A real world MCDA application: evaluating software

E. Paschetta(*), A. Tsoukiàs(+)
(*) CSELT spa, via G. Reiss Romoli 274, 10148 Torino, Italy
e-mail: Elio.Paschetta@cselt.it
(+) LAMSADE - CNRS, Université Paris Dauphine
75775 Paris Cedex 16, France
e-mail: tsoukias@lamsade.dauphine.fr

Abstract

A real world MCDA application on software evaluation is presented in the paper. The decision process concerned a big Italian company faced with the management of a call for tenders for a very important software acquisition. The decision aiding process is extensively presented and discussed, mainly as far as its products are concerned, that is:
- the problem situation;
- the problem formulation;
- the evaluation model;
- the final recommendation.
The results of the experience are discussed using the comments of the client of the study.

Keywords: software evaluation, decision aiding process, evaluation model, ordinal measurement.

Introduction

Although the use of the multicriteria methodology is now widely acknowledged in many sectors of human life, very few real world applications are reported in the literature (for noticeable exceptions see, Belton et al., 1997; Tabucanon and Chankong, 1989; Bana e Costa et al., 1997; Vincke, 1992b; Roy and Bouyssou, 1993). In all cases it is rare to find extended presentations of such real world cases where the whole decision aiding process is ex-post analysed, discussed and where lessons are obtained. Despite the fact that the multicriteria decision aid approach is deeply rooted in empirical research, very little operational validation is reported in the literature as already noticed in Bouyssou et al. (1993) and French (1998).

The paper tries to contribute in filling such a gap reporting a real world decision aiding process which took place in a large Italian firm late 1996 and early 1997 concerning the evaluation of offers following a call for tenders for a very important software acquisition problem. In the paper we will try to extensively present the decision process for which the decision aiding was requested, the actors involved, the decision aiding process, including the problem structuring and formulation, the evaluation model created and the multicriteria method adopted.

More precisely the paper is organized as follows. Section 1 introduces and defines some preliminary concepts which will be used in the rest of the paper such as decision process, actors, decision aiding process, problem formulation, evaluation model, etc. Section 2 presents the decision process for which the decision aid was requested, the involved actors and their concerns (stakes), the resources engaged and the timing. The role of the authors is also presented in this section. Section 3 describes the decision aiding process, mainly through the different "products" of such a process which are specifically analysed (the problem formulation, the evaluation model and the final recommendation). Section 4 contains a discussion about the experience conducted in which an ex-post analysis is performed concerning the decision process, the information available, the evaluation model and the methods used. The clients' and the analysts' comments on the experience are also included in this section. The conclusion emphasizes the open questions that the experience highlighted, indicating future necessary theoretical achievements and operational validations. All technical details are included in Appendix A (the ELECTRE-TRI type procedure used), while the complete list of the evaluation attributes is provided in Appendix B.

1 Preliminaries

All representations always follow a descriptive model used by the observers (in this case the authors). In the following we will try to introduce some key concepts and issues which may help the reader to better understand our presentation.

A basic concept adopted in our paper is the one of decision process (see Mintzberg et al., 1976; Jacquet-Lagrèze et al., 1978; Heurgon, 1982; Humphreys et al., 1983; Masser, 1983; Nutt, 1984; Ostanello, 1990 and 1997). We will not confine ourselves to the classic concept of decision process given by Simon (1957). We may emphasize that our concept of decision process does not refer to an individual behaviour as observed facing a decision situation (as studied by descriptive decision aiding approaches, see Svenson 1996). For us a decision process is the set of all activities performed around an "object" which evokes the concerns of one or more individuals, groups, formal or informal organizations. We call actors all participants involved in a decision process. We call objects (of the decision process) all the concerns of all the actors involved in the decision process. Normally the actors allocate to their concerns (objects) resources including knowledge, ideas, time, etc. (for more details on this type of descriptive model see Ostanello and Tsoukiàs, 1993).

A first assumption in our approach (typical of the multicriteria decision aid (MCDA) approach) is that the decision aiding activity is an interaction of at least two distinct actors, the "client" and the "analyst". Normally the client is involved in a decision process. We call such interaction a "decision aiding process" and we admit that also other actors can be involved in this specific task. A decision aiding process has as object a concern of the client on which (s)he considers the allocated resources insufficient and estimates necessary the help of another actor (the analyst). The way by which the decision aiding process is structured depends on how the analyst perceives the client concerns, the problem situation in which the client finds (her)himself and on how the client perceives the analysts methodological knowledge (for more details on the concepts of client and analyst, see Checkland, 1981; Moscarola, 1984; Rosenhead, 1989; Norese, 1988).

There is no unique approach under which it is possible to describe the activities of the analyst in the decision aiding process since there is no unique decision aiding approach. The one adopted in this case is based on the MCDA methodology (see Roy, 1996; Vincke, 1992b) and its use in the software evaluation (see Morisio and Tsoukiàs, 1997). Basically such an approach considers three consecutive products of the decision aiding process:
- a representation of the problem situation;

- one or more problem formulations coherent with the problem situation representation;
- an evaluation model consistent with the problem formulation adopted.

More formally we have:

1. A problem situation model in time t (PS_t) is a triplet $\langle \mathcal{A}_t, \mathcal{O}_t, RS_t \rangle$ where:
   - $\mathcal{A}_t$ are the actors involved in the decision process in which the client find (her)himself and for which (s)he ask the help of the analyst;
   - $\mathcal{O}_t$ are the objects (concerns, stakes, etc.) of the actors involved in the process;
   - RS_t are the resources allocated by the actors to the objects in the problem situation.
   In other words, a problem situation model allows the analyst and the client to have a clear common view of what are the stakes at the moment, to identify the reasons for which the present situation is considered as problem for the client and to understand whether such a "problem" can be formulated in a more formal (less ambiguous) way.

2. A problem formulation Γ_{PS} (with respect to a problem situation) is a triplet $\langle A, V, \Pi \rangle$ where:
   - A is the set of actions which the client may adopt, execute, pursue in order to provide an answer to the problem situation in which (s)he find (her)himself;
   - V are the points of view under which the different actions may be considered by the client and the analyst;
   - Π is a problem statement which declares the purpose and type of evaluation of the alternatives. The purpose can be operational (a choice, a reject, a ranking), descriptive or conceptual. The type can be absolute or relative.
   A problem formulation therefore translates the concerns expressed in the problem situation model in a "formal" problem on which it may be possible to apply some techniques such as statistics, measurement, operational research, simulation, etc.

3. An evaluation model $\mathcal{M}_\Gamma$ (with respect to a problem formulation) is a n-uple $\langle A^*, D, E, M, G, \mathcal{U}, \mathcal{R} \rangle$ where:
   - A^* are the alternative actions which will be considered by the model;
   - D are the different dimensions under which the evaluation will be performed; it may be an hierarchy or a flat set;
   - E, M are the available (if any) measurements and their relative scales

associated with some of the elements of D;
- G is the set of criteria constructed on the basis of D;
- $\mathcal{U}$ is the uncertainty imported or created by the model and the representations adopted to handle it;
- $\mathcal{R}$ is the set of aggregation techniques that will be used in order to put together the evaluations of the elements of A^*, expressed on the different dimensions in D, in order to obtain a final result for the client consistent with the problem statement.
The evaluation model therefore presents the precise way under which the alternatives that the client considers as possible are going to be evaluated and includes any eventual prescription the analyst will provide to the client.

With respect to the evaluation model, we want to make some remarks.

1. The reader may notice that the evaluation model is the third product constructed in the decision aiding process. It is therefore the result of a deep abstraction with respect to the reality and therefore different (more or less) arbitrary "cuts" of the reality have been done. A clear construction of the three products and of the reasons for which the different cuts have been done may facilitate the understanding and the agreement of the client to the results of the decision aiding process and a better use of such results in the decision process.

2. Normally an evaluation model contains or create uncertainty. It can be imported in the model due to imprecision, to errors, to missing or inconsistent information, or it can be created by the model itself due to arbitrary reductions of the information or to implicit ambiguity of the model structure. The way by which uncertainty is represented is not neutral (and not objective) and not always straightforward since there exist different formalisms and approaches in modeling uncertainty. This is why uncertainty is considered as a part of the evaluation model, the elements of which have to be justified.

3. The evaluation model provides the client with a "condensed" representation of the alternatives, obtained aggregating data, measures and preferences about them. Again the choice of the aggregation procedures is not straightforward (since there exist different aggregation procedures) and not neutral (since different procedures result in different results). This is the reason for which the set of aggregation procedures is a part of the model, being choices which have to be justified.

4. Actually a decision aiding process generates a fourth product as a direct consequence of the evaluation model. Such a product is the "final recommendation". In fact it contains the suggestions and advises of the analyst to the client. Conventionally we may consider that such a product is more or less implicitly contained in the evaluation model (as soon as the aggregation procedure is applied to the available information the final recommendation might appear). Unfortunately this is not always the case. There are some further considerations and elaborations to perform in order to provide such a final product. Such further elaborations may include (without being limited to):
   - elaboration of a final prescription to the client;
   - sensitivity analysis of the result;
   - robustness analysis of the result.

From our preliminary formulation of the outline of our model of the decision aiding process products, we hope to make clear that:

- the decision aiding is the result of the interactions among at least two actors (the client and the analyst) which have to achieve a consensual representation of the reality through three models: the problem situation model, the problem formulation and the evaluation model;

- for the same reason such three models require a double validation:
  - from the client point of view, it has to satisfy his(her) expectations and necessities;
  - from the analyst point of view, it has to satisfy some formal properties of meaningfulness, correctness and formal coherence.

Nevertheless the construction of the three models has many degrees of freedom and is not unique. Moreover the double validation of the models may also help the legitimation of the decision aiding process towards the reference decision process.

2 The Decision Process

In the early 1996, a very large Italian company operating a network based service decided, as part of a strategic development policy, to get equipped with a Geographical Information System (GIS) on which all information concerning the structure of the network and the services provided all over the country was to be transferred. However, since this was quite a new technology, the Information Systems Department (ISD) of the company asked the affiliated research and development agency (RDA), and more specifically the department concerned with this type of information technology (GISD), to perform a pilot study of the market in order to orient the company towards an acquisition.

The GISD of the RDA noticed that:
- the market offered a very large variety of possibilities of software which could be used as a GIS for the company's purposes;
- the company required a very peculiar version of GIS that did not exist ready made in the market, but had to be created customizing and combining different modules of existing software besides ad-hoc written software for the companies purpose;
- the question of the ISD was very general, but also very committing because it included an evaluation for an acquisition and not a simple description of the different products;
- the GISD felt able to describe and evaluate different GIS products on a set of attributes (possibly some hundreds), but was not able to provide a synthetic evaluation, the purpose of which was even obscure (the use of a weighted sum was quite immediately left aside because perceived as "meaningless").

At this point of the process, the GISD came to know that within the RDA was operating a unit concerned with the use of the MCDA methodology in software evaluation (MCDA/SE) and presented this problem as a case study opening a specific commitment. The first author was thus involved into the process. On its turn the MCDA/SE unit responsible decided to activate his links with an academic institution in order to get more insight and advice on the problem which soon appeared to overcome the knowledge level of the unit at that time. The second author was thus involved into the process. At this point we can make the following remarks.

- The reference decision process for which the decision aiding was provided concerned the "acquisition of a GIS for X (the company)". The actors involved at this level are the IS manager of the company, the acquisition (AQ) manager of the company, the RDA, different suppliers of GIS software, some external consultants of the company concerned with software engineering.

- A first decision aiding process was established where the client is the ISD manager and the analyst is the GISD department of the RDA.

- A second decision aiding process was established where the client is the GIS department of the RDA and the analyst is the MCDA/SE unit. A third actor involved into this process is the "supervisor" of the analyst in the sense of someone supporting the analyst in different modeling tasks, providing him with expert methodological knowledge and framing his activity.

We will focus our attention on this second decision aiding process where four actors are involved: the ISD manager (background client), the GISD (or team of analysts) as domain experts and client (keep in mind their peculiar position of clients and analysts at the same time), the MCDA/SE unit as analyst and the supervisor.

A first advice of the analyst to the GISD was to negotiate a more specific commitment with their client such that their task could be more precise and better defined. After such a negotiation the frame of the GISD activity has been the "technical assistance to the IS manager in a bid, concerning the acquisition of a GIS for the company" and the specific task was to provide a "technical evaluation" of the offers that were expected to be submitted. For this purpose the GISD traced a decision aiding process outline where the main activities to be performed were emphasized besides the timing and submitted it to their client (see Figure 1). A call for tenders was therefore established and used in order to obtain some specific offers. At this point it is important to notice the following.

1. The call for tenders concerned the acquisition of some hundreds of software licenses, plus (possibly) the hardware platforms on which such software was expected to run, the whole budget being several millions of euros. From a financial point of view it represented a large stake for the company and a high level of responsibility for the decision makers.

2. From a procedural point of view the administration of a bid of this type is delegated to a committee which in this case included the IS manager, the AQ manager, a delegate of the CEO and a lawyer of the legal staff. Under such a perspective the task of the GISD (and of the decision aiding process) was to provide the IS manager with a "global" technical evaluation of the offers which could be used in the negotiations with the AQ manager (inside the committee) and the suppliers (outside the committee).

3. As already noticed before, the bid concerned not ready made software, but an assembly of existing modules of GIS software which was expected to be used in order to create ad-hoc software for the specific company's necessities. This generated two difficulties:

- evaluating a priori the software behaviour and its performance without being able to test it on specific company's cases;
- the timing of the evaluation (including the testing of the offers) could

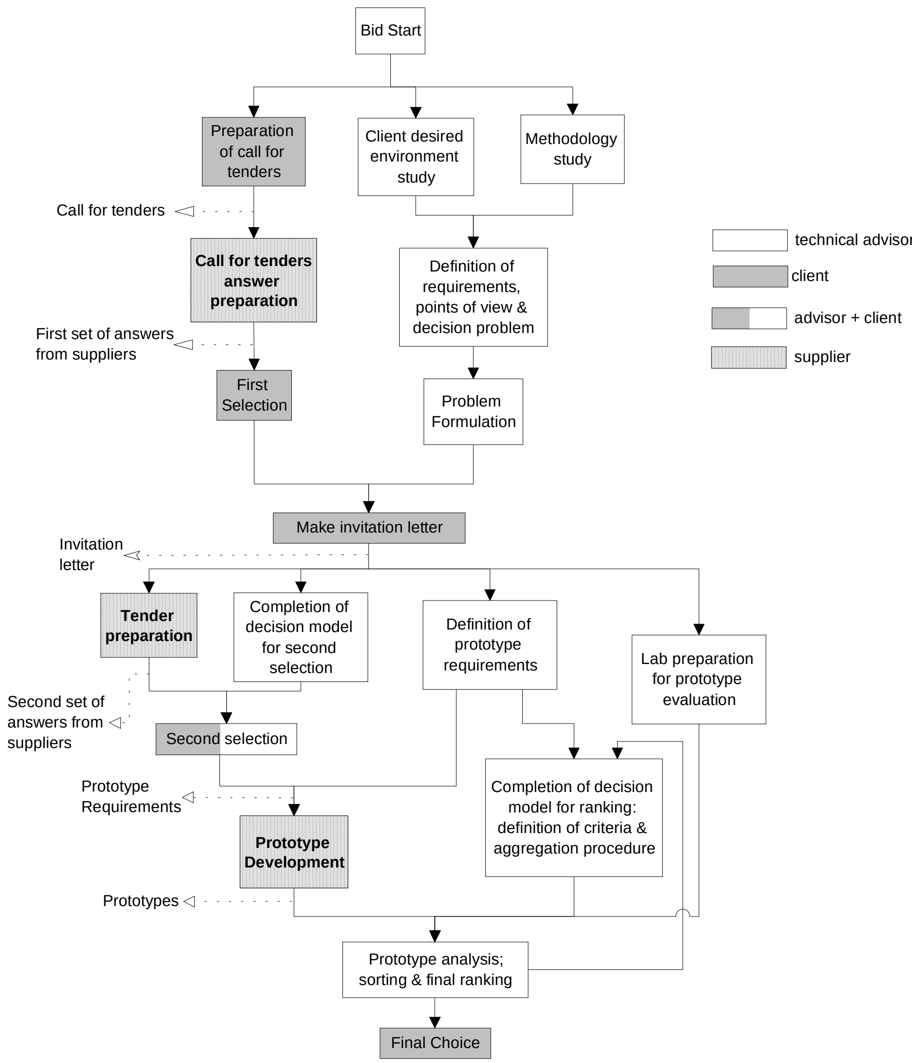

Figure 1: The bid process

Once the call for tenders has been prepared (including the software requirements sections, the tenders requirements section, the timing and evaluation procedure) a set of offers has been presented to the company and the technical evaluation activity has been settled. It is interesting to notice that the GISD staff in charge of this evaluation has been "supported" by some external consultants, software engineering experts of the company's sector and practically acting as delegates of the IS manager in the group. It is this extended group that signed the final recommendation to the IS manager.

A second step in the decision aiding process has been the generation of a problem formulation and of an evaluation model. Although we formally consider the two as two distinct products of the process, in reality and in this case specifically, they have been generated contemporaneously. We will discuss in detail the problem formulation and the evaluation model in the next section, but we can anticipate that the final formulation consisted of an absolute evaluation of the offers under a set of points of view which could be divided in two parts: the "quality evaluation" and the "performance evaluation". Although the set of alternatives was relatively small (only six alternatives have been considered), the set of attributes was extremely complex (as often happens in software evaluation). Actually the basic evaluation dimensions were seven, expanded in an hierarchy with 134 leaves resulting in 183 evaluation nodes.

A third and final step of the decision aiding process has been the elaboration of the final recommendation after all the necessary information for the evaluation had been obtained and the evaluation performed. We will discuss in detail such a construction in the next section, but we can anticipate that such an elaboration put in evidence some questions (substantial and methodological) which have not been considered before.

3 Decision Aiding

In the following we present the three products of the decision aiding process: the problem formulation, the evaluation model and the final recommendation. We have to remind that the problem formulation and a first outline of the evaluation model have been established while the call for tenders was under elaboration for two reasons:
- for legal reasons an outline of the evaluation model has to be included into the call for tenders;
- the evaluation model contains implicitly the software requirements of the offers which on its turn defines which is the information the tenders have to provide. For instance the call for tenders specified that a prototype was

asked in order to test some performances. The tenders therefore knew they had to produce a prototype in a certain time. The choice to introduce some tests is done during the definition of the evaluation model.

3.1 Problem Formulation

The set A was considered as the set of offers to be submitted after the call for tenders. A first idea to evaluate the tenders, besides the offers, was eliminated due to the particular technology where consolidated producers do not exist. The set of points of view was defined using the technical knowledge of the team of analysts and can be viewed in two basic sets. One concerning "quality" including specific technical features required for the software plus some ISO/IEC 9126 (1991) based dimensions and the second concerning the performance of the offered software to be tested on prototypes. No cost estimations were required by the client and so they were not considered in this set.

The only point that engaged a discussion in the team of analysts about the problem formulation was the problem statement Π. After some discussion the problem statement adopted was the one of an "absolute" evaluation of the offers both on a disaggregated level and on a global one. Actually the team of analysts interpreted the client's demand as a question of whether the offers could be considered intrinsically "good", "bad", etc. and not to compare bids among them. There were two reasons for this choice.

1. A simple ranking of the offers could conceal the fact that all of them could be of a very poor quality or satisfy to a very low level the software requirements. In other words it could happen that the best bid could be "bad" and this was incompatible with the importance and cost of the acquisition.

2. The team of analysts felt uncomfortable with the idea to compare merits (or de-merits) of an offer with merits (or de-merits) of another offer. A first informal discussion about the problem of compensation convinced them to overcome this question by comparing the offers to profiles for which they had sufficient knowledge.

If we interpret the concept of measurement in a wide sense (comparing the offers to pre-established profiles can be viewed as a measurement procedure) the result the team of analysts was looking for appeared to be the conclusion of repeated aggregation of measures. Using the terminology introduced by Roy (1996), the problem statement appeared to be an hierarchically organized sorting of the offers, the sorting being repeated to all levels of the hierarchy.

As it will become more clear in the following, such a problem formulation concerned essentially the "quality" points of view. When the client has been faced with the definition of the final recommendation he decided to elaborate a relative ranking to compare with the absolute evaluation. In fact, while the client desired to know how the offers behaved with respect to the quality requirements, he wanted to have some operational indications on hand (an acquisition recommendation) to offer to the ISD manager.

3.2 The Evaluation Model

The different components of the evaluation model have been specified in an iterative way. In the following we present their definition as it occurred in the decision aiding process.

The set of alternatives A^* has been identified as the set of offers legally accepted by the company in reply to the call for tenders. No preliminary screening of the offers was expected to be done. Although each offer was composed by different modules and software components, they have been considered as wholes.

The set of evaluation dimensions D was a complex hierarchy with seven root nodes, 134 leaves and 183 nodes in total (the complete list is available in Appendix B). The key idea was that each node of the hierarchy was an evaluation model itself for which the evaluation dimensions to aggregate and the aggregation procedure had to be defined. The whole hierarchy of nodes was subject to extensive discussion before arriving at a final version. Basically two elements have been considered in such a discussion:
- the separability of each sub-dimension with respect to the parent node, in the sense that each sub-node should be able to discriminate alone the offers with respect to the evaluation considered on the parent level;
- the presence of redundant nodes at the same level of evaluation in the sense of nodes carrying practically the same information about the features of the offers.

Before carrying on the definition of the model associated with each node, the problem of the aggregation procedure has been faced since it could influence the construction of such models. In our case the presence of ordinal information in almost all leaves and the problem statement which required a "repeated sorting" of the offers, oriented the team of analysts to choose an aggregation procedure based on the ELECTRE TRI method (see Yu Wei, 1992; see also appendix A for a detailed presentation of the procedure). At this point the team was ready to define for all nodes their specific evaluation models. In particular we had the following cases.

1. For all leave nodes an ordinal scale has been established. The avai-

lable technical knowledge consisted of different possible "states" in which an offer could find itself. For instance consider the leave nodes 1.1.1 (type of presentation on the user interface in the land-base management), 1.1.2 (graphic engine of the user interface in the land-base management), 1.1.3 (customization of the user interface in the land-base management). The possible states on these characteristics were:
1.1.1: standard graphics (SG), non standard graphics (NSG);
1.1.2: station M (M; graphic engine already adopted in other software used in the company), other acceptable graphic engine (OA), other non acceptable graphic engine (ON);
1.1.3: availability of a graphic tool (T), availability of an advanced graphic language (E), availability of a standard programming language (S), no customization available (N). In this case different possible combinations were possible (for instance T,E,S means availability of a graphic tool, an advanced graphic language and a standard programming language).

The three ordinal scales associated with the three nodes have been ($>$ representing the scale order):
1.1.1: SG $>$ NSG;
1.1.2: M $>$ OA $>$ ON;
1.1.3: T,E,S $>$ T,E $>$ T,S $>$ T $>$ E,S $>$ E $>$ S $>$ N.

2. For all parent nodes a brief descriptive text of what the node was expected to evaluate was provided. All parent nodes were equipped with the same number of classes: unacceptable (U), acceptable (A), good (G), very good (VG), excellent (E). Then two possibilities for defining the relationship between the values on the sub-nodes and the values on the parent nodes were established.

   2.1 When possible, an exhaustive combination of the values of the sub-nodes was provided. For instance consider node 1.1 (user interface of the land-base management) which has as sub-nodes the three evaluation models introduced previously. In this case we have the following evaluation model:
   - E: (T,E,S),M,SG or (T,E),M,SG or (T,S),M,SG;
   - VG: T,M,SG or (T,E,S),OA,SG or (T,E),OA,SG or (T,S),OA,SG;
   - G: T,OA,SG or (E,S),M,SG or E,M,SG;
   - A: all remaining cases except the unacceptable;
   - U: all cases where 1.1.1 is NSG, or 1.1.2 is ON, or 1.1.3 is N.

   2.2 When an exhaustive combination of the values was impossible an ELECTRE TRI procedure was used. For this purpose the

following information was requested:
- relative importance of the different sub-nodes;
- concordance threshold for the establishment of the outranking relation among the offers and the profiles;
- veto condition on the sub-node such that the value on the parent node could be limited (possibly unacceptable).

The relative importance of the sub-nodes and the concordance threshold have been established using a reasoning about coalitions. In other terms the team of analysts established which were the characteristics of the sub-nodes for which an offer could be considered as a very good offer (therefore should outrank the very good profile) and consequently compared the values of the relative importance parameters and of the concordance threshold. The veto condition has been established as the presence of the value "unacceptable" to a sub-node. The presence of a veto had the effect to produce an "unacceptable" value also at the level of the parent node. In other words the team of analysts considered any "unacceptable" value as a severe technical limit of the offer. The reader may notice that this is a very strong notion of veto among the ones used in the outranking based sorting procedures, but it was the one with which the team of analysts felt comfortable at the moment of the construction of the evaluation model. The team of analysts also used to establish very high concordance thresholds (never less than 80%, very often around 90%) which results in a very severe evaluation. Such a choice reflected the conviction of at least a part of the team of analysts that very strong reasons were necessary in order to qualify an offer as very good. Since the whole model was calibrated beginning from the very good value, this conviction had wider effects than the team of analysts could consider.

As an example for the definition of the importance parameters, we can take the node 1 (land-base management) which has eight sub-nodes:
1.1: User interface;
1.2: Functionality;
1.3: Development environment;
1.4: Administration tools;
1.5: Work flow connection;
1.6: Interoperability;
1.7: Integration between land-base products and the Spatial Data manager;
1.8: Integration among land-base products;

The relative importance parameters were established as follows: (1.1):4, (1.2):8, (1.3):5, (1.4):4, (1.5):1, (1.6):4, (1.7):8, (1.8):2 and the concordance threshold has been fixed as 29/36 (around 0.8). Such choices reflect the conviction that no coalition was acceptable excluding the nodes 1.2 or 1.7, and that the smallest acceptable coalition should necessarily include the nodes 1.2, 1.7, 1.3 and any two among the nodes 1.1, 1.4 and 1.6.

3. As already mentioned, the set of dimensions was built around two basic points of view which were the "quality" and the "performances". The first generated six evaluation dimensions which will be called hereafter the "quality attributes" or "quality criteria" or "quality part of the hierarchy" corresponding to six (among seven) of the root nodes of the model. The seventh root node (node 7, sub-nodes 7.1, 7.2, 7.3, 7.4) concerned the evaluation of the performances of the prototypes submitted to some tests by the team of analysts. Such performances were basically measured in the necessary time to execute a set of specific tasks under certain conditions and with some external parameters fixed.

   For instance consider the node 7.3 (performance under load). The dimension is expected to evaluate the performance of the prototype as the quantity of data that have to be elaborated increases. The value $v(x)$ (x being an offer) combined an observed measure $W_x(t)$ and an interpolated one $T_x(t)$ (t representing the data load, the interpolation being not necessarily linear). The combination is obtained in this case through the following formula:

$$v(x) = \int W_x(t) T_x(t) dt$$

   However, in this case there are no external profiles to which compare the performances because the prototypes are created ad-hoc, the technology was quite new and there were no standards of what a "very good" performance could be. An ordinal scale has been created in this case considering the best performances as "first", all performances presenting a difference of more than 5% and less than 20% "second", all performances presenting a difference of more than 20% and less than 25% "third", all performances presenting a difference of more than 25% and less than 50% "fourth" and all performances presenting a difference of more than 50% "fifth". The same model has been applied to all sub-nodes of the node 7. A sorting procedure could then be established in order to obtain the final evaluation.

This process has been repeated for all the intermediate nodes up to the seven root nodes representing the seven basic evaluation dimensions. Such a process took four to five months before all the nodes were equipped with their evaluation model and generated several discussions in the team of analysts mainly of technical nature (concerning the specific contents of the values on each node). The most discussed concept of the model was the concordance threshold and the veto condition since part of the team considered that the required levels were extremely severe. However, since such an approach corresponded to a cautious attitude, it finished to prevail in the team and finally has been accepted. The length of the process is justified not only by the quantity of nodes to define, but also because the team of analysts for each node was obliged to define a new measurement scale and a precise measurement aggregation procedure. Although this process can be often qualified as "subjective measurement", it was the only way to obtain some meaningful values for the offers.

The set of criteria to be used, in case a preference aggregation was requested comparing the alternatives among them, was defined as the seven root nodes equipped with a simple preference model: the weak order induced by the ordinal scale associated with each of these nodes.

No exogenous uncertainty was considered in the evaluation model. The information provided by the tenders concerning their offers was considered as reliable and the use of ordinal scales avoided the problems of imprecision or of measurement errors. This reasoning, however, is less true as far as the node 7 and its sub-nodes are concerned, but the team of analysts felt sufficiently confident with the tests and did not analyze further the problem. On the other hand some endogenous uncertainty appeared as soon as the model was put into practice (the offers being available). We shall discuss more in details this problem in the next section (concerning the elaboration of the final recommendation), but we can anticipate that the problem was created by the "double" evaluation provided by the ELECTRE TRI aggregation type adopted, consisting of an "optimistic" and a "pessimistic" evaluation which may not necessarily coincide.

The evaluation model has been coded in a formal document which has been submitted (and explained) to the final client receiving his consensus. It is worth notice that the final client as such was not able to participate to the elaboration of the model (technical details, parameters establishment, etc.). Part of the team of analysts (some of the external consultants) were acting as his delegates. The establishment of the evaluation model and its acceptance by the client opened the way for its application on the set of offers received and for the elaboration of the final recommendation.

3.3 The final recommendation

The evaluation of the six offers, which were submitted after the call for tenders, has been elaborated in two main steps. The first one consisting of evaluating the six "quality attributes", and the second one consisting of testing the prototypes provided by the tenders.

A specific problem which raised in the first step was the generation of uncertainty due to the aggregation procedure. The ELECTRE TRI type procedure adopted produces an interval evaluation consisting of a lower value (the pessimistic evaluation) and an upper value (the optimistic evaluation). When an alternative has a profile on the sub-nodes which is very different from the profiles of the classes on the parent node then, due to the incomparabilities that occur comparing the alternative to the profiles, it may happen that the two values do not coincide (see more details in Appendix A). If the user of the model is not able to choose one of the two evaluations in an hierarchical aggregation, this can be a problem since at the next aggregation, the sub-nodes may have evaluations expressed on an interval. For this purpose the following procedure has been adopted. Two distinct aggregations were done, one where the lower values were used and the other one where the upper values were used. Each of these may produce on its turn a lower value and an upper value. At the next aggregation step the lowest of the two lower values and the highest of the two upper values are used. This is a cautious attitude and has the drawback to widen the intervals as the aggregation goes up the hierarchy. However, in the specific case this effect did not occur and the final result on the six dimensions is presented in Table 1 (in the following we will represent with Ci the criteria and with Oi the alternatives).

	O1	O2	O3	O4	O5	O6
C1	A-A	G-G	A-VG	A-G	G-VG	A-A
C2	A-A	G-VG	A-VG	A-VG	G-G	A-G
C3	A-A	G-G	A-VG	G-G	A-A	A-A
C4	A-G	G-VG	A-VG	G-VG	A-VG	A-G
C5	U-U	G-VG	G-G	A-G	G-VG	U-U
C6	A-A	VG-VG	E-E	VG-VG	G-G	VG-VG

Table 1: The values of the alternatives on the six quality criteria.

Another modification introduced in the aggregation procedure concerned the use of the veto concept. As already mentioned, in the evaluation model a strong veto concept has been used such that the presence of an "unacceptable" on any node (among the ones endowed with such veto power) could result in a global "unacceptable" value. However, during the evaluation of the offers weaker concepts of veto appeared necessary. The idea was that certain values could have a "limitation" effect of the type: "if an offer has the value x on a son node then cannot be more than y on the parent node".

The results on the node 7 concerning the performances of the prototypes are presented in Table 2. We remind that such a result is an ordinal scale obtained by aggregating the four scales defined as explained in the previous section. Therefore it could be considered more as a ranking than as an absolute evaluation. For this reason the team of analysts decided to use such an attribute only in order to rank the different offers after their sorting obtained using the six quality attributes. The team of analysts experimented three different aggregation scenarios for this purpose corresponding to three different hypothesis about the importance of the performance attribute.

	O1	O2	O3	O4	O5	O6
C7	A-A	G-G	G-G	A-A	E-E	A-A

Table 2: The values of the alternatives on the performance criterion.

1. The performance attribute is considered equivalent to the set of the six quality attributes. This scenario represents the idea that the tests on the software performances correspond to the only "real" or "objective" measurement of the offers, and therefore it should be viewed as a validation of the result obtained through the subjective measurement done on the six quality attributes. The aggregation procedure consisted of using the six quality attributes as criteria equipped with a weak order. Since the evaluations on some of the six attributes were under the form of an interval, an extended ordinal scale was defined in order to induce the weak order: $E > VG > G-VG > G > A-VG > A-G > A > U$. The importance parameters are (1.):2, (2.):2, (3.):4, (4.):1, (5.):4, (6.):2 and the concordance threshold is 12/15 (0.8). The six orders are the following:
   - $O5 > O2 > O3 > O4 > O1, O6$;
   - $O2 > O5 > O3 > O4 > O6 > O1$;
   - $O2 > O4 > O3 > O5, O1, O6$;

- $O2, O4 > O3, O5 > O1, O6$;
- $O2, O5 > O3, O4 > O1, O6$;
- $O3 > O2 > O6, O4 > O5 > O1$.

The final result is presented in Table 3. Ranking (in this and the two following) has been obtained testing both “net flow” and “repeated choice” procedures (see Vincke, 1992a). The final ranking (it was the same for both procedures) is given in Figure 2a (it is worth notice that the indifference obtained in the final ranking corresponds to incomparabilities obtained at the aggregation step). An intersection was therefore done with the ranking obtained on the node 7. resulting in a final ranking reported in Figure 2b

	O1	O2	O3	O4	O5	O6
O1	1	0	0	0	0	0
O2	1	1	1	1	1	1
O3	1	0	1	0	0	1
O4	1	0	0	1	0	1
O5	1	0	0	0	1	1
O6	1	0	0	0	0	1

Table 3: The outranking relation aggregating the six quality criteria.

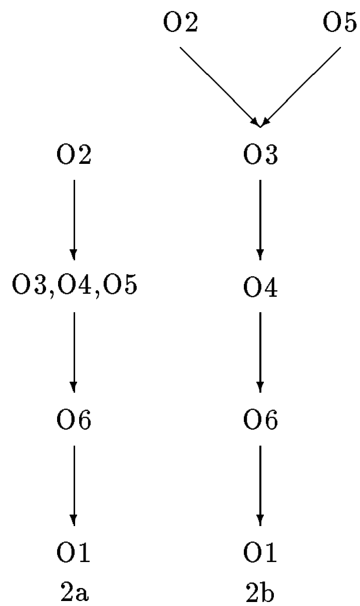

Figure 2a: The final ranking using the six quality criteria.
Figure 2b: The final ranking as intersection of the six quality criteria and the performance criterion.

2. The performance attribute is considered of secondary importance, to be used in order to distinguish among the alternatives assigned in the same class using the six quality attributes. In other words the main evaluation to be considered was the one using the six quality attributes and the performance evaluation was only a supplement enabling an eventual further distinction. Such an approach considered the performance evaluation of low confidence and did not want to assign a high importance to it. A lexicographic aggregation has been therefore applied using the six quality criteria as in the previous scenario and applying the criterion performance to the equivalence classes of the global ranking. The final ranking is $O2 > O5 > O3 > O4 > O6 > O1$.

3. A third approach consisted of considering the seven attributes as seven criteria to be aggregated in order to obtain a final ranking assigning them a reasoned importance parameter. The idea was that while the client could be interested in having the absolute evaluation of the offers

(result obtainable only using the six quality attributes), he could also be interested by a ranking of the alternatives which could help him in the final choice. Under such a point of view the absolute evaluations on the six quality attributes have been transformed in rankings as in the first scenario adding the seventh attribute as a seventh criterion. The seven weak orders are the following:
- $O5 > O2 > O3 > O4 > O1, O6$;
- $O2 > O5 > O3 > O4 > O6 > O1$;
- $O2 > O4 > O3 > O5, O1, O6$;
- $O2, O4 > O3, O5 > O1, O6$;
- $O2, O5 > O3, O4 > O1, O6$;
- $O3 > O2 > O6, O4 > O5 > O1$;
- $O5 > O2, O3 > O4, O6, O1$.
The importance parameters are (1.):2, (2.):2, (3.):4, (4.):1, (5.):4, (6.):2, (7.):4 and the concordance threshold 16/19 (more than 0.8). The final result is reported in Table 4.

	O1	O2	O3	O4	O5	O6
O1	1	0	0	0	0	0
O2	1	1	1	1	0	1
O3	1	0	1	0	0	1
O4	1	0	0	1	0	1
O5	1	0	0	0	1	1
O6	1	0	0	0	0	1

Table 4: The outranking relation aggregating the seven criteria.

The final ranking, using net flow procedure was: $O2 > O5 > O3, O4 > O6 > O1$, while using repeated choice procedure was: $O2, O5 > O3, O4 > O6 > O1$. The net flow result was finally adopted because it provided a single winner.

Finally and after some discussions with the client the third scenario has been adopted and used as the final result. The two basic reasons were:
- while it was meaningful to interpret the ordinal measures on the six quality attributes as weak orders representing the client's preferences, it was not meaningful to translate the weak order obtained on the performance attribute as an ordinal measure of the offers;
- the first and second scenario, implicitly adopted two extreme positions

concerning the importance of the performance attribute which correspond to two different "philosophies" present in the team of analysts, but not to the client's perception of the problem. The importance parameters and the concordance threshold adopted in the final version allowed to define a compromise among these two extreme positions expressed during the decision aiding process.

In fact, an importance parameter of 4 is associated to the performance criterion which, combined with the concordance threshold of 16/19, implies that it is impossible to an alternative to outrank another if its value on the performance criterion is worst (and this satisfied the part of the team of analysts who considered the performance criterion as a critical evaluation of the offers). On the other hand, giving a regular importance parameter to the performance criterion avoided the extreme situation in which all other evaluations could be irrelevant. The final ranking obtained respects this idea and the outranking table allowed all members of the team of analysts to understand it.

A final question which appeared while the final recommendation was elaborated was whether it could be possible to provide a numerical representation to the values obtained by the offers and to the final ranking. Soon it was clear that the question originated from the will of the final client to be able to negotiate with the AQ manager on a monetary basis since it was expected that (s)he would introduce the cost dimension in the final decision.

For this purpose an appendix was included in the final recommendation where the following was emphasized:
- it is possible to give a numerical representation to both the ordinal measurement obtained using the six quality attributes and to the final ranking obtained using the seven criteria, but it was meaningless to use such a numerical representation in order to establish implicit or explicit tradeoffs with a cost criterion;
- it is possible to compare the result with a cost criterion following two possible approaches:
1.) either induce an ordinal scale from the cost criterion and then, using an ordinal aggregation procedure, construct a final choice (then the negotiation should concentrate on defining the importance parameters, the thresholds etc.);
2.) or establish a value function of the client using one of the usual protocols available in the literature in order to obtain the tradeoffs among the quality evaluations, the performance evaluations and the cost criterion (then the negotiation should concentrate on such a value function);
- the team of analysts was available to conduct also this part of the decision

aid process in the case the client desired it.

The final client was very satisfied with the final recommendation and was also able to understand the reply about the numerical representation. However, he decided to conduct personally the negotiations with the AQ manager and so the team of analysts terminated its task with the delivery of the final recommendation.

4 Discussion

Some months after the conclusion of the process and the delivery of the final report, we asked to our client (the team of analysts) to discuss with us their experience and to answer some questions concerning the methodology used, how they perceived it, what did they learn and what was their appreciation. The discussion was conducted in a very informal way, but the client provided us with some written remarks which were also reported during a conference presentation (see Fiammengo et al., 1997)

This section is partly based on this discussion and remarks, besides our personal reflection on our experience.

4.1 The process

From the presentation of the process we can make the following observations.

1. It was extremely important for our client (the team of analysts) to understand his role in the process, what his client (the ISD manager) expected from them and what they were able to provide. In fact the problem situation at the beginning of the process was absolutely not clear. Moreover our client considered extremely relevant to be able to understand which were the expectations of the other actors involved in the process both for strategic reasons (having to do with organizational problems of the company) and operational reasons (recommend something reliable in a clear and sound way for all the actors involved in the bid).

   Reporting the client's remarks: *"....MCDA was very useful in organizing the overall process and structure of the bids evaluation: which were the important steps to do, how to define the call for tenders,....", "....MCDA was used in background for the whole decision process. Under such a perspective it turned out to be very useful because every activity had a justification....", "....as a formal process MCDA guaranteed greater control and transparency to the process....", "A complex* *process, like a bid, could be greatly eased by the usage of any process centered methodology."*

It is the last sentence which clearly puts emphasizes the necessity for the client to have a support along the whole process and for all its aspects, which could be able to take into account what in the decision process was happening. We actually agree with their comment that *"any process centered methodology could be useful"*, and we consider that their positive perception of MCDA depends on the fact that it was the first decision aiding approach process oriented they came in knowledge. Two other points seem to be also relevant.

2. Justification. As already reported the client considered the approach as useful because *"every activity was justified"*. A major concern for people involved into complex decision processes is to be able to justify their behaviour, recommendations and decisions towards a director, a superior in the company's hierarchy, an inspector, a committee, etc. Such a justification applies both to how a specific result has been obtained and to how the whole evaluation has been conducted.

   For instance, in our case, the choice of the final aggregation was justified by a specific attitude towards the two basic evaluation "points of view": the technical information and the performance of the prototypes. For our client to be able to recognize the correspondence between an aggregation procedure and an operational attitude was extremely important because it allowed them to better argue against the possible objections of their client.

3. Formality. Again we recall the client's remarks: *"....as a formal approach MCDA generated greater control and transparency...."*. Complex decision processes are based on human interactions and these are based on the intrinsic ambiguity of human communication (which for this reason is very efficient). However, such an ambiguity may result in an impossibility to understand and ultimately to propose viable solutions. Moreover when important stakes are considered (as in our case), decision makers may consider dangerous to make a decision without having a clear idea of the consequences of their acts. The use of a formal approach enables to reduce ambiguity (without completely eliminating it) and thus appears to be an important support to the decision process.

4.2 The information

Despite the fact we had a large amount of information to handle in our model the case did not present any exogenous uncertainty since the client considered the basic data and its judgments as reliable and felt confident with them. However, some remarks are possible.

1. The basic information available was of the type "subjective ordinal measurement". With this term we want to indicate that each alternative could be described by a vector of 134 elementary information which were in the large majority either subjective evaluations of experts (mostly part of the analyst team, our client) of the type "good", "acceptable", etc., or descriptions of the type "operational system X", "compatible with graphic engine Y", etc. The latter were expressed in nominal scales, while the former were expressed in ordinal scales. It was almost impossible that the experts could be able to give more information than such an order and was exactly such type of information which pushed the client to look for another evaluation model instead of the usual weighted sum widely diffused in software evaluation manuals and standards (see ISO/IEC 9126, 1991 and IEEE, 1992).

2. The length of the evaluation process was such that the information available at the beginning of the process could be no more valid at the end of it. This was partly due to the very rapid evolution of the GIS technology which in six months could completely innovate the state of the art. Another observation done by part of the team of analysts was that towards the end of the process due to the knowledge acquired in this period (mainly due to the process itself), they could revise some of their judgments. Actually the length of the evaluation was considered as a negative critical issue in the client's remarks.

   The final report did not consider any revision to the evaluations since in the context of a call for tenders, it could be considered unfair to modify the evaluations just before the final recommendation.

   We consider that this is a critical issue for decision support and decision aiding processes. Information is valid only for a limited period of time, and consequently the same is true for all evaluations based on such information. Moreover the client himself may revise or update his perception of the information and modify his judgments. This is rarely considered in decision aiding methodologies. While for relatively short decision aiding processes the problem may be irrelevant, it is sure that in long processes such a problem cannot be neglected and requires a specific consideration.

3. As already mentioned no exogenous uncertainty was imported in the model. However, as reported in section 3 some endogenous uncertainty was created due to the interval evaluation provided by the specific aggregation procedure and the client's unwillingness to choose one of the two values each time. The client was keen to consider the pessimistic and optimistic evaluation as bounds of the "real" value, but there was no uncertainty distribution on the interval. The problem has been handled in an ad-hoc way (as reported in section 3). We consider however, that the problem of interval evaluation on ordinal scales is an open theoretical problem which deserves more consideration in the future (very little is available in the literature to our knowledge: see Roubens and Vincke, 1985; Vincke, 1988; Pirlot and Vincke; 1997, Tsoukiàs and Vincke, 1999).

4. Obtaining the information was not a difficult task, but a time consuming process and required the establishment of an ad-hoc procedure during the process (see figure 1).

   We consider that this is also a critical issue in a decision aiding process. Often gathering and obtaining the relevant information for a decision aiding model is considered as a second level activity and therefore neglected from further specific considerations. But it can be a problem such to invalidate the problem formulation adopted if not solved. Moreover the information used in an evaluation model results from the manipulation of the rough information available at the beginning of the process. We can consider that the information is constructed during the decision aiding process and cannot be viewed as a simple input.

4.3 The model

As already mentioned in section 3, the evaluation model used in this case has been the result of a long process of interactions among the members of the team of analysts (our client) and the analysts (ourselves). Some considerations follow.

1. The model was structured around a complex hierarchy of "attributes", generally transformed into criteria through the introduction of a preference model associated to them. This is a typical situation in software evaluation (see Blin and Tsoukiàs, 1999; Stamelos and Tsoukiàs, 1998). The process of defining such an hierarchy emphasized two key issues:
   - the choice of the attributes to use;
   - the semantics of each attribute.

With respect to the first issue, a usual attitude of technical committees in charge to evaluate complex objects (as in our case) is to define an "excellence list" where every possible aspect of the object is considered. Such a list is generally provided by the literature, the experience, international standards, etc. The result is that such a list is an abstract collection of attributes independent from the specific problem on the hand, thus containing redundancies and conceptual dependencies which can invalidate the evaluation. Our client was aware of the problem, but had no knowledge and tool that could enable him to simplify and reduce the first version of the list they had defined. The repeated use of a coherence test (in the sense of Roy and Bouyssou, 1993) for each intermediate node of the hierarchy allowed to eliminate a significant amount of redundant and dependent attributes (more than 30%) and to better understand the semantics of each attribute used.

The coherence test was performed on a discoursive basis of the type: "consider a set of son nodes and choose arbitrarily one, then ask: does a difference of value on this node enables to discriminate two alternatives *ceteris paribus*"? If the answer is YES then the son node can be considered relevant, f the answer is NO then the son node is presumably irrelevant and the whole model of the parent node has to be discussed.

Despite this work, the client, in their ex-post considerations, wrote: *"....it was not necessary to be so detailed in the evaluation; the whole process could be faster because we needed the software for a due date; it could be preferable to use a limited number of criteria....".* On the other hand it is also true that it is only after the process that the client was able to determine which were the effectively significant criteria which discriminated among the alternatives.

With respect to the second issue, we pushed the client to provide us with a short description of each attribute and when a preference model was associated to it, a short description of the model (why a certain value was considered better than another). Such an approach helped the client both to eliminate redundancies (before using the coherence test which is time consuming) and to better understand the contents of the evaluation model. For instance at a certain time of the hierarchy definition process there has been a discussion about some attributes which could be considered as leaves as well as the final level of the hierarchy. Those were the so-called "process attributes", i.e., they were intended to evaluate special functionality inside different processes (in this context "process" means a chunk of functionality aiming at supporting a stream of activities of a software). In fact one can consider a process attribute (at the final level) and then subdivide it in quality aspects, or alternatively consider single independent quality aspects whose evaluation depends on how the process attribute is considered. The final choice was to put process attributes to the final level because they were coming out immediately from the evaluation scope.

Such an activity helped also the client to realize that for almost all the intermediate nodes of the hierarchy they needed an absolute evaluation of the alternatives, thus implicitly defining the problem statement of the model.

2. An important discussion with the client was the distinction between measures and preferences. As already reported the basic information consisted either of observations concerning the offers (expressed in nominal scales), or of expert judgments (expressed in ordinal scales of value of the type "good", "acceptable", etc.). All the intermediate nodes were expected to provide information of the second type. Clearly all nominal scales had to be translated into ordinal, associating a preference model on the elements of the nominal scale of the attribute. Under such a perspective it was important for the client to understand on what they were expressing their preferences. Actually the client did not compare the alternatives among them, but to a-priori defined (by the client) standards of "good", "acceptable", etc. When preferences were asked to be formulated, they concerned the elements of the nominal scales and not the alternatives themselves. The preference among the alternatives was expected to be induced once the alternatives could be "measured" by the attributes.

   Under a certain point of view we can claim that, except for the final aggregation level, the client needed to aggregate ordinal measures and not preferences (in the sense of pairwise comparisons of the alternatives). Such an observation greatly helped the client to understand the nature and scope of the evaluation model and ultimately define the problem statement of the model. Moreover the discussion on the different typologies of measurement scales helped the client to understand the problem of choosing an appropriate aggregation procedure.

3. The client greatly appreciated his involvement in the establishment of the evaluation model which turned out to be a product considered of their ownership (from their ex-post remarks: *"....this (the involvement) turned out to be important....for the acceptability of the evaluation results"*). The fact that each node of the hierarchy has been

discussed, analyzed and finally defined by the team of analysts (our client) allowed them to understand what were the consequences for the global level. It enabled them also to explain the contents of the model to their client and justify the final result on grounds of their own knowledge and experience and not on the procedure adopted.

In other words we can claim that the model was validated during its construction. Such an approach helped both the acceptability of the model and the final result, eased the discussion when the question of the final aggregation was settled, and definitely legitimated the model to the eyes of our client and of their client.

4.4 The method

The method adopted in this case study in order to aggregate the information and construct the final evaluations was a variant of the ELECTRE TRI procedure (see Yu, 1992; see also Appendix A). On the use of such a method we have the following remarks.

1. The key parameters used in the method are the profiles (to which the alternatives are compared in order to be classified in a specific class), the importance of each single criterion for each parent criterion classification, and the concepts of concordance threshold and veto condition.

   For each intermediate node, such parameters have been extensively discussed before reaching a precise numerical representation. As already mentioned in section 3, the relative importance of each criterion and the concordance threshold have been established using the reasoning based on the identification of the "winning coalitions", enabling the outranking relation to hold. Under this point of view such importance parameters do not convey any value information, but just help to define when a global preference holds or not. The veto condition was initially perceived as a theoretical possibility of no practical use (mainly due to its very severe restrictions), but very soon the client realized its importance, mainly when it was necessary to get an incomparability instead of an indifference which was a counterintuitive situation when very different objects were compared. Further on and as soon as the veto conditions have been understood by the client, they decided to introduce such a concept all the times they wanted to distinguish between positive reasons (for the establishment of the outranking relation) and negative reasons (against the establishment of the outranking relation), since they are not necessarily complementary and have to be evaluated in a separate and independent way.

The profiles have been established using the knowledge of the team of analysts (experts on their domain) which were able to identify the minimal requirements in order to qualify an object in a certain class. It is interesting to notice that the intuitive idea of a profile for the client was the one of a typical object of a class and not of the lower bound. However, the shift from the intuitive idea to the one used in the case study was immediate and presented no problems. Remain the fact that the distinction among the two concepts of profile is crucial, while the lower bound approach appears to be less intuitive than the typical element one.

2. The whole method (and the model) was implemented on a spreadsheet. This was of great importance because spreadsheets are a basic tool for communication and work in all companies and enable an immediate understanding of the results. Moreover, they enable on-line what-if operations when specific problems, concerning precise information and/or evaluation, appeared during the discussions inside the team of analysts. The experimental validation of the model was greatly eased by the use of the spreadsheet.

   Further on it helped the acceptability and legitimation of the model through the idea that *"if it can be implemented on a spreadsheet it is sufficiently simple and easy to be used by our company"*. In fact some of the critics of the client about the approach adopted in this case were that *"....MCDA is not yet an universally known method...."*, *"....seems less intuitive than other well known techniques as the weighted sum..."*, *"....it is time consuming to apply a new methodology...."*, all these problems limiting the acceptability of the methodology towards the client of our client and the company, more generally. Being able to implement the method and the model on a spreadsheet was, for them, a demonstration that, although new, complex and apparently less intuitive the method was simple and easy and therefore legitimately used in the decision process.

3. It is sure that there was space (but no time) to experiment more variants and methods as far as the aggregation procedure and the construction of the final recommendation were concerned. Valued relations, valued similarity relations, interval comparisons using extended preference structures, dynamic assignment of alternatives to classes and other innovative techniques were considered too "new" by the client who considered already as a revolution (compared with the company's standards) the use of an approach different from the usual

grid and weighted sum. To their eyes, the fact to be able to aggregate in a correct and meaningful way the ordinal information available was more than satisfactory as they report in their ex-post remarks: *"....pointed out that it was not necessary to always use ratio scales and weighted sums, as we thought in advance, but that it was possible to use judgments and aggregate them...."*.

5 Conclusion

In this paper we extensively present and discuss a real world application of MCDA in the domain of software evaluation. The problem was to support an analysis team in following a call for tenders for the acquisition of a GIS, such an acquisition representing an important stake for the company.

The paper presents the whole decision aiding process, the actors and the objects included in the process, besides the main products of such an activity:
- a recognition of the problem situation enabling more general recommendation to how the process could be conducted and how the team of analysts could participate;
- some problem formulations in order to better understand what the analysts team was expected to provide to its client (a technical evaluation of the offers to the bid, but including a value judgment of the type "good", "acceptable", etc.);
- an evaluation model suitable for the problem formulation previously adopted, which implied an extensive work on the definition of the hierarchy of attributes and the aggregation procedure to use;
- a final recommendation to the client of the team of analysts about the offers considered in the bid.

On the basis of our experience and of some ex-post remarks provided (informally) by our client. we present in the paper a discussion of the most relevant issues (to our opinion) of the decision aiding process.

Such key issues include the usefulness of any process oriented decision aiding approach in supporting the whole decision process of the client, the importance of a clear understanding by the client of the model constructed, the acceptability of the methodology in terms of easiness and simplicity, the legitimation of the client in the process and of the analysts role, the flexibility in using specific tools (like aggregation procedures and their relative parameters) in a way which should combine correctness in manipulating the information (formal meaningfulness) and clients understanding (clients meaningfulness).

Although our client appeared totally satisfied by our collaboration and the obtained results, we are sure that more and better was possible (as always happen in such cases). But this has to do with our human limits. Next time we will try better.

Acknowledgments

The authors gratefully acknowledge the contribution of Daniele Buosi, Anna Fiammengo, Ivano Iob, Paolo Maffioli and Gianfranco Panarotto who practically conducted the experience and allowed us to report it. It is clear that the opinions and conclusions expressed in the paper strictly represent the personal point of view of the authors and do not engage in whatever direction their institutions.

References

1. Bana e Costa C., Ensslin L., Corrêa E.C., Vansnick J.C., (1997), "Decision Support Systems in Action: Integrated Application in a Multicriteria Decision Aid Process", forthcoming in *European Journal of Operational Research.*

2. Belton V., Ackermann F., Shepherd I., (1997), "Integrated Support from Problem Structuring through Alternative Evaluation Using COPE and VISA", *Journal of Multi-Criteria Decision Analysis*, vol. 6, 115 - 130.

3. Blin M.J., Tsoukiàs A., (1999), "Evaluation of COTS using multicriteria methodology", in Proceedings of the 6th European Conference on Software Quality, 429 - 438.

4. Bouyssou D., Perny P., Pirlot M., Tsoukiàs A., Vincke Ph., "The manifesto of the new MCDA era", (1993), *Journal of Multi-Criteria Decision Analysis*, vol. 2, 125 - 127.

5. Checkland P., (1981), *Systems Thinking Systems Practice*, J. Wiley, Chichester.

6. Fiammengo A., Buosi D., Iob I., Maffioli P., Panarotto G., Turino M., (1997), "Bid management of software acquisition for cartography applications", presented in *AIRO 97 conference*, Aosta.

7. French S. (1998), "Decision Making not Decision Theory", *Journal of Multi-Criteria Decision Analysis*, vol. 7, 303.

8. Heurgon E., (1982), "Relationships between decision making process and study process in OR interventions", *European Journal of Operational Research*, vol. 10, 230 - 236.

9. Humphreys P.C., Svenson O., Vári A., eds., (1983), *Analysis and aiding decision processes*, North Holland, Amsterdam.

10. IEEE 92, The Institute of Electrical and Electronics Engineers, (1992), *Standard for a Software Quality Metrics Methodology*, December 1992

11. ISO/IEC 9126, (1991), *Information Technology - Software Product Evaluation, Quality Characteristics and guidelines for their use*, ISO, Genève.

12. Jacquet-Lagrèze E., Moscarola J., Roy B., Hirsch G., (1978), "Description d'un processus de décision : quelques concepts", *Cahier du LAMSADE N. 13*, Université Paris Dauphine, Paris.

13. Masser I., (1983), "The representation of urban planning- processes: an explorary review", *Environment and Planning B*, vol. 10, 47 - 62.

14. Mintzberg H., Raisinghani D., Theoret A., (1976), "The structure of unstructured decision processes", *Administrative Science Quarterly*, vol. 21, 246 - 272.

15. Morisio M., Tsoukiàs A., (1997), "IUSWARE: A formal methodology for software evaluation and selection", *IEE proceedings on software engineering*, vol. 144, 162 - 174.

16. Moscarola J., (1984), "Organizational Decision Processes and ORASA intervention", in R. Tomlinson, I. Kiss, (eds.), *Rethinking the process of Operational Research and Systems Analysis*, Pergamon Press, Oxford, 169 - 186.

17. Norese M.F., (1988), "A multi-dimensional model by a multi-actor system", in B.R. Munier and M.F. Shakun (eds.), *Compromise, Negotiation and Group Decision*, Kluwer Academic, Dordrecht, 263 - 276.

18. Nutt P.C., (1984), "Types of organizational decision processes", *Administrative Science Quarterly*, vol. 19, 414 - 450.

19. Ostanello A., (1990), "Action Evaluation and Action Structuring different decision aid situations reviewed through two actual cases", in C.A. Bana e Costa (ed.), *Readings in MCDA*, Springer Verlag, Berlin, 36 - 57.

20. Ostanello A., (1997), "Validation Aspects of a Prototype Solution Implementation to solve a complex MC problem", in J. Clímaco (ed.), *Multicriteria Analysis*, Springer Verlag, Berlin, 61 - 74.

21. Ostanello A., Tsoukiàs A., (1993), "An explicative model of 'public' interorganizational interactions", *European Journal of Operational Research*, vol. 70, 67 - 82.

22. Pirlot M., Vincke Ph., (1997), *Semi-orders: properties, representations, applications*, Kluwer, Dordrecht.

23. Rosenhead J., (1989), *Rational Analysis of a Problematic World*, J. Wiley, New York.

24. Roubens M., Vincke Ph., (1985), *Preference Modeling*, Springer Verlag, Berlin, LNEMS 250.

25. Roy B., (1996), *Multicriteria Methodology for Decision Aiding*, Kluwer Academic, Dordrecht, translation of *Méthodologie multicritère d'aide à la décision*, Economica, Paris, 1985.

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

27. Simon H.A., (1957), "A behavioural model of Rational Choice", in *Models of Man*, J. Wiley, New York, 241 - 260.

28. Stamelos I., Tsoukiàs A., (1998), "Software Evaluation Problem Situations", *Cahier du LAMSADE N. 156*, Université Paris Dauphine, Paris.

29. Svenson O., (1996), "Decision Making and the search for fundamental psychological regularities: what can be learned from a process perspective?", *Organizational, Behavior and Human Decision Processes*, vol. 65, 252 - 267.

30. Tabucanon M.T., Chankong V., (eds.), (1989), *Multiple Criteria Decision Making: applications in industry and service*, Asian Institute of Technology, Bangkok.

31. Tsoukiàs A., Vincke Ph., (1999), "On the generalization of interval orders", submitted.

32. Vincke Ph., 1988, "P,Q,I preference structures", in J. Kacprzyk, M. Roubens, eds., *Non conventional preference relations in decision making*, LNEMS 301, Springer Verlag, Berlin, 72 - 81.

33. Vincke Ph., (1992a), "Exploitation of a crisp relation in a ranking problem", *Theory and Decision*, vol. 32, 221 - 240.

34. Vincke Ph., (1992b), *Multicriteria Decision Aid*, J. Wiley, New York.

35. Yu, W., (1992), "Aide multicritère à la décision dans le cadre de la problématique du tri: méthodes et applications", PhD Thesis, LAMSADE, Université Paris Dauphine, Paris.

Appendix A

The basic concepts adopted in the ELECTRE TRI type procedure used are the following.

- A set A of alternatives a_i, $i = 1, \cdots, m$.
- A set G of criteria g_j, $j = 1, \cdots, n$. To each criterion g_j a relative importance w_j is associated.
- Each criterion g_j is equipped with an ordinal scale $\mathcal{E}_j$ with degrees e_j^l, $l = 1, \cdots, k$.
- A set $\mathcal{P}$ of profiles p_h, $h = 1, \cdots, t$, p_h being a collection of degrees, $p_h = \langle e_1^h, \cdots, e_n^h \rangle$, such that if e_j^h belongs to profile p_h, e_j^{h+1} cannot belong to profile p_{h-1}.
- A set $\mathcal{C}$ of categories c_λ, $\lambda = 1, \cdots, t+1$, such that the profile p_h is the upper bound of category c_h and the lower bound of category c_{h+1}.
- An outranking relation $S \subset A \times \mathcal{P} \cup \mathcal{P} \times A$, where $s(x, y)$ should be read as "x is at least as good as y".
- A set of preference relations $\langle P_j, I_j \rangle$ for each criterion g_j such that:
  - $\forall x \in A, \;\; P_j(x, e_j^h) \;\Leftrightarrow\; g_j(x) > e_j^h$,
  - $\forall x \in A, \;\; P_j(e_j^h, x) \;\Leftrightarrow\; g_j(x) < e_j^h$,
  - $\forall x \in A, \;\; I_j(x, e_j^h) \;\Leftrightarrow\; g_j(x) = e_j^h$,

  $>$ being induced by the ordinal scale associated with criterion g_j.

The procedure works in two basic steps.

1. Establishment of the outranking relation on the basis of the following rule:

$$s(x, y) \;\Leftrightarrow\; C(x, y) \wedge \neg D(x, y)$$

where

$$\forall x \in A,\; y \in \mathcal{P} : \; C(x, y) \;\Leftrightarrow\; \sum_{j \in G^{\pm}} w_j \geq c \wedge \sum_{j \in G^{+}} w_j \geq \sum_{j \in G^{-}} w_j$$

$$\forall y \in A,\ x \in \mathcal{P}:\ C(x,y)\ \Leftrightarrow$$

$$\Big(\sum_{j\in G^{\pm}} w_j \geq c\ \wedge \sum_{j\in G^{+}} w_j \geq \sum_{j\in G^{-}} w_j\Big)\ \vee\ \sum_{j\in G^{+}} w_j > \sum_{j\in G^{-}} w_j$$

$$\forall (x,y) \in A\times\mathcal{P}\cup\mathcal{P}\times A:\ \neg D(x,y)\ \Leftrightarrow\ \sum_{j\in G^{-}} w_j \leq d \wedge \forall g_j\ \neg v_j(x,y)$$

where
- $G^{+}\ =\ \{g_j \in G:\ P_j(x,y)\}$
- $G^{-}\ =\ \{g_j \in G:\ P_j(y,x)\}$
- $G^{=}\ =\ \{g_j \in G:\ I_j(x,y)\}$
- $G^{\pm}\ =\ G^{+}\cup G^{=}$
- c: the concordance threshold $c \in [0.5, 1]$
- d: the discordance threshold $d \in [0, 1]$
- $c + d \neq 1$
- $v_j(x,y)$: veto, expressed on criterion g_j, of y on x.

2. When the relation S is established, then assign any element a_i on the basis of the following rules.

   2.1 pessimistic assignment
   - a_i is iteratively compared with $p_t \cdots p_1$,
   - as soon as is established $s(a_i, p_h)$, then assign a_i to category c_h.

   2.2 optimistic assignment
   - a_i is iteratively compared with $p_1 \cdots p_t$,
   - as soon as is established $s(p_h, a_i) \wedge \neg s(a_i, p_h)$, then assign a_i to category c_{h-1}.

The pessimistic procedure finds the profile for which the element is not worst. The optimistic procedure finds the profile against which the element is surely worst. If the optimistic and pessimistic assignments coincide, then no uncertainty exists for the assignment. Otherwise, an uncertainty exists and should be considered by the user.

Appendix B: list of attributes

1 LAND-BASE MANAGEMENT

1.1 User interface

1.1.1 Graphics type

1.1.2 Graphics engine adequacy

1.1.3 Interface personalization

1.2 Functionality

1.2.1 Availability

1.2.2 Adequacy

1.2.2.1 Planes analysis functions

1.2.2.2 Topological connectivity functions

1.2.2.3 Graphical rendering functions

1.3 Development environment

1.3.1 Libraries personalization

1.3.2 Development support tools

1.3.3 Debugging support tools

1.3.4 Code documentation

1.3.4.1 Documentation support tools

1.3.4.2 Code browsing

1.3.5 Documentation Quality

1.3.5.1 Completeness

1.3.5.2 Documentation support type

1.3.5.3 Information retrieval ease

1.3.5.4 Contextual help

1.4 Administration tools

1.4.1 User administration functions

1.4.2 Software configuration management

1.4.3 Performance data collection

1.5 Work flow connection

1.6 Interoperability

1.7 Integration between Land-base products and the Spatial Data Manager

1.7.1 Vectorial data products integration

1.7.2 Descriptive data products integration

1.7.3 Raster data products integration

1.7.4 Digital Terrain Model products integration

1.8 Integration among Land-base products

1.8.1 Interfaces integration

- 1.8.2 Data sharing

2 GEOMARKETING

- 2.1 User interface
  - 2.1.1 Graphics type
  - 2.1.2 Graphics engine adequacy
  - 2.1.3 Interface personalization
- 2.2 Functionality
  - 2.2.1 Availability
  - 2.2.2 Adequacy
    - 2.2.2.1 Planes analysis functions
    - 2.2.2.2 Graphical rendering functions
- 2.3 Development environment
  - 2.3.1 Libraries personalization
  - 2.3.2 Development support tools
  - 2.3.3 Debugging support tools
  - 2.3.4 Code documentation
    - 2.3.4.1 Documentation support tools
    - 2.3.4.2 Code browsing
  - 2.3.5 Documentation Quality
    - 2.3.5.1 Completeness
    - 2.3.5.2 Documentation support type
    - 2.3.5.3 Information retrieval ease
    - 2.3.5.4 Contextual help
- 2.4 Administration tools
  - 2.4.1 Software configuration management
- 2.5 Interoperability
- 2.6 Integration between Geomarketing products and the Spatial Data Manager
  - 2.6.1 Vectorial data products integration
  - 2.6.2 Descriptive data products integration
  - 2.6.3 Raster data products integration
- 2.7 Integration among Geomarketing products
  - 2.7.1 Interfaces integration
  - 2.7.2 Data sharing

3 PLANNING, DESIGN, IMPLEMENTATION AND OPERATING SUPPORT

- 3.1 User interface
  - 3.1.1 Graphics type

4.1.1 Graphics type

4.1.2 Graphics engine adequacy

4.1.3 Interface personalization

4.2 Functionality

4.2.1 Availability

4.2.2 Adequacy

4.2.2.1 Planes analysis functions

4.2.2.2 Topological connectivity functions

4.2.2.3 Graphical rendering functions

4.2.2.4 Network schema creation

4.3 Development environment

4.3.1 Libraries personalization

4.3.2 Development support tools

4.3.3 Debugging support

4.3.4 Code documentation

4.3.4.1 Documentation support tools

4.3.4.2 Code browsing

4.3.5 Documentation Quality

4.3.5.1 Completeness

4.3.5.2 Documentation support type

4.3.5.3 Information retrieval ease

4.3.5.4 Contextual help

4.4 Administration tools

4.4.1 Software configuration management

4.4.2 Performance data collection

4.5 Interoperability

4.6 Integration between this process products and the Spatial Data Manager

4.6.1 Vectorial data products integration

4.6.2 Descriptive data products integration

4.6.3 Raster data products integration

4.7 Integration among this process products

4.7.1 Interfaces integration

4.7.2 Data sharing

5 SPATIAL DATA MANAGER

5.1 Data base properties

5.1.1 Fundamental properties

5.1.2 Transaction typology support