. The MCDA software VIP Analysis (Dias and Clímaco 2000), for instance, can be used to compute the maximum and minimum value that an alternative can achieve, based on solving a linear program to maximize (or minimize) Equation (2) subject to constraints on k_1, \dots, k_5 (such constraints as usually named *partial information*)

Figure 3 depicts the maximum and minimum value achieved by each alternative when constraints bound the maximum ratio of weights to an order of magnitude: $k_i \leq 10k_j$, for all pairs (i, j) . This bound implies that an increase of δ in one value function cannot be worth more than an increase of $10 \times \delta$ in some other value function. Besides the ten farms, Figure 3 also presents three reference values: 0.5 (poor environmental performance, hindering environmental sustainability), 0.75 (an intermediate reference) and 1.0 (good environmental performance, compatible with an environmentally responsible farm). It is already possible to observe that farm a_5 is quite good, farm a_4 is always above 0.75, and farms a_3, a_9, a_1 and a_6 are always above 0.5.

VIP Analysis can also perform a pairwise comparison of the alternatives, computing the maximum difference of value between two alternatives (Figure 4). For instance, the (a_1, a_2) cell indicates that $\max\{v(a_1) - v(a_2)\} = 0.378$, *i.e.*, in a direct comparison a_1 can beat a_2 by a difference of 0.378 units of global value for some vector (k_1, \dots, k_5) respecting the maximum ratio of 10. But the (a_2, a_1) cell indicates that a_2 can also beat a_1 (by a difference of 0.448 units of global value) for some other vector that also respects the same constraints. Therefore, without more bounds on the weights it is not possible to tell which one has more value. Some cells have a negative value, *e.g.*, $\max\{v(a_6) - v(a_3)\} = -0.019$. This means that farm a_6 can never have more value than farm a_3 for weights respecting the maximum ratio of 10. Although the value ranges overlap (Figure 3), when a vector (k_1, \dots, k_5) yields a high value for a_6 , it yields an even higher value for a_3 .

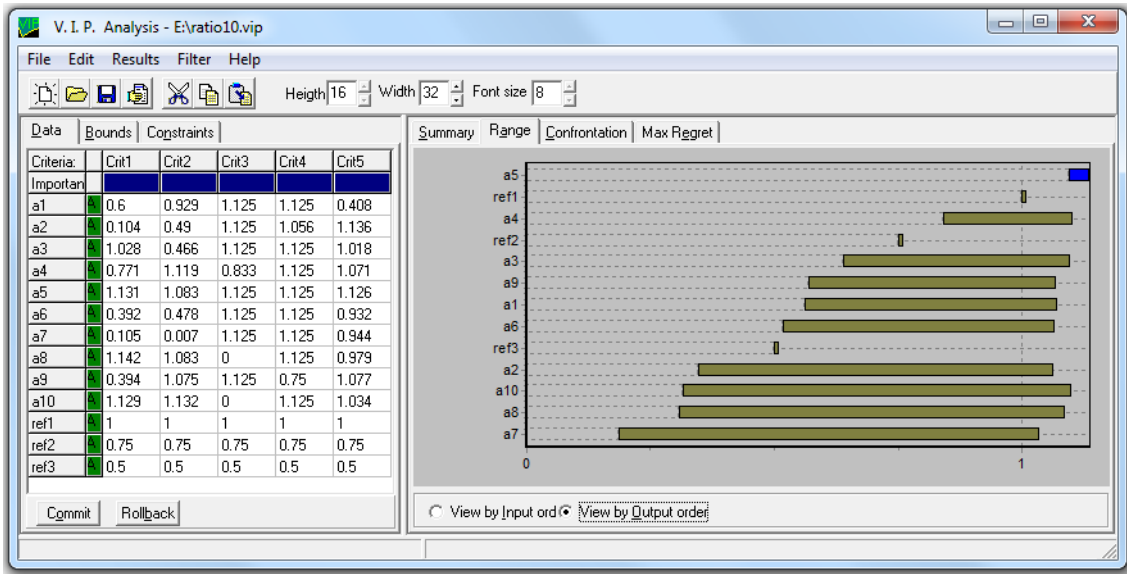


Figure 3. Maximum and minimum value if the maximum ratio of weights is 10

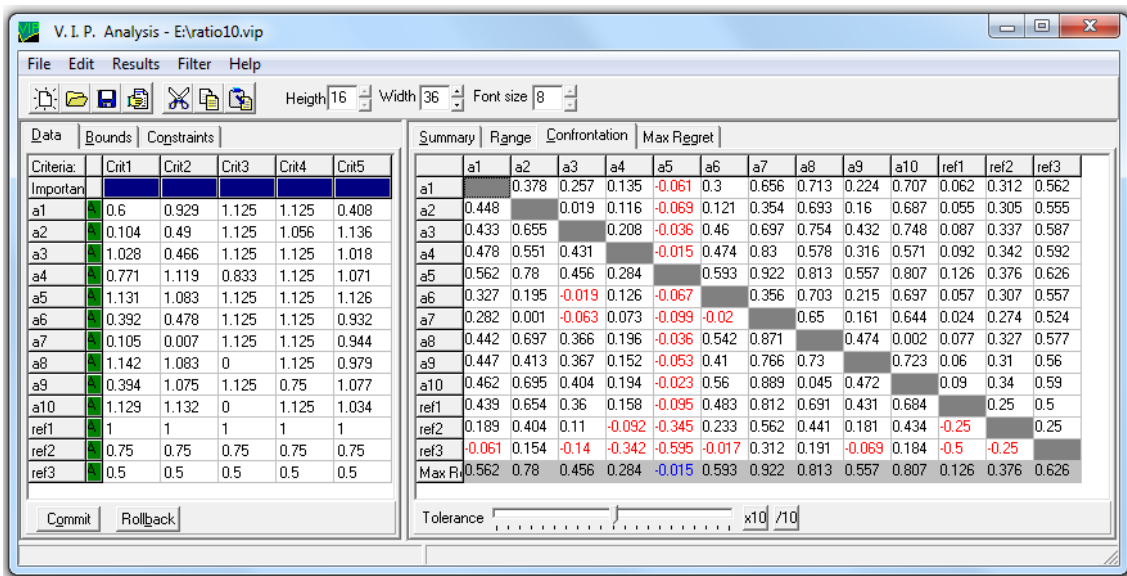


Figure 4. Maximum value differences if the maximum ratio of weights is 10. Negative values mean the row alternative cannot attain more value than the column alternative

A decision maker might consider that a maximum ratio of 10 is already large enough. A maximum ratio of 3 might already be considered acceptable, stating that $k_i \leq 3k_j$, for all pairs (i,j) . This means that an increase of 0.25 value units in one value function cannot be worth more than an increase of 0.75 value units in some other value function.

A decision maker might also be able to compare 0.5→1.0 swings, stating for which value function he or she would choose to increase value from 0.5 to 1.0, which value function would be the second choice, and so on, but without attaching a numerical rating to such swings. As an illustration, consider that the decision maker has the preferences presented in Table 4. These preferences would imply that $k_1 \geq k_2 \geq k_3 \geq k_4 \geq k_5$. Indeed, if the 0.5→1.0 swing in $v_i(\cdot)$ is preferred to a 0.5→1.0 swing in $v_j(\cdot)$, then by Equation (2):

$$k_i \times 1.0 + k_j \times 0.5 \geq k_i \times 0.5 + k_j \times 1.0 \Leftrightarrow k_i \geq k_j.$$

Table 5 summarizes results for a maximum ratio of 10 (corresponding to Figure 3) and a maximum ratio of 3, with and without the ranking constraints $k_1 \geq k_2 \geq k_3 \geq k_4 \geq k_5$. As would be expectable, the value ranges for a maximum ratio of 3 are narrower than those for a maximum ratio of 10, and the value ranges when the ranking constraints are added are narrower than the value ranges when the ranking constraints are not considered.

Ranking	Function	Value swing 0.5 → 1.0 corresponds to improvement:
First choice	v_1	0.75 → 1.00 (of target storage capacity per LU)
Second choice	v_2	1.25 → 1.00 (of target LU per LAUAP)
Third choice	v_3	0.60 → 0.20 (of storage structures near water lines)
Fourth choice	v_4	0.30 → 0.70 (of storages with collection of rainwater)
Fifth choice	v_5	2.50 → 6.00 (of covered stall area in m ² per cow)

Table 4. Comparison of 0.5 → 1.0 swings, implying $k_1 \geq k_2 \geq k_3 \geq k_4 \geq k_5$

	Farm	Maximum ratio of 10	Maximum ratio of 3
Only bounds	a1	[0.561, 1.062] (0.812)	[0.689, 0.965] (0.827)
	a2	[0.346, 1.055] (0.701)	[0.567, 0.959] (0.763)
	a3	[0.640, 1.087] (0.864)	[0.813, 1.029] (0.921)
	a4	[0.842, 1.092] (0.967)	[0.903, 1.050] (0.977)
	a5	[1.095, 1.126] (1.111)	[1.108, 1.123] (1.116)
	a6	[0.517, 1.057] (0.787)	[0.644, 0.950] (0.797)
	a7	[0.188, 1.024] (0.606)	[0.392, 0.881] (0.637)
	a8	[0.309, 1.077] (0.693)	[0.618, 1.003] (0.811)
	a9	[0.569, 1.060] (0.815)	[0.744, 0.998] (0.871)
	a10	[0.316, 1.090] (0.703)	[0.631, 1.020] (0.826)
Bounds and ranking	a1	[0.685, 0.932] (0.809)	[0.770, 0.903] (0.837)
	a2	[0.346, 0.782] (0.564)	[0.567, 0.782] (0.675)
	a3	[0.792, 1.001] (0.897)	[0.861, 0.974] (0.918)
	a4	[0.847, 0.984] (0.916)	[0.923, 0.984] (0.954)
	a5	[1.109, 1.126] (1.118)	[1.113, 1.122] (1.118)
	a6	[0.517, 0.810] (0.664)	[0.644, 0.810] (0.727)
	a7	[0.188, 0.661] (0.425)	[0.392, 0.661] (0.527)
	a8	[0.761, 1.059] (0.910)	[0.798, 0.975] (0.887)
	a9	[0.569, 0.884] (0.727)	[0.744, 0.884] (0.814)
	a10	[0.774, 1.077] (0.926)	[0.813, 0.994] (0.904)

Table 5. Value intervals subject to constraints on the scaling weights (values in parentheses correspond to the midpoint of the value interval)

In practice, an entity evaluating these farms could be interested in selecting one alternative (*e.g.*, to award a prize), or to rank the alternatives, or to sort the alternatives among performance classes (*e.g.*, class A corresponds to a value of 1 or higher, class B to values between 0.75 and 1, class C to values between 0.5 and 0.75, and class D to a value below 0.5).

This entity could wish to evaluate farms with or without information on the relative importance (weight) of the value functions. Without such information, a decision maker could define the maximum ratio among weights that would be acceptable, imposing a limit of 10, or 3, as illustrated, or some other bound. Additionally, the decision maker could define a ranking of these weights, either $k_1 \geq k_2 \geq k_3 \geq k_4 \geq k_5$ or some other ranking. This leads to value intervals as

presented in Table 5. Another option is to elicit specific values for the weights by rating swings or through trade-off questions.

If no specific scaling weights are elicited and a result similar to Table 5 is obtained, the conclusions of the study can be based on different aspects of the results. Consider for instance the results for a maximum ratio bounded to 3 and with ranking constraints (lower right part of Table 5). Table 6 presents some conclusions based on different options.

The first row indicates which conclusions are robust, *i.e.*, which are true for all vectors (k_1, \dots, k_5) than satisfy the maximum ratio and ranking constraints. Clearly a_5 is the best alternative, since its minimum value is higher than the maximum value of any other farm. The analysis of maximum value differences (not depicted, but analogous to Figure 4) yields a partial ranking of the alternatives. The value ranges ensure six of the farms have a precise classification, but three farms could either be in class B or C (depending of the weights) and one farm could be in class C or D.

If crisp rankings and classifications are sought, then they must be based on a specific weights vector. The second row of Table 6 presents the conclusions considering the worst case for each farm (choosing for each farm the weights that yield lower global value). This corresponds to a perspective of “minimum assured value”. The third row presents results considering the best case (maximum value) for each farm, corresponding to a “benefit of doubt” perspective (Cherchye et al. 2007). The fourth row considers the midpoint obtained by averaging the best and worst cases.

If specific weights are elicited from the decision maker then the result will also be a complete ranking of the alternatives or a crisp classification. This requires asking the decision maker to provide ratings for the swings, or to answer trade-off questions, as described in Section 3.2. If, for instance, $k_1=0.293$, $k_2=0.236$, $k_3=0.192$, $k_4=0.155$ and $k_5=0.124$, then the ranking and the classification of the farms would be the same as the last row on Table 6 (equal to the case of midpoint value). These specific weights correspond to the centroid of the set of weights that comply with the maximum ratio and ranking constraints. Such centroid weights have been shown to constitute a good approximation when the decision maker does not provide a specific vector (Sarabando and Dias 2009).

<i>Part of the result</i>	<i>Best farm</i>	<i>Ranking of the farms</i>	<i>Classification of the farms</i>
Robust conclusions	a_5		a_5 : A $a_1, a_3, a_4, a_8, a_{10}$: B a_2, a_6, a_9 : B or C a_7 : C or D
Minimum value	a_5	$a_5 > a_4 > a_3 > a_{10} > a_8 > a_1 > a_9 > a_6 > a_2 > a_7$	a_5 : A $a_1, a_3, a_4, a_8, a_{10}$: B a_2, a_6, a_9 : C a_7 : D
Maximum value	a_5	$a_5 > a_{10} > a_4 > a_8 > a_3 > a_1 > a_9 > a_6 > a_2 > a_7$	a_5 : A $a_1, a_2, a_3, a_4, a_6, a_8, a_9, a_{10}$: B a_7 : C
Midpoint value	a_5	$a_5 > a_4 > a_3 > a_{10} > a_8 > a_1 > a_9 > a_6 > a_2 > a_7$	a_5 : A $a_1, a_3, a_4, a_8, a_9, a_{10}$: B a_2, a_6, a_7 : C

Table 6. Summary of conclusions concerning best choice, ranking and classification (Class A: [1.0,1.25], Class B: [0.75,1.0], Class C: [0.5,0.75[, Class D: [0.0,0.5[)

6. Conclusion

This chapter provided an overview of the methodology of MCDA for environmental performance assessment and illustrates it for a case study. MCDA can be used to evaluate alternative products, processes, locations, facilities, etc., explicitly considering multiple evaluation criteria. Alternatives are evaluated on a single dimension at a time, and then these evaluations are aggregated to provide a synthesis of how the alternatives perform. The described panorama of MCDA is quite varied, but only one approach is developed in more detail and illustrated here: the additive value function. It is a popular approach that can be used for choice, ranking, or classification problems.

The case study concerns the assessment of environmental performance of Portuguese dairy farms, choosing as an illustration the ten farms located in the Cambeses parish of Barcelos. Based on a farm characterization database and the advice of domain experts, five criteria are taken into account for the evaluation.

The use of the additive value function requires parameters reflecting the subjective preferences of a decision maker, as do most MCDA methods. For environmental performance assessment (or, more generally, for sustainability assessment), however, the decision maker will hardly be a single individual. Entities carrying out such an assessment commonly seek that preference-related parameters, such as criteria weights, result from a consensus among experts and/or stakeholders. They may also seek to avoid setting subjective parameters. Sometimes such entities choose to use “equal weights” to avoid subjective statements. However, equal weights already have an implicit subjective judgment in most methods. In the additive value function, these mean that an increase of δ in one value function is as valued (subjectively) as an increase of δ in another value function.

A sounder option to avoid setting subjective parameters is to use methods that deal with partial information, such as the VIP Analysis illustrated in this chapter. It is possible to set constraints that are widely acceptable among stakeholders and experts, such as a maximum ratio among weights of 10 (or even 3). If there is a consensus on other constraints such as a ranking of the weights, this can also be considered.

In this case study, using a maximum ratio of 3 and a ranking of the weights was already sufficient to provide a rich set of robust conclusions about the ranking and the classification of dairy farms. If less information is used (*e.g.*, other cases in Table 5), a perspective of “minimum assured value” or a perspective of “benefit of doubt” can be still be used to rank or sort the alternatives. The first perspective will highlight the weak aspects of each alternative, favoring those that do not present poor performances on any criterion. The second perspective corresponds to letting each alternative choose the weights that put it under the best possible light, thereby allowing some freedom to pursue very good performance on some criteria even if performance on some other criteria can be quite poor.

The flexibility of MCDA and the way it makes the assessment transparent, inviting stakeholder participation, has contributed to the increasing adoption of this methodology in many organizations, private and public, including government (*e.g.*, the UK Department for the Environment, Transport and the Regions has commissioned an MCDA manual to provide the central government guidance on the application of these techniques (Dodgson et al. 2000)).

Acknowledgements

This work is partially supported by Portuguese Foundation for Science and Technology under project grants PEst-OE/ EEI/UI308/2014 and COMPETE PTDC/SEN-TRA/117251/2010-FCOMP-01-0124-FEDER-021495, as well as project EMSURE (CENTRO 07-0224-FEDER-002004).

References

- Bana e Costa, C.A., J.-M. De Corte, and J.-C. Vansnick (2012), 'MACBETH.' *International Journal of Information Technology & Decision Making* 11 (02): 359–387.
- Belton, Valerie, and Theodor Stewart (2002), *Multiple Criteria Decision Analysis. An Integrated Approach*. Boston: Kluwer.
- Beuthe, M., and G. Scannella (2001), 'Comparative Analysis of UTA Multicriteria Methods.' *European Journal of Operational Research* 130 (2) (April): 246–262.
- Bollinger, D., and J. Pictet (2008), 'Multiple Criteria Decision Analysis of Treatment and Land-Filling Technologies for Waste Incineration Residues.' *Omega* 36 (3) (June): 418–428.
- Bouyssou, D. (1993), 'Décision Multicritère Ou Aide Multicritère.' *Newsletter of the European Working Group 'Multicriteria Aid for Decisions', Series 2, Spring*.
- Bouyssou, Denis, Thierry Marchant, Marc Pirlot, Patrice Perny, and Alexis Tsoukiàs (2000), *Evaluation and Decision Models: A Critical Perspective*. Kluwer.
- Bouyssou, Denis, Thierry Marchant, Marc Pirlot, Alexis Tsoukiàs, and Philippe Vincke (2006), *Evaluation and Decision Models with Multiple Criteria*. Berlin and Heidelberg: Springer.
- Brans, J.P., Ph. Vincke, and B. Mareschal (1986), 'How to Select and How to Rank Projects: The Promethee Method.' *European Journal of Operational Research* 24(2), 228–238.
- Brucker, K. De, C. Macharis, and A. Verbeke (2013), 'Multi-Criteria Analysis and the Resolution of Sustainable Development Dilemmas: A Stakeholder Management Approach.' *European Journal of Operational Research* 224 (1): 122–131.
- Cherchye, L., W. Moesen, N. Rogge, and T. Van Puyenbroeck (2007), 'An Introduction to 'benefit of the Doubt' Composite Indicators.' *Social Indicators Research* 82: 111–145.
- Covas, M. T., C.A. Silva, and L. C. Dias (2013), 'Multicriteria Decision Analysis for Sustainable Data Centers Location.' *International Transactions in Operational Research* 20 (3) (May 11): 269–299.
- Dias, L.C., and J.N. Clímaco (2000), 'Additive Aggregation with Variable Interdependent Parameters: The VIP Analysis Software.' *Journal of the Operational Research Society* 51 (9): 1070–1082.
- Dias, L.C., and V. Mousseau (2003), 'IRIS: A DSS for Multiple Criteria Sorting Problems.' *Journal of Multi-Criteria Decision Analysis* 12 (4-5) (July): 285–298.
- Dias, Luis C., and A. R. Domingues (2014), 'On Multi-Criteria Sustainability Assessment: Spider-Gram Surface and Dependence Biases.' *Applied Energy* 113 (January): 159–163.
- Dodgson, J., M. Spackman, A. Pearman, and L. Philips (2000), 'Multi-Criteria Analysis: A Manual'. London: Department for the Environment, Transport and the Regions.
- DRAEDM (2007), 'Entre Douro E Minho's Regional Development Plan for Dairy Farming' (in Portuguese). Direcção Regional de Agricultura de Entre Douro e Minho.

- Ehrgott, Matthias (2005), *Multicriteria Optimization*. 2nd ed. Berlin and Heidelberg: Springer.
- Ehrgott, Matthias, and Theodor J. Stewart (ed.) (2010), *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems*. Springer.
- Figueira, J., V. Mousseau, and B. Roy (2005), 'ELECTRE Methods.' In *Multiple Criteria Decision Analysis: State of the Art Surveys*, edited by José Figueira, Salvatore Greco, and Matthias Ehrgott, 133–153. Springer.
- Forman, E. H., and S. I. Gass (2001), 'The Analytic Hierarchy Process — An Exposition.' *Operations Research* 49 (4): 469–486.
- Goodwin, Paul, and George Wright (2014). *Decision Analysis for Management Judgement*. 5th ed. Wiley.
- Grabisch, M., and C. Labreuche (2010), 'A Decade of Application of the Choquet and Sugeno Integrals in Multi-Criteria Decision Aid.' *Annals of Operations Research* 6(1), 1–44.
- Huang, H., L. Zhang, Z. Liu, and J. W. Sutherland (2010), 'Multi-Criteria Decision Making and Uncertainty Analysis for Materials Selection in Environmentally Conscious Design.' *The International Journal of Advanced Manufacturing Technology* 52 (5-8) (June 8): 421–432.
- Huang, I. B., J. Keisler, and I. Linkov (2011), 'Multi-Criteria Decision Analysis in Environmental Sciences: Ten Years of Applications and Trends.' *The Science of the Total Environment* 409 (19) (September): 3578–94.
- Ishizaka, A., and A. Labib (2011), 'Review of the Main Developments in the Analytic Hierarchy Process.' *Expert Systems with Applications* 38 (11) (May): 14336–14345.
- Ishizaka, Alessio, and Philippe Nemery (2013), *Multi-Criteria Decision Analysis: Methods and Software*. Chichester: Wiley.
- Jacquet-Lagrèze, E., and Y. Siskos (2001), 'Preference Disaggregation: 20 Years of MCDA Experience.' *European Journal of Operational Research* 130: 233–245.
- Keeney, R. L. (2004), 'Framing Public Policy Decisions', *International Journal of Technology, Policy and Management*, 4 (2), 95–115.
- Keeney, R. L. (2006), 'Using Preferences for Multi-Attributed Alternatives' *Journal of Multi-Criteria Decision Analysis* 14: 169–174.
- Keeney, R. L., and A. Sicherman (1983), 'Illustrative Comparison of One Utility's Coal and Nuclear Choices.' *Operations Research* 31 (1): 50–83.
- Keeney, Ralph L., and Howard Raiffa (1993). *Decisions with Multiple Objectives—preferences and Value Tradeoffs*. Cambridge & New York: Cambridge University Press.
- Lahdelma, R., and P. Salminen (2001), 'SMAA-2 : Stochastic Multicriteria Acceptability Analysis for Group Decision Making.' *Operations Research* 49 (3): 444–454.
- Lahdelma, R., P. Salminen, and J. Hokkanen (2002), 'Locating a Waste Treatment Facility by Using Stochastic Multicriteria Acceptability Analysis with Ordinal Criteria.' *European Journal of Operational Research* 142 (2) (October): 345–356.

- Linkov, I., and E. Moberg (2012), *Multi-Criteria Decision Analysis: Environmental Applications and Case Studies*. CRC Press.
- Marchant, Th. (1996), 'Valued Relations Aggregation with the Borda Method.' *Journal of Multi-Criteria Decision Analysis* 5 (2): 127–132.
- Martel, Jean-Marc, and B Matarazzo (2005), 'Other Outranking Approaches.' In *Multiple Criteria Decision Analysis: State of the Art Surveys*, edited by José Figueira, Salvatore Greco, and Matthias Ehrgott, 198–219. Springer.
- Marttunen, M., and M. Suomalainen (2005), 'Participatory and Multiobjective Development of WaterCourse Regulation - Creation of Regulation Alternatives from Stakeholders ' Preferences' *Journal of Multi-Criteria Decision Analysis* 49: 29–49.
- Munda, G. (2004), 'Social Multi-Criteria Evaluation: Methodological Foundations and Operational Consequences.' *European Journal of Operational Research* 158 (3): 662–677.
- Munda, Giuseppe (1995), *Multicriteria Evaluation in a Fuzzy Environment*. Heidelberg: Physica Verlag.
- Munda, Giuseppe (2005), 'Multiple Criteria Decision Analysis and Sustainable Development.' In *Multiple Criteria Decision Analysis: State of the Art Surveys*, edited by José Figueira, Salvatore Greco, and Matthias Ehrgott. Springer.
- Mustajoki, J., R. P. Hämäläinen, and M. Marttunen (2004), 'Participatory Multicriteria Decision Analysis with Web-HIPRE: A Case of Lake Regulation Policy.' *Environmental Modelling & Software* 19 (6): 537–547.
- Myllyviita, T., A. Holma, R. Antikainen, K. Lähtinen, and P. Leskinen (2012), 'Assessing Environmental Impacts of Biomass Production Chains – Application of Life Cycle Assessment (LCA) and Multi-Criteria Decision Analysis (MCDA).' *Journal of Cleaner Production* 29-30 (July): 238–245.
- Neves, L.P., L.C. Dias, C.H. Antunes, and A.G. Martins (2009), 'Structuring an MCDA Model Using SSM: A Case Study in Energy Efficiency.' *European Journal of Operational Research* 199 (3) (December): 834–845.
- Oberschmidt, J., J. Geldermann, J. Ludwig, and M. Schmehl (2010), 'Modified PROMETHEE Approach for Assessing Energy Technologies.' *International Journal of Energy Sector Management* 4 (2): 183–212.
- Pohekar, S.D., and M. Ramachandran (2004), 'Application of Multi-Criteria Decision Making to Sustainable Energy planning—A Review.' *Renewable and Sustainable Energy Reviews* 8 (4): 365–381.
- Ramanathan, R. (2001), 'A Note on the Use of the Analytic Hierarchy Process for Environmental Impact Assessment.' *Journal of Environmental Management* 63 (1): 27–35.
- Rosenhead, J. (1996), 'What's the Problem? An Introduction to Problem Structuring Methods.' *Interfaces* 26(6): 117–131.

- Roth, S., S. Hirschberg, C. Bauer, P. Burgherr, R. Dones, T. Heck, and W. Schenler (2009), 'Sustainability of Electricity Supply Technology Portfolio.' *Annals of Nuclear Energy* 36 (3): 409–416.
- Roy, B. (1991), 'The Outranking Approach and the Foundations of Electre Methods.' *Theory and Decision* (31): 49–73.
- Roy, Bernard (1996), *Multicriteria Methodology for Decision Aiding*. Dordrecht: Kluwer Academic.
- Saaty, T. L. (2008), 'Decision Making with the Analytic Hierarchy Process.' *International Journal of Services Sciences* 1(1), 83–98.
- Sarabando, P, and L.C. Dias (2009). 'Multiattribute Choice With Ordinal Information: A Comparison of Different Decision Rules.' *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 39 (3) (May): 545–554.
- Stewart, T. J. (1996), 'Robustness of Additive Value Function Methods in MCDM.' *Journal of Multi-Criteria Decision Analysis* 5 (4): 301–309.
- Van Calker, K.J., P.B.M. Berentsen, C. Romero, G.W.J. Giesen, and R.B.M. Huirne (2006), 'Development and Application of a Multi-Attribute Sustainability Function for Dutch Dairy Farming Systems.' *Ecological Economics* 57 (4) (June): 640–658.
- Wang, J.-J., Y.-Y. Jing, C.-F. Zhang, and J.-H. Zhao (2009), 'Review on Multi-Criteria Decision Analysis Aid in Sustainable Energy' *Renewable and Sustainable Energy Reviews* 13: 2263–2278.
- Yager, R.R. (1988), 'On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking.' *IEEE Transactions on Systems, Man, and Cybernetics* 18(1), 183-190.
- Yoon, Kwangsun, and Ching-Lai Hwang (1995), *Multiple Attribute Decision Making: An Introduction*. Thousand Oaks (CA): Sage Publications.