Integration of Multi-Criteria Tools in Geographical Information Systems

Karim LIDOUH

Thèse présentée en vue
de l’obtention du grade de
Docteur en Sciences de l’Ingénieur

ULB ECOLE POLYTECHNIQUE DE BRUXELLES

Computer and Decision Engineering (CoDE)
Ecole polytechnique de Bruxelles
Université libre de Bruxelles

Belgium
November 12, 2013
Acknowledgments

I would like to begin by thanking my two supervisors, Esteban Zimányi and Yves De Smet, for giving me the opportunity to integrate their teams and work on some very interesting subjects. This thesis is the result of almost 6 years of research during which I have benefited plenty of all the encounters I made.

I wish to thank the members of my jury for their interest in my work: Olivier Debeir, Bertrand Mareschal, Vincent Mousseau, Marc Pirlot, Philippe Vincke, Eleonore Wolff. I especially thank Bertrand Mareschal, Marc Pirlot, Philippe Vincke, and Eleonore Wolff for the yearly meetings we had to discuss my progress.

All of the colleagues I have met in the different departments I was involved in also deserve my most sincere thanks for their support, their encouragements, and help during the course of this thesis: Maryam Bagherikahvarin, Issam Banamar, Mohamed Ayman Boujelben, Nguyen Anh Vu Doan, Stefan Eppe, Julien Roland, Renaud Sarrazin, Dimitri Van Assche, and Céline Verly of the CoDE-SMG unit for the discussions we had and their judicious remarks and comments; Serge Boucher, Frédéric Servais, Stijn Vansummeren, and Boris Verhaegen of the CoDE-WIT unit for their advice; Catherine Dehon, Jean-Jacques Droesbeke, El Maati Allaoui, Christopher Bruffaerts, Virginie Maghe, Mohamed Sraieb, and Catherine Vermandele of the Solvay Brussels School and the Faculté des Sciences Sociales et Politiques for their encouragement.

Of course I could not possibly list all of the friends and family members who gave me the little push I needed to go on whenever I needed it. This has been a long and tedious journey, but I learned and experienced things I would not have otherwise.
Abstract

For a little over twenty years, researchers have worked on integrating multi-criteria aggregation procedures (MCAP) to GIS. Several notable contributions have brought this field to what it is today. After studying the course of MCDA-GIS integration through several works, we question the future of such an attempt. Indeed most works that aim for an integration do not survive long after their direct purpose has been fulfilled. We end up understanding through a critical review of the existing systems that technical integration means nothing if it is not visible to the user on an operational level.

We therefore propose several contributions to improve the usability of MCDA methods in the context of GIS. One of our works consists in adapting the PROMETHEE-GAIA methodology to be used on maps for spatially referenced problems. To do so, we define symbols/glyphs that display select parts of the results obtained through the PROMETHEE and GAIA methods. This allows for the comparison of alternatives’ profiles and characteristics based on their geographic location which wasn’t possible before. This adaptation helps us combine multicriteria and geographic aspects in an entirely new way. We also propose some extensions of the GAIA method to improve the quality of the results and reduce the risk of wrong interpretations to be made due to losses of data.
# Contents

<table>
<thead>
<tr>
<th>Acknowledgments</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Problem</td>
<td>1</td>
</tr>
<tr>
<td>Objective</td>
<td>2</td>
</tr>
<tr>
<td>Outline</td>
<td>2</td>
</tr>
<tr>
<td>Contributions</td>
<td>3</td>
</tr>
<tr>
<td>Publications</td>
<td>5</td>
</tr>
</tbody>
</table>

## 1 Decision Support Tools in MCDA

1.1 Introduction ........................................... 7
1.1.1 Historical Background ............................ 7
1.1.2 Mono-Objective vs Multicriteria Decision Aid .... 9
1.1.3 Modeling Features ................................. 10
1.2 Main Concepts .......................................... 12
1.2.1 Definitions ......................................... 12
1.2.1.1 Action/Alternative ............................... 12
1.2.1.2 Criterion ......................................... 13
1.2.1.3 Problem .......................................... 14
1.2.1.4 Dominance and Efficiency ........................ 16
1.2.2 Classification of methods .......................... 16
1.3 The American School .................................... 17
1.3.1 Weighted Sum ........................................ 17
1.3.2 AHP .................................................. 19
1.3.3 TOPSIS .............................................. 20
1.3.4 Graphical Support Tools ............................ 22
1.4 The European School .................................... 25
4.2.3 The Late 2000’s: A Look at the Future

5 Reasons and Objective for the Integration

5.1 Introduction

5.2 The Integration of MCDA and GIS

5.3 Technical Aspects of the Integration

5.4 Usability Aspects of the Integration

5.5 Conclusion

6 Extensions and Variants of Existing MCDA Methods

6.1 Introduction

6.2 The GAIA Methods

6.3 The Weighted GAIA Methods

6.4 Numerical Examples and Simulations
Introduction

Nowadays Geographical Information Systems (GIS) have become quite popular. Being available in simplified forms such as Google’s famous Google Maps® or Google Earth® services, Microsoft’s Bing Maps®, or the numerous GPS devices that guide us, GIS are now essential to our everyday life. However these simple and easy to use interfaces hide complex systems that have many more uses than just calculating routes. Indeed these systems are used by companies, governments, and several other associations to analyze spatial data and produce maps. In order to improve these systems, developers have enhanced them by including additional analysis tools in their design. In this thesis we will take interest in the integration of multi-criteria tools with GIS: tools that focus on supporting decision-making with multiple criteria.

Problem

For a little over twenty years, researchers have worked on integrating multi-criteria aggregation procedures (MCAP) to GIS. Several notable contributions have brought this field to what it is today. Within this thesis, we take a look at what has been achieved and what is needed as the next breakthrough.

The first works that focussed on having a conjoint use of multi-criteria methods and spatial aspects can be traced back to the 1950’s [Tyrwhitt, 1950], however it was only in the 1990’s that the first actual implementations of those two tools came to make an appearance [Eastman et al., 1993, Laaribi et al., 1996]. This delay was mainly caused by the time it took for geographical information systems to truly become standardized and start using well defined sets of tools. That step was necessary in order for them to become democratized and open to the public. When that happened, the integration of several statistics and operations research tools in GIS felt natural as these systems strive to become complete spatial decision support systems (SDSS).
Objective

Multi-criteria Decision Aiding (MCDA) was born from the need of taking several objectives into account when making decisions [Evans and Steuer, 1973, Keeney, 1972]. It is a discipline that encompasses several aggregation procedures to be used in choice problems, ranking problems, sorting problems, and many more. It also proposes tools that can help in the modeling of a problem or the analysis of its results. Through the years this field has given birth to numerous methodologies designed to reflect the different ways of thinking of a decision maker. These many tools and their respective properties make this field quite complex and difficult to take in for novices.

On the other hand Geographical Information Systems have come a long way and developed a set of spatial analysis tools. They can help with the exploration and evaluation of potential solutions and their visualization functionalities allow them to present results to the decision makers. Indeed, by using the contiguity and proximity relationships as well as overlay functionalities, GIS are able to select solutions that verify certain constraints or compute spatial evaluations for them. What these systems are missing is the possibility to aggregate several factors in order to compare these alternatives.

Spatial decision problems usually target existing situations with several stakeholders and objectives. It is therefore clear that by using multi-criteria methods along with other spatial analysis tools we can avoid neglecting some aspects and obtain a fair result. Unfortunately, the path towards a complete integration of these two disciplines is difficult both conceptually and technically. That is why we will first focus on determining the actual benefits that could come from a complete integration and orient our efforts towards the remaining gaps.

Outline

In the first chapter, we will be taking an overview of the main tools of the MCDA field and draw attention to the ones that can best combine themselves with GIS, namely visual or supporting tools for the application of MCDA methods.

Through a case study involving the Greater London Region, we will compare a few selected methods in the second chapter. This will allow us to demonstrate how useful the support tools in MCDA can be to find a robust result. These can even help when some information about the problem is missing or the analyst is not in direct contact with the decision maker.

The third chapter presents the analysis tools used in GIS as well as some concepts of information visualization. We deem it important to point out what possible spatial criteria could be evaluated using a GIS. For that purpose, GIS naturally have several
options linked to the proximity on contiguity relationships between alternatives. They are unfortunately lacking in terms of analytical functionalities. The tools available for information visualization will later help us in adapting MCDA methods for cartography purposes.

The fourth chapter is a state of the art of the MCDA-GIS integration during the last thirty years. It will describe the works that have been done by selected researchers on the integration between MCDA and GIS. We will use it to summarize the several conclusions that have been expressed and use them to establish clear objectives for future research. Indeed several tracks have been neglected up till now and may prove to be of great value for the design of integrated systems.

In the fifth chapter, we take a look at some integration strategies and explore the reasons that have led to such developments. Even though several researchers had already identified reasons for the MCDA-GIS integration, we find it necessary to reassess those motivations, now that the field has advanced and some solutions exist. We take a look at some of the existing MCDA-GIS implementations and try to see if the reasons objectives enunciated in the third chapter still hold.

In the sixth chapter, we study some extensions to existing MCDA methods. We see how to improve the quality of the results in the GAIA method by taking the weights into account. This variant of the visual tool will help us make better interpretations once we start using it on spatially referenced problems.

Finally, after a reevaluation of the tracks yet to be explored and that present an interest for the MCDA-GIS integration, we take a deeper look at one of them: the visual representation of MCDA results using GIS. From all the possible gains of integrating MCDA and GIS, this one seems to have the greater value for improving the dialogue between an analyst and a decision maker. We therefore focus on adapting some visual MCDA tools for use on geographical maps. The GAIA visualization tool presented in the first chapter will turn out to be well suited for this purpose. Through two types of glyphs/symbols, we will try to represent the information given to us by the tool on a map. We will then propose some enhancements in order to solve some of the issues we encounter such as loss of data.

Contributions

In this thesis, after studying the course of MCDA-GIS integration through several works, we question the future of such an attempt. Indeed most works that aim for an integration do not survive long after their direct purpose has been fulfilled. We end up understanding through a critical review of the existing systems that the answer is to realize what the
user wants.

As such, we will continue our works while aiming to improve the usability of any system that brings together MCDA and GIS. To do so, we focus on visual tools and the help they bring to communicate and justify results after an analysis. We decide to base our work on the existing MCDA visual tool GAIA, which is part of the PROMETHEE family of outranking methods. The reason for this choice is that there are hardly any articles on works that tried to combine the PROMETHEE methodology with GIS and none tried to represent the results obtained with GAIA on maps. Our strategy however, will not be reduced to a simple implementation of the tool. Instead, we will break down the results obtained and completely adapt them to a representation intended for cartography purposes.

As the transformation of results from the GAIA plane to a geographic map suffers from a certain loss of data, we also develop a variant of the GAIA tool designed to increase the quality of the results displayed. By taking additional parameters into account before applying the GAIA procedure, we obtain a representation of the decision problem that better reflects the view of the decision maker. We can thereby make better interpretations of the results and
Publications

The works described in this thesis have led to several publications in refereed journals and conference proceedings:

- **Information Visualisation IV09 [Lidouh et al., 2009b]:**
  

- **Algorithmic Decision Theory ADT2009 [Lidouh et al., 2009a]:**
  

- **OR Society Conference OR52 [Lidouh et al., 2010]:**
  

- **Industrial Engineering and Engineering Management IEEM2010 [Verly et al., 2010]:**
  

- **Symposium Series in Computational Intelligence SSCI2011 [Lidouh et al., 2011]:**
  
  Karim Lidouh, Yves De Smet and Esteban Zimanyi. "An Adaptation of the GAIA Visualization Method for Cartography: Using the HSV color

• International Journal of Information and Decision Sciences IJIDS [Nemery et al., 2011]:


• Tourism Management [Ishizaka et al., 2013]:


• International Journal of Multicriteria Decision Making IJMCDM [Lidouh, 2013]:


• Business Intelligence [De Smet and Lidouh, 2013]:

Chapter 1

Decision Support Tools in MCDA

1.1 Introduction

Multi-criteria Decision Aid (MCDA) offers a set of models that can be formalized in order to support an real-life decision problem. Known under several names such as Multi-criteria Decision Aid (MCDA), Multi-criteria Decision Making (MCDM), Multi-criteria Analysis (MCA), it is above all a decision aid activity that exists to support the decision maker instead of making the decision in his place [Roy, 1993]. The particularity of MCDA compared to other decision aiding techniques is that it tries to take several factors into account when giving a recommendation.

In this chapter, we present some of the differences between MCDA and other operational research approaches, we define the main concepts, and we take a look at two of the main approaches used to solve decision problems. After describing some of the methods and their associated tools, we proceed with an illustrative example that allows us to apply different methods and compare their results and the way they work in practice.

The first section will explain the framework in which MCDA is set by explaining its origin and particular approach. Section 1.2 will describe some of the main concepts in MCDA. Finally sections 1.3 and 1.4 will describe two different approaches in MCDA: the weighted summation and the outranking methods.

1.1.1 Historical Background

MCDA evolved from notions that came from very different backgrounds: goal programming [Zionts and Wallenius, 1976], sequential decision processes [Howard and Kimball, 1959], utility theory [Keeney and Raiffa, 1976], directed network of preferences [Roy, 1971] ... It is therefore difficult to set a date at which it appeared or was established.

After World War II, occidental countries witnessed an important economic growth. This development gave rise to numerous studies in the economic field and public manage-
ment. In particular, managers focused on the subject of evaluating alternatives and choosing an optimal solution. Mathematical programming and cost-benefit analyzes gained in popularity and dominated other optimization methods until the 1970’s [Nijkamp et al., 1990]. These methods however were based on the maximization of economic well-being and could hardly take qualitative factors into account.

Indeed linear programming optimization methods failed to include all the aspects of a realistic decision problem such as the socio-economic aspects of land-use management, the environmental impacts of transportation networks, or the ethical responsibility in societal problem analysis. The main reason for that, is that these methods are based on the evaluation of a single criterion, formulated by an objective function. At that time most approaches in operational research mainly pursued the evaluation of alternatives in a normative or prescriptive way. They assumed that the information is complete and thorough, that the decision maker is rational and that the solution to be found should be optimal. In his social choice theory, Arrow listed the axioms that an aggregation function should verify [Arrow, 1951].

It was the restrictive framework of application of these methods that led researchers to explore different approaches. Several were proposed but the most famous was the one by Simon [Simon, 1960]. It suggested decomposing the decision process in three steps: Intelligence (gathering of the information about the problem), Design (generating the possible solutions and defining the criteria), Choice (evaluating the solutions and choosing the best one). This structure in which the evaluation is but a step in a bigger process was later extended to include several other details and constitute different approaches. Its main appeal is that it is based on a descriptive approach in which it works on the representation of the decision maker’s preference structure and tries to explain how decisions should be made in the reality. It assumes that the information used is subjective and incomplete and that the decision maker’s cognitive abilities are limited and do not allow them to be fully rational. Furthermore, the solution to be found is expected to be satisfactory instead of optimal.

In the beginning of the 1970’s, it became clear that linear programming methods and cost-benefit analyzes needed to be extended in order to take several conflicting factors into account. This gave birth on the one hand to multiobjective mathematical programming as the natural extension of linear programming methods [Evans and Steuer, 1973] and on the other hand to multicriteria analysis as an extension to the cost-benefit analysis [Keeney, 1972, Keeney and Raiffä, 1976]. After 1975, several multicriteria evaluation methods were developed and built the foundations of multicriteria decision aiding as we know it [Keeney and Raiffä, 1976, Roy, 1976, Saaty, 1980, Brans, 1982, Roy, 1991].
1.1.2 Mono-Objective vs Multicriteria Decision Aid

For over 50 years, optimization models have helped solve decision problems and are still widely used. However nowadays, we are faced with an increasing number of problems that require us to take several non-economic aspects into account. We can therefore distinguish two very different approaches to problem solving: the first one focusing on the optimization of a single criterion and aiming at finding an optimal solution, and another one taking several conflicting factors into account in order to find a best compromise solution.

Mono-objective problems can be formulated as follows:

\[ x^* = \text{argmax}\{ g(x) | x \in A \} \]  

where \( g \) is the objective function that needs to be optimized (maximized in this specific case), \( A \) is the set of feasible solutions and \( x^* \) is the optimal solution. These are problems in which the set of alternatives has been defined explicitly (i.e. enumeration) or implicitly (i.e. by formulating constraints) and where we are looking for the solution \( x \) for which the objective function \( g(x) \) takes an optimal value. As an example, we could think of a transportation problem for which we aim to find the minimum total cost. This would depend on several inputs which would all be monetized to be included in a single function (fuel, time, salaries, \( CO_2 \) emissions...).

Mono-objective problems are typically well defined and mathematically well written. However as these models focus on optimizing a single objective (e.g. maximizing profit, minimizing costs, minimizing delays, etc). In the real world, decision problems are rarely mono-objective and usually involve taking several objectives or opinions into account simultaneously. For example, the construction of a new train track would cause several issues to be considered. These could be of very different natures: economic (cost, return), social (mobility, noise), environmental (pollution).

Multi-criteria problems will aim at optimizing several objective functions at once:

\[ \text{opt}\{ g_1(x), g_2(x), ..., g_q(x) | x \in A \} \]

where \( g_1, g_2, ..., g_q \) are the objective functions. Multi-criteria models are richer since they are closer to the real problems. Unfortunately, by pursuing several objectives, the problem becomes mathematically poorly written. Indeed, in such problems, there is usually no optimal solution \( x \) that is better than all the other alternatives when considering all criteria simultaneously.

As there is no optimal solution to be found, multi-criteria problems will typically follow two trains of thought: (1) the objective view that consists in finding a set of Pareto-optimal solutions with respect to a set of criteria, and (2) the subjective view that takes...
the preferences of the decision maker into account to find the compromise that best suits them. In order to find the solution that would best satisfy them, the analyst will need to collect information about the preferences, the weights, and several other characteristics of the problem. Due to the way the alternatives will be compared and the models are built, studies of multi-criteria problems will often exhibit distinctive features.

1.1.3 Modeling Features

One of the features brought by MCDA is the possibility to model preferences using more elaborate relationships than before. When comparing two alternatives and considering only one single criterion, we can identify two cases represented by the following relationships:

1. The indifference relationship $I$: The first alternative $a$ is not better than the second one $b$ and $b$ is not better than $a$. It is written $aIb$.

2. The preference relationship $P$: The first alternative $a$ is better than the second one $b$. It is written $aPb$.

We refer the interested reader to the works of Roy for more complete definitions of these relationships [Roy, 1976, Roy, 1991, Roy, 1996].

When considering several criteria however, we are no longer limited to just these two situations.

The first feature one can identify in multi-criteria problems is the possibility of incomparability between alternatives. It happens in cases where the decision maker is unable to choose the best alternative between two possibilities due to the heterogeneity of their profiles or the imperfection of the information he possesses [Schärlig, 1985].

Figure 1.1 shows us a situation with 4 alternatives evaluated according to two criteria (to be minimized) where the alternatives $b$ and $c$ are incomparable. Indeed, we can see that alternative $a$ is better or at least as good as any of the three other alternatives and that alternative $d$ is never better than the others. However alternatives $b$ and $c$ are such that each one of them is better than the other on a specific criterion. That means that without any additional information, it is impossible to tell with certainty which one is better. That situation can be represented by a graph that shows the relations “is considered better or at least as good as” between each pair of alternatives. The numbers in the boxes indicate the rank associated to each alternative. Even though this example uses only two dimensions, it can be expanded to take more criteria into account.

The second feature is the intransitivity of the relations between alternatives. Typically, a binary relationship $R$ (i.e. a collection of ordered pairs of elements) will be transitive if
Figure 1.1: A case of incomparability between two alternatives

the following expression between three elements is true:

\[ aRb \land bRc \implies aRc \]  

meaning that if \( a \) is related to \( b \) and \( b \) is related to \( c \), then \( a \) is related to \( c \).

We can show that this is not true for multi-criteria relations in two cases:

1. Indifference relation: The indifference relation could be seen as the opposite of a significant difference between two alternatives. When the differences are so small that they do not matter, two alternatives can be similar enough that the decision-maker will be indifferent between the two. But indifference does not mean equality and it would be a mistake to assume that it is transitive \( \text{Jacquet-Lagrèse and Siskos, 1982} \).

We can illustrate this property with the following example (inspired from \( \text{Luce, 1956} \)): Consider a series of 20 tea cups that all contain an equal amount of tea but a different amount of sugar. Tea cup 1 contain slightly less sugar than tea cup 2, tea cup 2 contains slightly less sugar than tea cup 3, and so on. If a person were to taste the different cups from 1 to 20, she would hardly be able to distinguish the cups as the difference between two successive cups is too small. However, when comparing cup 1 and cup 20, the cumulative difference would be so important that the person would be able to immediately tell which one has more sugar in it.

2. Preference relation: Even though it might be difficult to believe, preference relations can also be intransitive in some cases. These cases are not so uncommon that they would have to be considered exceptions as explained by Gardner \( \text{Gardner, 1974} \).

Once again, we can illustrate such a case with an example: Let us consider a soccer/football championship between three teams. Three games take place where all pairs of teams compete. The results are shown in Figure 1.2 For the sake of this
example, we will consider that a team that has won against another is preferred to the losing team. We see that team A beats team B with a score of 3 – 0, team B beats team C with a score of 2 – 1, and team C gets a score of 2 – 1 against team A. In such a case, we see that the relation “has won against” or the preference relationship is not transitive. The same situation can be encountered in many more cases where several criteria are considered and the alternatives that excel in regard to some aspects always have drawbacks.

![Table showing scores of teams against each other]

**Figure 1.2: Intransitivity of the preference relation between three alternatives**

In the next section we will define the main concepts behind MCDA and the different types of problems it can help us solve.

### 1.2 Main Concepts

In this section, we will cover the definitions of the most commonly used terms in MCDA and explain some of the notions on which MCDA methods are based. We will see that the varying characteristics of these methods can help us construct a typology. That in turn will help us adequately choose the methods to be applied to a specific problem.

#### 1.2.1 Definitions

##### 1.2.1.1 Action/Alternative

In MCDA, actions will represent the elements in a decision problem that need to be evaluated (e.g. options, projects, candidates, potential decisions ...). In MCDA the word action does not refer to an activity but to any type of possible solution or outcome [Laaribi, 2000]. In spatial problems, we often prefer the term “alternative” that better fits the spatial entities to be evaluated when they are exclusive. Indeed, in most cases the possible solutions to be compared can be points, lines, areas, volumes, or combinations of them that represent the different geographical entities (locations, routes, parcels...).
Roy also defines the notion of potential action as a real or fictional action that is judged feasible for the problem [Roy, 1985]. The construction of the set of all potential actions $A$ will depend on the types of actions it contains: either the actions are all exclusive and the set is globalized or the actions can be put together to form solutions and the set is considered fragmented.

Vincke further explains how the set of actions can be defined before continuing with the decision process [Vincke, 1992]:

- In extension (i.e. by enumerating all its elements) when it is finite and small enough for this task. For example, it could be a list of locations to be evaluated for a given activity.

- By comprehension (i.e. by describing the property that characterizes it or defining mathematical constraints) when it is infinite or finite but too big to be enumerated. For example, we could delimit a territory and try to find the areas in it that are suitable for a given purpose.

### 1.2.1.2 Criterion

In MCDA, the term criterion can have several meanings across the literature. Most of the time each criterion in a problem is a function that is defined on the set of alternatives $A$ (i.e. takes an alternative as parameter) and takes its value in a totally ordered set (i.e. associates to each alternative the preference or evaluation that the decision maker gives it).

Some authors such as Nijkamp use the term criterion when referring to discrete problems and the term objective when referring to continuous problems [Nijkamp et al., 1990]. Others have also pointed out the difference between attributes and criteria: attributes are the untouched characteristics of the alternatives (e.g. price, size, type of soil, quality...); as soon as some additional information is added to them (such as the preferences of the decision maker regarding those characteristics), the attributes become criteria [Pomerol and Barba-Romero, 1993].

In spatial decision problems, it is also interesting to differentiate between explicitly spatial criteria and implicitly spatial ones. Explicitly spatial refers to the characteristics of an alternative that can be obtained directly and are a direct consequence of its spatial properties (location, size, shape, contiguity...). Implicitly spatial criteria however refer to characteristics for which spatial data is required to compute the level of achievement attained by the alternative (cost of disposing of waste, exposure to noise from several sources, visual impact of a construction...). They usually involve the use of spatial attributes such as distance, proximity, slope, elevation...
Sometimes the meaning of a criterion can be mistaken as a constraint. However constraints indicate conditions that have to be met, whereas criteria indicate factors that we wish to improve without necessarily setting a threshold to be attained. In a spatial problem, constraints would be used to delimit the areas to be considered for feasibility reasons, whereas criteria would be used to score them.

1.2.1.3 Problem

All methods in MCDA exist for the sake of solving a specific problem. These can involve choosing an alternative from a given set, assign elements to different categories, compute an evaluation for each entity and rank them...

Among all methods that allow us to model decision problems, we can distinguish two categories based on the types of actions and criteria they use. The first category of problems involves a finite number of alternatives that are explicitly defined. These can be referred to as multi-attribute problems or multi-criteria evaluation problems.

Inside this first category, Roy defines the three main types of decisions as follows [Roy, 1976]:

- Choice problem \((P\alpha)\): It is a selection approach that consists in determining a subset \(A^*\) of all possible alternatives \(A\) which contains the best one(s). One possible example would be the location selection problem that aims a finding one or several locations that suit a specific purpose. The routing problem is another example that uses routes as the feasible alternatives.

- Sorting problem \((P\beta)\): It is a segmentation approach that consists in assigning each alternative of \(A\) to one of several previously defined categories (not necessarily ordered). Land use problem often involve designating the most appropriate activity to assign to certain locations.

- Ranking problem \((P\gamma)\): It is a classifying approach that aims at ordering the alternatives in \(A\) from best to worst. Used essentially to rate alternatives by comparing to other possible solutions, it is suitable to assess the risk associated to certain locations. Under some circumstances it can also be used for selection purposes or classification problems (see below).

With time, some other types of problems that could be solved using MCDA were added to the previous ones:

- Description problem \((P\delta)\): Added later by Roy, this problem simply aims at describing the alternatives and their consequences. The GAIA visualization tool (seen in
Section 1.4.2) is one of the rare methods that focus on describing the alternatives of a problem.

- Nominal or ordinal classification: This is a partition of the sorting problem previously defined in two separate problems. In both cases of classification, categories or classes are defined using reference profiles, limiting profiles, central objects, and so on. When these categories are not linked by any relation, we refer to the decision as nominal classification. However when the categories can be ordered from best to worst or vice-versa, we call the problem ordinal classification [Doumpos and Zopounidis, 2002, Jajuga et al., 2002, Nemery and Lamboray, 2008].

- Clustering: It is a problem that involves grouping alternatives based on their properties. It differs from the sorting problem in that the categories are defined a posteriori.
A second category of problems contains those where the alternatives are implicitly defined (i.e. using constraints) and are either present in very high numbers (infinite or non-countable) or need to be constructed during the decision making process. These are commonly called multi-objective problems or multi-criteria design problems [Evans and Steuer, 1973]. In a spatial context this would require developing new alternatives for possibly addressing a problem. One example would be the partitioning of a territory in segments that respect certain constraints [Tavares-Pereira et al., 2007a].

1.2.1.4 Dominance and Efficiency

The relations of dominance and efficiency can help reduce the number of alternatives considered by eliminating the ones that are "beaten" by others. In problems where the number of alternatives to be evaluated or compared is too important, this can help reduce the workload. In general, we will say that an alternative \( a \) dominates an alternative \( b \) if \( a \) is better or at least as good as \( b \) on all the considered criteria [Vincke, 1992]:

\[
\triangle \iff g_k(a) \geq g_k(b) \quad \forall k \in \{1, 2, ..., q\} \tag{1.4}
\]

with at least one of the inequalities being strict in a problem where all criteria need to be maximized.

We can then define efficiency as the state of an alternative that is not dominated. In other words, an alternative \( a \) from \( A \) is efficient if for any action \( b \in A \) there is at least one criterion \( k \) for which:

\[
g_k(a) > g_k(b) \tag{1.5}
\]

A first step in some approaches can thus consist in eliminating all dominated solutions from \( A \). By doing that we keep only the efficient solutions otherwise known as Pareto optimal solutions. This relationship however is poor as in some cases all the feasible solutions could be Pareto optimal.

1.2.2 Classification of methods

Several types of methods exist based on the combinations of alternatives, criteria, problems to be solved, and various underlying characteristics. This has motivated researchers to try and propose classifications to better select appropriate methods for given problems. Some however warn their readers about the danger of using a multicriteria method to select an appropriate multicriteria method [Guitouni and Martel, 1998], [Tacle and Duckstein, 1992].
The criteria used to distinguish approaches are generally the following [Greene et al., 2011]: Number of decision makers, phases of the decision process to be supported, number of objectives, number of alternatives, existence of constraints, risk tolerance, uncertainty, measurement scales and units, computational resource capacity, forward or backward problem solving, and the experience of the analyst and decision makers. All of these aspects when considered, should allow an analyst to identify a subset of methods that are suitable for a given problem [Guitouni and Martel, 1998, Laaribi et al., 1996].

In the next two sections we present the two most important trains of thought when considering multicriteria aggregation methods. The first is the American school which includes aggregation approaches based on value/utility functions and optimization. The second is the European (mainly French) school which introduced aggregation approaches based on the outranking relationship.

1.3 The American School

The first MCDA approach that appeared was derived from linear programming. It is an approach that relies heavily on precise knowledge and judgements and aims to reach an optimal decision through the use of utility/value functions and multi-objective optimization. The methods that evolved from it try to aggregate the evaluations of each alternative independently in order to obtain a score for them. This approach is also referred to as complete aggregation [Keeney and Raiffa, 1976, Keeney and Raiffa, 1993]. This technique is one of the first to ever be implemented in a geographic context. Its simplicity and relatively low computational cost are the reasons for that. The next subsection presents the weighted sum model, while the next one focuses on the AHP method that can also be used to determine criteria weights. Then we describe TOPSIS, a method that uses a different approach based on reference points.

1.3.1 Weighted Sum

The weighted sum method is the most popular multi-criteria decision method, mainly because of its simplicity. As its name indicates, it is a simple sum of weighted scores:

\[ p_a = \sum_{i=1}^{n} x_{ai}w_i \]  \hspace{1cm} (1.6)

where \( p_a \) is the priority score of action \( a \), \( x_{ai} \) is the score of action \( a \) on criterion \( i \), \( w_i \) is the weight of criterion \( i \), and \( n \) is the number of criteria. We suppose once again that the aim is to maximize all the criteria.
Due to its simple and very easy to explain methodology, the weighted summation is suitable for problems where the decision maker has a greater involvement in the modeling phase. Indeed the approach offers the possibility of visualizing and communicating intermediate and final results which can help support and improve negotiations between stakeholders. Numerous software tools implement this approach and complete it with additional tools such sensitivity analyzes and visualization possibilities [Belton and Stewart, 2002; Janssen, 2001].

The weighted summation however also has some drawbacks [Marler and Arora, 2010; Mareschal and Brans, 2002]. One of them being the several assumptions that are to be made before using it:

- Compensability between criteria: adding criteria evaluations to one another naturally means that profiles that exhibit different performances might end up having similar scores (e.g. a low performance on a given attribute has less impact when there is a high performance on another one).
- Additivity of attributes: This supposes that there is no interaction between the attributes taken into account (i.e. all criteria are independent) which is rarely true in practice.

Applying the weighted summation model also leads to some difficulties to be tackled.

- Loss of information due to normalization: Normalizing by the maximum or minimum of the evaluations, or by the difference between the nadir and zenith (or utopia) points are but a few examples of the possibilities offered in the weighted sum approach [Sunar and Kahraman, 2001]. Each of these possibilities will of course influence the results to a certain extent and have to be used carefully. Some of these approaches can also lead to issues like rank reversal in the results as the processes sometimes depend on the alternatives considered in the decision problem [Belton and Gear, 1985]. Let us note that this problem also occurs in other methods [Mareschal et al., 2008; Saaty and Vargas, 1984b].
- Weight assignment: Even though several methods exist to help assign weights to the different criteria, this step stays a difficult one. Indeed, in the weighted summation, setting up the weights is equivalent to the act of determining substitution coefficients for the compensation that will be taking place (i.e. the decision maker typically has to answer questions of the type “how much of a gain do I need on a given criterion in order to make up for a given loss on another?”).
1.3.2 AHP

The Analytic Hierarchy Process (AHP) by Saaty is a widely used tool for multicriteria decision aid [Saaty, 1980, Saaty, 2002, Saaty, 2005]. It can be found in several areas as presented by [Vaidya and Kumar, 2006] and [Ho, 2008]. Among others, AHP has been used in the fields of engineering [Golden and Wasil, 1986, Muralidhar et al., 1990], manufacturing [Akgunduz et al., 2002, Shang and Sueyoshi, 1995], industry [Ngai, 2003, Radcliffe and Schniederjans, 2003], logistics [Korpela and Lehmusvaara, 1999, Zhou et al., 2000], health-care [Lee and Kwak, 1999] and services [Badri, 1999]. One of the distinctive features of AHP is to build a matrix by asking the decision maker to compare all pairs of actions and criteria using a verbal scale. In the standard version of the method, the normalized principal right-hand eigenvector of this matrix allows to compute the score associated to each action and the weight associated to each criterion (as seen later in Equation 1.10). In this section we will see how AHP can be used to establish a weight distribution for the criteria of a decision problem. A similar approach is used to determine the evaluations of alternatives. This technique is widely used in spatial analysis problems and even exists in some commercial GIS either as a script or a MCDA module (as will be seen in Section 5.3).

As enunciated previously, one of the most important steps of the AHP method is to build a matrix $A$ where each element $A_{ij}$ ($i, j = 1, ..., n$) represents the relative importance of the criterion $i$ over the criterion $j$. To express this relative importance, the decision maker can make use of a verbal scale. The latter is then transformed into a fundamental scale of absolute numbers taking integer values between 1 and 9. A complete explanation of this scale can be found in [Saaty, 2005]. Furthermore, the elements of this matrix need to respect the following consistency property:

$$A_{ij} = 1/A_{ji} \quad \forall i, j$$

(1.7)

When a 9 level scale is used, the matrix $A$ takes its values in the following set $\{1/2, 1/3, ..., 1/9, 1, 2, ..., 9\}$.

In an ideal case, the matrix is consistent, i.e. it naturally verifies the following property:

$$A_{ij} = A_{ik} \times A_{kj} \quad \forall i, j, k$$

(1.8)

In that case, we suppose that the weights matrix $A$ can be written in the following
way:

\[
A = \begin{pmatrix}
\frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\
\frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n}
\end{pmatrix}
\]  \tag{1.9}

where \(w_1, w_2, \ldots, w_n\) are the criteria weights. To find the vector \(w (w_1, w_2, \ldots, w_n)\), we can solve the following equation:

\[
Aw = nw
\]  \tag{1.10}

where the vector \(w\) is the right hand eigenvector of \(A\) (if \(A\) is consistent). We then normalize \(w\) by the sum of its elements to make it unique.

If \(A\) is not consistent, a certain level of inconsistency may still be accepted. Saaty defines the consistency index C.I. to have a measure of the consistency level of a given matrix [Saaty, 1990]:

\[
C.I. = \frac{\lambda_{max} - n}{n - 1}
\]  \tag{1.11}

where \(\lambda_{max}\) is the largest eigenvalue. Saaty also defined a set of reference values R.I., the Random Consistency Index, which depend on the number of alternatives. When C.I. is smaller or equal to 10% of R.I., the level of inconsistency is deemed not acceptable to apply the eigenvector method.

The process to establish the evaluation matrix is the same as this one. The only difference is that instead of comparing criteria based on their importance, the decision maker will be asked to compare the alternatives based on their preferences.

### 1.3.3 TOPSIS

TOPSIS (Technique for Preference by Similarity to the Ideal Solution) was developed by Hwang and Yoon [Hwang and Yoon, 1981; Lai et al., 1994; Yoon, 1980]. The goal is to simultaneously minimize the distance of an action to an ideal action (an action which has the best scores on all criteria) and maximize the distance from an anti-ideal action (an action which has the worst scores on all criteria). Its advantage is the limited subjective inputs needed from the decision-maker. The only subjective inputs are the weights given to the criteria. The method implemented in DECERS [Sullivan et al., 2009] for example, is the classical TOPSIS based on six steps:

1. The evaluations of \(n\) actions with respect to \(m\) criteria are collected in a \(n \times m\) decision matrix \(X = (x_{ai})\), where \(x_{ai}\) is the evaluation of action \(a\) for criterion \(i\).
2. The decision matrix is normalized to give us a matrix $R = (r_{ai})$ where:

$$r_{ai} = \frac{x_{ai}}{\sqrt{\sum_{a=1}^{n} x_{ai}^2}} \quad (1.12)$$

for $a = 1, ..., m$ and $i = 1, ..., n$.

The normalization is necessary to compare criteria measured on different units (e.g. Pounds, years ...).

3. A weighted normalized decision matrix is constructed by multiplying the normalized decision matrix $r_{ai}$ by the criteria weights $w_i$.

$$v_{ai} = w_i \cdot r_{ai} \quad (1.13)$$

4. An ideal (or zenith) and an anti-ideal (or nadir) action are constructed by collecting the best and worst score on each criterion in the normalized decision matrix.

- Ideal action:

$$A^+ = v_1^+, ..., v_m^+ \quad (1.14)$$

where, $v_i^+ = \max_a(v_{ai})$, if $i$ is to be maximized; $\min_a(v_{ai})$, if $i$ is to be minimized

- Anti-ideal action:

$$A^- = v_1^-, ..., v_m^- \quad (1.15)$$

where, $v_i^- = \max_a(v_{ai})$, if $i$ is to be minimized; $\min_a(v_{ai})$, if $i$ is to be maximized

5. Calculate the distance for each action to:

- the ideal action:

$$d_a^+ = \sqrt{\sum_i (v_i^+ - v_{ai})^2} \quad (1.16)$$

where $a = 1, ..., m$

- the anti-ideal action:

$$d_a^- = \sqrt{\sum_i (v_i^- - v_{ai})^2} \quad (1.17)$$

where $a = 1, ..., m$

6. Calculate relative closeness coefficient of each action:

$$C_a = \frac{d_a^-}{d_a^+ + d_a^-} \quad (1.18)$$
The closeness coefficient is between 0 and 1, where 1 is the preferred action. If an action is closer to the ideal than the anti-ideal action, then \( C_a \) approaches 1, whereas if an action is closer to the anti-ideal than to the ideal action, \( C_a \) approaches 0. Opricovic and Tzeng have published a review on the subject with simple example, where an extreme action evaluated on two criteria, is preferred over a superior compromise [Opricovic and Tzeng, 2004]. We cannot generalize this surprising result but certainly the Euclidean distance (\( L_2 \)) used in Equations 1.16 and 1.17 which enlarges high distances [Lai et al., 1994], may lead to different results than methods based on Manhattan distances (\( L_1 \)).

### 1.3.4 Graphical Support Tools

There are several techniques that can help an analyst to adjust the parameters of a method or give some information on the results obtained.

First of all, most of the software that exists for applying MCDA methods has graphical views of the data that has been entered or of the results it generates as bar charts or spider charts. A commonly used tool to represent profiles are the parallel scales (see Figure 1.4). These are constituted of a set of axes that show the evaluations of the alternatives for each criterion. As the evaluations are joined for each alternative, it is easy to identify the profiles and compare them [Inselberg and Dimsdale, 1991].

![Parallel scales](image)

**Figure 1.4: Parallel scales**

Some programs then add some graphical tools that are specific to the methods proposed. For example all programs that implement AHP will give the possibility of entering the hierarchy of criteria and alternatives as a graph.

And last but not least some programs offer graphical tools that allow the user to
conduct a sensitivity analysis on the results obtained [Triantaphyllou and Sanchez, 1997].

Some of the most common tools that do this are:

- **Line weights**: This tool allows the user to select a given criterion and see what would happen to the final result if this weight were to change (while keeping all other weights proportional). The final score of each alternative is represented as a line that is a function of the weight. That means that there is a line for each alternative and that the intersections indicate the weight values where there is a change in the preference relation.

In Figure 1.5 we can see that for the chosen weight value of 25% alternative \( a \) is the best, \( d \) is the worst, and \( b \) and \( c \) have the same score. If the weight were to be increased to 50% then alternative \( c \) would become the best and alternative \( a \) would no longer be preferred to \( b \).

![Figure 1.5: Line weights representation for 4 alternatives](image)

- **Walking weights**: The walking weights are a dynamic version of the line weights which allow the user to experiment manually with changes to the weights and observe the results on a series of graphs. For this reason, this tool cannot be fully illustrated in these pages. This tool is also used with other methodologies like the PROMETHEE methods (see Section 1.4).

Expert Choice (http://expertchoice.com), one of the software packages that implements the AHP and weighted sum approaches offers the possibility of applying several sensitivity analyzes using graphical tools (see Figure 1.6).
Another system that offers sensitivity analysis tools is DECERNS SDSS (http://www.decerns.com). It implements several MCDA methods and allows the user to use line weights or walking weights with most of them (see Figure 1.7).
Some other tools have been developed to visualize parts of a decision problem. One of them focuses on displaying the intransitive data present in preference matrices \cite{Lidouh et al., 2009a}.

1.4 The European School

The second main approach in MCDA was derived from cost-benefit analysis \cite{Keeney, 1972, Keeney and Raiffa, 1976}. It considers the possibility of imprecision in the evaluation criteria and that the optimal decision is not always achievable. This approach gave birth to outranking methods. These follow a different approach than the weighted summation. They were born in Europe and aim at giving more importance to the relations between alternatives \cite{Roy and Vanderpooten, 1996}. The outranking relation will be built through a series of pairwise comparisons of the alternatives. This difference in methodology aims at avoiding some of the drawbacks related to the weighted summation but will introduce some other downside in its stead.

1.4.1 PROMETHEE Methods

The PROMETHEE\(^1\) methodology belongs to the family of the multi-criteria outranking methods \cite{Vincke, 1992}. In it, the actions are first pairwise compared on each criterion according to the decision-maker’s preferences, resulting in scores on each criterion. These scores are then aggregated into a global score, which lead to the PROMETHEE I or PROMETHEE II ranking \cite{Brans, 1982, Brans and Vincke, 1985}. In PROMETHEE I, the resulting ranking is a partial pre-order whereas in PROMETHEE II the resulting ranking is a complete pre-order. Several successful cases have been compiled in \cite{Behzadian et al., 2010}.

The idea behind PROMETHEE is to avoid using the evaluations \(f_k(a)\) of each alternative as is but to compare them pair by pair. The method can be applied to any discrete set of alternatives \(\mathcal{A}\) and starts by computing the differences between each pair of alternatives and for every criterion.

\[
\mathcal{A} = \{a_1, a_2, ..., a_n\} \\
\forall k \in \{1, 2, ..., q\}, \quad \forall a_i, a_j \in \mathcal{A} : \quad d_k(a_i, a_j) = f_k(a_i) - f_k(a_j)
\]

A preference function is then applied to all these differences to convert them into preference degrees that take values between 0 and 1. For each criterion, and for each

\(^1\)Preference Ranking Optimization METHod for Enrichment Evaluations
ordered pair of actions, the decision-maker thus expresses his preference by means of this value. A preference degree of 0 means that the decision-maker does not consider the difference between the two alternatives to be significant and a preference degree of 1 means that the difference is strong enough to make him prefer the better alternative without a doubt.

\[ P_k : \mathbb{R} \to [0, 1] : d_k(a_i, a_j) \mapsto P_k(d_k(a_i, a_j)) \]  

(1.21)

with \( P_k \) being a non-decreasing function that is null for negative values.

An example of such a function can be seen in Figure 1.8. In this example, the preference function requires different parameters such as the indifference threshold \( q_i \) and the preference threshold \( p_i \). If the difference \( d_k(a_i, a_j) \) between the scores of action \( a_i \) and \( a_j \) on criterion \( k \) is higher than \( p_i \), the action \( a_i \) is preferred over \( a_j \). If \( d_k(a_i, a_j) < q_i \), then action \( a_i \) and \( a_j \) are indifferent. Several typical shapes are proposed by Brans et al. for the preference functions like the linear, the step or the Gaussian preference function [Brans and Mareschal, 2005].

![Figure 1.8: Preference function](image)

Finally the method aggregates these preference degrees to compute, for each alternative, either a score per criterion (called the unicriterion net flow \( \phi_k(a_i) \) in Equation 1.23) or a global score (called the net flow \( \phi(a_i) \) in Equation 1.24) which leads to the complete ranking of PROMETHEE II.

\[ P(a_i, a_j) = \sum_{k=1}^{q} \omega_k \cdot P_k(d_k(a_i, a_j)) \]  

(1.22)
\[ \phi_k(a_i) = \frac{1}{n-1} \sum_{a_j \in A} [P_k(a_i, a_j) - P_k(a_j, a_i)] \] (1.23)

\[ \phi(a_i) = \frac{1}{n-1} \sum_{a_j \in A} [P(a_i, a_j) - P(a_j, a_i)] = \sum_{k=1}^{q} \phi_k(a_i) \cdot \omega_k \] (1.24)

The higher the net flows, the better the rank of an action. A deeper discussion on the net flow scores can be found in the literature [Brans and Mareschal, 2005, Mareschal et al., 2008].

It is also possible to define \( \phi^+(a_i) \) and \( \phi^-(a_i) \), respectively the positive and negative net flows:

\[ \phi^+(a_i) = \frac{1}{n-1} \sum_{a_j \in A} [P(a_i, a_j)] \] (1.25)

\[ \phi^-(a_i) = \frac{1}{n-1} \sum_{a_j \in A} [P(a_j, a_i)] \] (1.26)

We then use the intersection of the rankings obtained by these two flows to establish the partial ranking of PROMETHEE I.

PROMETHEE is not the only outranking method. It was indeed preceded by the ELECTRE methods that use a similar methodology [Roy, 1991, Schärig, 1996]. Both share the same advantage which is to try and eliminate part of the compensation problem we encountered with the weighted sum. However, while pairwise comparisons bring this strength, they also have a higher computational cost. This will make applying these approaches difficult on problems with high numbers of alternatives. In spatial problems this means that only discrete problems with limited alternatives can be tackled using outranking methods. Continuous problems like the ones where we try to evaluate each pixel on a map, will have to rely on other methods. In these cases, another possibility is the use of approximated approaches like the one proposed by Eppe et al. [Eppe and De Smet, 2014].

Techniques like PROMETHEE also have a series of graphical tools at their disposal. These can range from the simplest ones like bar charts for the display of rankings and spider charts for the display of profiles, to more complex tools like the walking weights for the exploration of the results obtained. The walking weights were introduced for the PROMETHEE methodology with the PROMCALC software [Brans and Mareschal, 1994] and indeed work well in this case since the last step of the ranking method is a weighted sum of the unicriterion net flows (Equation 1.24). Software packages like DECERNS DSS which gather several methodologies like the weighted sum, AHP, PROMETHEE, TOPSIS show that these same sensitivity analysis tools can be used by several methods alike.
The PROMETHEE methodology however has a visualization tool that was dedicated to it and allows the user to explore the problem’s data before the aggregation step. This tool is referred to as GAIA.

1.4.2 GAIA Visualization Tool

GAIA\textsuperscript{2} is a procedure that produces a two-dimensional view of the multi-criteria problem using a Principal Component Analysis (PCA). To do that we define the matrix $\Phi$ that contains all the unicriterion net flows of the decision problem (as seen in Equation 1.27). We then apply a PCA on the associated variance-covariance matrix $C$.

\begin{align*}
\Phi &= (\phi_k(a_i)) \quad \forall a_i \in \mathcal{A}; k \in \{1, 2, \ldots q\} \quad (1.27) \\
nC &= \Phi^\top \Phi \quad (1.28)
\end{align*}

Once we have selected the two eigenvectors $u$ and $v$ with the highest associated eigenvalues $\lambda_1$ and $\lambda_2$, we can project all the elements of the decision problem on the plane they define: the coordinates of each alternative $\alpha_i$, an axis for each criterion $e_k$, and the weights vector $\omega$ that will represent the objective.

\begin{align*}
\alpha_i : (\phi_1(a_i), \phi_2(a_i), \ldots, \phi_k(a_i), \ldots, \phi_q(a_i)), \quad \forall a_i \in \mathcal{A} \quad (1.29) \\
e_k : (0, 0, \ldots, 1, 0, \ldots, 0) \quad k \in \{1, 2, \ldots, q\} \quad (1.30) \\
\omega : (\omega_1, \omega_2, \ldots, \omega_k, \ldots, \omega_q) \quad (1.31)
\end{align*}

The projections of each of these elements compose what is called the GAIA plane (see Figure 1.9). Finally, it is possible to quantify the percentage of information kept by the plane using the following formula:

$$
\delta = \frac{\lambda_1 + \lambda_2}{\sum_{j=1}^{k} \lambda_j} 
$$

(1.32)

The elements on the GAIA plane each have a very specific meaning concerning the decision to be made:

- **Delta value $\delta$**: These results wouldn’t be complete without an indication on their reliability. The delta value (i.e. the amount of information preserved by the plane)
will give us a confidence level for the results and will have to be indicated alongside them.

- **Positions of the criteria**: The orientation of the axes will indicate which criteria are compatible and which ones are in conflict. Having two compatible criteria means that we can easily find alternatives that excel in both of them simultaneously or on the contrary that some alternatives have bad evaluations on them both. Two criteria will be in conflict when they point towards very different directions. That in turn would mean that it is very difficult to find an action that presents good scores on both criteria.

Furthermore the size of the criteria axes will point out the discriminant criteria within the problem. Indeed, since the plane has been chosen to capture the maximum variation of the actions, the criteria that do not present a high enough variation of the evaluations will likely end up being orthogonal to the plane.

- **Position of the decision stick** $\pi$: In this multivariate view, the indication of an objective is of high importance. It will reflect the importance that the decision maker has given to each criterion. Of course, changing the weights will modify the decision stick and make it point in another direction.

- **Relative positions of the alternatives**: Groups of alternatives on the plane will represent solutions with similar profiles. Indeed, even though there is loss of data due to the projection, two alternatives that are similar will be projected close to
each other on the plane.

- **Positions of the alternatives (according to the criteria):** The location of an alternative on the plane will give us an indication on the type of profile it has. It will point out the strongest and weakest features of a solution. By taking a look at the actions in the direction of each criterion, we can identify the ones that are the best for each factor of the decision.

- **Positions of the alternatives (compared to the decision stick):** When projected on the decision stick, the alternatives take positions similar to the ones from the PROMETHEE ranking. Even though the ranking inferred from a projection could present differences due to loss of data, it still is an interesting use of the tool when more precise information is not available.

Having a GAIA plane such as illustrated in Figure 1.9, the decision maker may easily draw conclusions about his decision problem. However, the results we extract from the GAIA plane are but an approximation of the reality. Because of the loss of data due to the projection on the plane, some of the actions might not be well represented in two dimensions. For example, alternatives that seem close on the plane, might actually be apart from each other but have projections that are close.

There can thus be some distortions between the raw data and their GAIA representation for low $\phi(a_i)$ values. At first, the ranking of the actions in the GAIA plane, obtained by their projections on the decision stick, is not always completely consistent with the PROMETHEE II ranking. Furthermore, the ranking of the actions on a particular criterion is not always coherent with the uni-criterion net flows.

Several variants to the method have been proposed to address these problems [Hayez et al., 2009, Nemery et al., 2011]. One of them, the Weighted GAIA variant is described in Chapter 6 of this thesis.

### 1.5 Conclusion

In this chapter we presented the domain of multicriteria decision aiding, its definitions, concepts, and the problems it tackles. We also presented two of the main approaches used by MCDA methods: the weighted summation and the outranking relation. While each approach has its advantages, none is perfect and they all have problems they cannot solve.

In the next chapter we describe a study that was applied on a location selection problem. The aim of this work was to select the most suitable borough in the Greater London region in order to build a new large casino and obtain efficient social and economic
impacts. We will use this example to compare the results we obtain with three different methods.
Chapter 2

Illustrative Case for MCDA Approaches

2.1 Introduction

The work presented in this chapter is a comparative illustration of several multi-criteria methods used for a location selection analysis. Its aim involves choosing a suitable borough in the region of Greater London to construct a large casino. Currently 17 of the 26 large casinos in London are located in the borough of Westminster which is known to generate the highest revenue in tourist spending. However, in 2007 the Casino Advisory Panel (CAP) recommended the borough of Newham as the most suitable area for a new casino instead of Westminster. By taking two views into consideration (one focused on profitability and the other on social benefits), we evaluate the alternatives using the weighted sum, the TOPSIS and the PROMETHEE methods. Each one of these represents one of the types of approaches that we explored in the previous chapter. The results are compared to the proposals submitted to the CAP for validation. We find that the PROMETHEE and the Weighted Sum Method are more suitable than TOPSIS for solving this problem.

This study was published as a conference paper at OR52 in 2010 [Lidouh et al., 2010]:


An extended version of that study was later published as an article in Tourism Management in 2013 [Ishizaka et al., 2013]:

Site selection is a strategic problem that is regularly encountered in management and marketing studies as testified by the numerous published articles collected in recent surveys [Farahani et al., 2010, Revelle and Eiselt, 2005, ReVelle et al., 2008, Smith et al., 2009]. As in the case of the location of new industrial plants, bank branches, shops, hospitals or schools, the location of a casino is an important decision, because this raises strategic, regional and local considerations [Hannigan, 2007]. However, this topic has been seldom studied in a multi-criteria context. In this chapter, we will review the decision of the Casino Advisory Panel (CAP) in 2007 to recommend Newham as the area in which a large casino should be licensed in Greater London instead of Westminster, which accounts for already 17 out of the 26 existing casinos. This decision has been questioned [McMahon and Lloyd, 2006]. Why did the CAP recommend a permission to build in Newham, which has no previous track record of casinos? Why were other boroughs not considered? In this study, we provide some answer suggestions to these questions using a multi-criteria approach: (1) First, we model the problem based on a literature review of casino location benefits. (2) In a second step, as it is a Multi-Criteria Decision Analysis (MCDA) problem, we apply MCDA methods, which belong to three different families and require minimal subjective input from the decision-makers. Indeed, methods like AHP, when used to compare alternatives require the decision maker’s input at several levels whereas the methods chosen here can be applied by gathering the knowledge ourselves and compensating for the absence of accurate preference information by applying sensitivity analysis techniques. This is an essential feature as we do not have access to the original CAP. For these reasons, we select PROMETHEE (for the outranking family), Weighted Sum Method (for the full aggregation family) and TOPSIS (for the distance based family) to solve this location problem. Our analysis shows that the PROMETHEE and the Weighted Sum Method methods support the CAP decision. TOPSIS, however, results in a different recommendation.

2.2 Problem Description


Westminster’s proposal [Hodgson, 2006] highlights the strong assets of the borough: the high revenue generated by tourism, the high proportion of people in the highest socio-economic categories, the presence of London’s iconic attractions and the high concentration of hotels (40% of the hotels in London are in this borough). The presence of already 17 casinos, which represent 75% of the casinos in London and 14% in the United Kingdom,
ensures a proven location as local inhabitants are accustomed to this kind of premises. The social impact of a new casino in such similar environment would be small. There are even a few areas in the borough far from the commercial area that need some regeneration (see Figure 2.1), although probably less than in other places in London. Based on these arguments, the borough of Westminster bid for two additional large casino licenses.

Newham’s proposal [Heraty, 2006] highlights the fact that the borough is in need of regeneration and lies within the Thames Gateway (identified as a national priority for regeneration). Several conclusions of studies and statistics were included to support this observation. Figure 2.1 shows that Newham lies at the heart of areas needing regeneration and the Council is committed to reduce poverty. It is also London’s best connected borough through road, rail and underground and therefore has significant visitor potential. The report stated that it would be ensured that residents of Newham would benefit from the job opportunities generated by the casino.

![Figure 2.1: Map of the 20% most deprived areas in Greater London (Adapted from the Office for National Statistics licensed under the Open Government Licence v.1.0.). Newham is among those areas, while Westminster has few deprived neighborhoods.](image)

The Casino Advisory Panel (CAP, 2007) took its decision based on two criteria:

- Area in need of regeneration (as measured by employment and other social deprivation data) and which is likely to benefit in those terms from a new casino.

- Area which wants to license a new casino and is likely to find a company willing to open a casino in the area.

The methodology used is not specified in the report but it is our belief that a consensus decision was reached through internal debate rather than using a specific multi-criteria method. The Casino Advisory Panel has recommended Newham for hosting a new casino.
2.3 Model

For our analysis, we used the web application Spatial Decision Support Systems (SDSS) DECERNS. This online tool has already been successful in solving several decisions on land use planning and management [Sullivan et al., 2009]. It incorporates three popular multi-criteria decision methods belonging to different families: PROMETHEE, Weighted Sum Method and TOPSIS. Compared to other techniques, these MCDA methods require a minimal subjective input from the decision-makers, which is ideal in our case as we do not have access to the original CAP advisers. As all these MCDA methods aim to select one action from a set of \( m \) possible actions \( A = a, b, ..., m \) or to rank them on the basis of \( n \) criteria \( C = c_1, c_2, ..., c_n \), the table of score is entered only once in DECERNS and then particular parameters for each method are selected. The next sections will describe in detail each method used in this study.

Our model for the new casino location contains two main branches reflecting the two main objectives of the decision [CAP, 2007]:

- The first objective is to attract a casino company, therefore the number of customers should be maximized. For this purpose, the profile of the gamblers is defined through a literature search and corresponding criteria are selected.
- The second objective aims to regenerate deprived boroughs, therefore social advantages of a casino are modeled.

2.3.1 Profile of a Gambler

This section defines the typical profiles of casino customers. An accurate profile is essential as it allows us to identify the relevant criteria for maximizing the profit of a casino in London. According to Goodman [Goodman, 1995], there are two basic types of normal gamblers: convenience and tourist gamblers.

Most academic studies have focused on the negative impacts of gambling and specially on pathological gambling [Afifi et al., 2010, McBride et al., 2010]. The research on gambling has explored why people become pathological gamblers rather than why and who are the gamblers in the general population. These researches may have reinforced the negative perception that the public have about gambling as a general activity. However, problem gamblers are a small minority according to studies in UK: 0.6% [Orford et al., 2003, Wardle et al., 2007] and 1.4% [McBride et al., 2010]. This observation is also valid in other countries: 1−2% in Australia [Walker and Dickerson, 1996] and 0.5−2% in Canada [Marshall and Wynne, 2004] and 3% in the United States [Kessler et al., 2008]. We will therefore neglect them in this study.
2.3.2 Convenience Gambler

Convenience gamblers are customers living near the casino. Some surveys have identified the general profile of gamblers. In United Kingdom, the British Gambling Prevalence Survey was undertaken to help the Gambling Commission to understand the nature and scale of gambling in Great Britain and then to regulate the commercial gambling. The first survey has been published in 2000 [Sproston et al., 2000] and the second in 2007 [Wardle et al., 2007]. A random sample of 9003 individuals participated in the second survey, which have been interviewed on several types of gambling activities (National lottery draw, bingo, online gambling, casino, etc). According to the survey, 2 million adults (4% of the population) gambled in a casino within the last 12 months. Among casino gamblers there are three times more men than women, and men also spend more: £34/week for a man compared to £3/week for a woman. If we consider only the table games in a casino, young (Table 2.1), single (Table 2.2) and white (Table 2.3) persons represent the majority of the customers. Let us however note that even though other ethnicities are not represented as much their presence in the global population is also small. As such 9% of asian people is actually quite high.

Table 2.1: Casino customers by age [Wardle et al., 2007]

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Gamblers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–24</td>
<td>33 thoroughly readjusted as the 16–18 years are not allowed to enter a casino.</td>
</tr>
<tr>
<td>25–34</td>
<td>32</td>
</tr>
<tr>
<td>35–44</td>
<td>16</td>
</tr>
<tr>
<td>45–54</td>
<td>12</td>
</tr>
<tr>
<td>55–64</td>
<td>4</td>
</tr>
<tr>
<td>65–74</td>
<td>4</td>
</tr>
<tr>
<td>75+</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.2: Casino customers by marital status [Wardle et al., 2007]

<table>
<thead>
<tr>
<th>Marital status</th>
<th>Gamblers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married/living</td>
<td>21</td>
</tr>
<tr>
<td>as married</td>
<td>21</td>
</tr>
<tr>
<td>Separated/divorced</td>
<td>50</td>
</tr>
<tr>
<td>Single</td>
<td>7</td>
</tr>
<tr>
<td>Widowed</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2.3: Casino customers by ethnicity group [Wardle et al., 2007]

<table>
<thead>
<tr>
<th>Ethnicity group</th>
<th>Gamblers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>36</td>
</tr>
<tr>
<td>Black</td>
<td>18</td>
</tr>
<tr>
<td>Asian</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>36</td>
</tr>
</tbody>
</table>

Persons with a higher qualification and a high salary are most likely to visit casinos (Table 2.4). As both criteria have a strong correlation, we will retain only the educational qualification in our gambler profile.

Each Tables 2.1 2.4 can be utilized to estimate the number of customers in each borough. We used all four separately and calculated an average in the hierarchy (Section
Table 2.4: Casino customers by highest educational qualifications [Wardle et al., 2007]

<table>
<thead>
<tr>
<th>Highest educational qualification</th>
<th>Degree or higher</th>
<th>Professional below degree</th>
<th>A-levels</th>
<th>GCSEs/ Other qualification</th>
<th>No qualification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamblers (%)</td>
<td>27</td>
<td>14</td>
<td>27</td>
<td>18</td>
<td>9</td>
</tr>
</tbody>
</table>

2.3.3 Tourist Gambler

Local gamblers are only one portion of the customers. The number of tourist gamblers may be very high, especially in synergistic tourist destinations, characterized by the presence of multiple casinos, thousands of hotels, high quality restaurants, nightclubs and recreational activities such as spas, shopping areas, theaters. The spending levels may change if a new casino is opened in a borough where currently there is no casino. This effect depends if gambling is the primary reason of the travel or a side activity. Recent researches tend to show that tourism gambling is often only a secondary activity. For example, in Las Vegas, almost 100% of the gamblers were tourists in 2009 [Research, 2009a]. However, the first purpose of their visit was for vacation in 40% of the cases and only 13% declared to be tourist’s gambler. In Laughlin [Research, 2009b], 99% of the gamblers were tourists in 2009, but only 20% were tourists’ gamblers, whilst in 48% of the cases the first purpose of their visit was for vacation. Finally, the mix of games has an influence on the type of customers [Zemke and Shoemaker, 2009]. Local gamblers prefer video poker games, whilst tourists’ gamblers play mostly with slots.

In order to quantify the tourism attractiveness of each borough, we will use the tourism spending figure. The amount of money spent by tourists in each borough is easily available through the use of national databases such as will be seen later on.

2.3.4 Social and Urban Benefits

In the United States and Australia, a fundamental justification for casino development has been its potential role as a social, economical and urban development tool [Hannigan, 2007]. Cities and regions see the economic benefit of new investments and tax resources, especially if customers are nonlocal, as in any other economic development activity, i.e. tourism, plants, etc [Barrow et al., 2004] [Burmania, 2010] [Leven et al., 1998]. Casinos are often considered a catalyst for the development of a tourism industry: restaurants, hotels, live entertainment venues [Felsenstein et al., 1999]. Their economic benefits have been acknowledged in several studies [Long, 1996] [Perdue et al., 1999] [Roehl, 1999]. In addition, they could also provide social activities and contribute to the wellbeing of the local community [McMahon and Lloyd, 2006]. However, if a casino development is not coupled...
with a careful community planning, effective implementation and constant evaluation and reassessment, the local residents may be affected from negative effects, as traffic congestion, noise, car-parking problems and reduction of the affordability of houses [McMahon and Lloyd, 2006]. Nevertheless, local residents who perceive personal benefits from having a casino in their community are more likely to support it, which is explained by the social exchange theory [Kang et al., 2008, Lee et al., 2010].

Residents have recognized that casinos have a positive impact on employment [Long, 1996, Roehl, 1999]. The job opportunities are numerous [AGA, 2012, Andersen, 1996, Harrah’s, 2000, Rose, 1998]; for instance:

- Gaming operations: machine technicians, cashiers, dealers, table games supervisors.
- Casino services: security, food and beverages, retail, purchasing, maintenance and facilities specialists.
- Marketing: public relations, market research, advertising professionals.
- Human resources: employee relations, compensation, staffing, training specialists.
- Finance and administration: lawyers, audit, payroll, income control.

It has been estimated that a new casino in Newham could generate about 300 additional jobs [Council, 2006]. As the casino will provide training for employees, its presence is beneficial for inhabitants with unskilled jobs.

On the other side, the common perception is that gambling increases criminal activities. This concern often arises because of the historic connection between gambling and organized crime. However, considerable effort has been done to control organized crime and corruption. Some studies did not find any significant negative changes in unemployment, bankruptcy or crime after casinos opened [Koo et al., 2007]. Other studies have observed an increase of all types of crimes apart from murders in the post-casino construction [Friedman et al., 1989, Grinols and Mustard, 2006, Hakim and Buck, 1989]. However, these studies do not take into account the increase of the population. If the resident population and the average daily number of visitors are combined, the proportional crime rate is reduced [Curran and Scarpitti, 1991] and its net increase is far less than in ski resorts [Park and Stokowski, 2011]. The majority of respondents of surveys do not perceive significant increase in disruptive influences [Stütt et al., 2005]. The introduction of National Lottery, scratch cards, online gambling and bingo served to popularize and legitimize gambling as a more acceptable social activity [McMahon and Lloyd, 2006].
<table>
<thead>
<tr>
<th>Number of customers</th>
<th>Regional &amp; social benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tourism</td>
</tr>
<tr>
<td>Barking &amp; Dagenham</td>
<td>45,1</td>
</tr>
<tr>
<td>Barnet</td>
<td>199,2</td>
</tr>
<tr>
<td>Bexley</td>
<td>103</td>
</tr>
<tr>
<td>Brent</td>
<td>121,8</td>
</tr>
<tr>
<td>Bromley</td>
<td>160,7</td>
</tr>
<tr>
<td>Camden</td>
<td>751,3</td>
</tr>
<tr>
<td>City of London</td>
<td>303,9</td>
</tr>
<tr>
<td>Croydon</td>
<td>204,3</td>
</tr>
<tr>
<td>Ealing</td>
<td>233,1</td>
</tr>
<tr>
<td>Enfield</td>
<td>173,1</td>
</tr>
<tr>
<td>Greenwich</td>
<td>98,5</td>
</tr>
<tr>
<td>Hackney</td>
<td>82,3</td>
</tr>
<tr>
<td>Hammersm. &amp; F.</td>
<td>244,9</td>
</tr>
<tr>
<td>Haringey</td>
<td>92,4</td>
</tr>
<tr>
<td>Harrow</td>
<td>112,2</td>
</tr>
<tr>
<td>Havering</td>
<td>153,5</td>
</tr>
<tr>
<td>Hillingdon</td>
<td>406,1</td>
</tr>
<tr>
<td>Hounslow</td>
<td>135,5</td>
</tr>
<tr>
<td>Islington</td>
<td>189,2</td>
</tr>
<tr>
<td>Kensington &amp; Ch.</td>
<td>1225,7</td>
</tr>
<tr>
<td>Kingston up .Th.</td>
<td>135,8</td>
</tr>
<tr>
<td>Lambeth</td>
<td>212,7</td>
</tr>
<tr>
<td>Lewisham</td>
<td>78,1</td>
</tr>
<tr>
<td>Merton</td>
<td>85,6</td>
</tr>
<tr>
<td>Newham</td>
<td>179,3</td>
</tr>
<tr>
<td>Redbridge</td>
<td>102,1</td>
</tr>
<tr>
<td>Richmond up . Th.</td>
<td>116</td>
</tr>
<tr>
<td>Southwark</td>
<td>217,4</td>
</tr>
<tr>
<td>Sutton</td>
<td>76,6</td>
</tr>
<tr>
<td>Tower Hamlets</td>
<td>251,8</td>
</tr>
<tr>
<td>Waltham Forest</td>
<td>73,3</td>
</tr>
<tr>
<td>Wandsworth</td>
<td>143</td>
</tr>
<tr>
<td>Westminster</td>
<td>3688,7</td>
</tr>
</tbody>
</table>

Table 2.5: Performance matrix used
Another point to consider for the decision is the possibility of synergies with revitalization projects planned by the authorities for the most deprived areas of Greater London [Authority, 2009]. This regeneration technique has been used in several other resorts [McMahon and Lloyd, 2006]. The selected borough could thereby benefit from improved transports, improved health and security systems and the presence of additional attractions in the borough, etc.

All the arguments mentioned above will be diluted if a new casino is constructed in a neighborhood of existing ones. For a maximal impact, a borough with no casino already established is preferred.

### 2.3.5 Criteria Hierarchy

Based on our literature review of the main criteria for casino location (Sections 2.3.1 and 2.3.4), we have constructed a hierarchy of criteria in Fig. 2.2. The next paragraphs explain how the data of Table 2.5 have been collected.

![Figure 2.2: Hierarchy of the criteria](image)

Each of these criteria will be calculated as follows:

- **Tourism spending**: As the spending data of tourists in casinos is not publicly available, we have assumed that it is proportional with the Local Area Tourism Impact (LATI) model. The LATI model has been developed by the London Development Agency, in conjunction with Greater London Authority (GLA) Economics. It estimates the overseas and domestic visitors spending in each borough of London. As we do not differentiate between overseas and domestic spending,
both data have been summed up (Table 2.5). Raw data can be found on: [http://data.london.gov.uk/datafiles/art-culture/tourism-spend-borough.xls](http://data.london.gov.uk/datafiles/art-culture/tourism-spend-borough.xls)

- **Age**: The expected number of customers of a borough can be calculated with the Bayes’ theorem:

  \[ P(A|B) = P(B|A).P(A)/P(B) \] (2.1)

  The conditional probability of a person visiting a casino, given his/her age group, is calculated by:

  \[ P(visitcasino|agegroup) = \]
  \[ P(agegroup|visitcasino).P(visitcasino)/P(agegroup) \] (2.2)

  where \( P(agegroup|visitcasino) \) is given in Table 2.1.

  \( P(visitcasino) = 4\% \) [Wardle et al., 2007].

  \( P(agegroup) \): The age distribution of each borough has been found in the London Datastore ([http://data.london.gov.uk/](http://data.london.gov.uk/)). The resident age population has been estimated by the UK office for national statistics by interval of 5 years.

- **Marital status**: The conditional probability of a person visiting a casino, given his/her marital status, is calculated by the Bayes’ theorem 2.1:

  \[ P(visitcasino|maritalstatus) = \]
  \[ P(maritalstatus|visitcasino).P(visitcasino)/P(maritalstatus) \] (2.3)

  where \( P(maritalstatus|visitcasino) \) is given in Table 2.2.

  \( P(maritalstatus) \): The marital status in each borough has been found in the London Datastore. The data have been collected during the census 2001. As the data on the casino customers by marital status are less detailed than the census, we have merged the numbers of persons married and living as married. The number of persons separated but still legally married and the divorced persons have also been merged for the same reason.

- **Ethnicity**: The conditional probability of a person visiting a casino, given his/her ethnicity, is calculated by the Bayes’ theorem 2.1:

  \[ P(visitcasino|ethnicity) = \]
  \[ P(ethnicity|visitcasino).P(visitcasino)/P(ethnicity) \] (2.4)
where $P(ethnicity|visitcasino)$ is given in Table 2.3.

$P(ethnicity)$: The ethnicity in each borough has been found in the London Datastore. The data have been collected during the census 2001. As the data on the casino customers by ethnicity are less detailed than the census, we have merged the numbers into white, black, Asian and other ethnicity groups.

- **Qualifications**: The conditional probability of a person visiting a casino, given his/her qualifications, is calculated by the Bayes’ theorem 2.1

\[
P(visitcasino|qualifications) = \frac{P(qualifications|visitcasino).P(visitcasino)}{P(qualifications)}
\]

where $P(qualifications|visitcasino)$ is given in Table 2.4.

$P(qualifications)$: The qualifications in each borough have been found in the London Datastore. The data have been collected by the UK Office for National Statistics through a survey in 2008.

- **Competitors**: The list of large casinos has been found on the Guide on London Casinos ([http://www.guidetolondoncasinos.com/](http://www.guidetolondoncasinos.com/)). As the number of competitors has to be minimized, we use the formula

\[
Max(Competitors) - x
\]

where $x$ is the data in the particular borough. This function is linear as its aim is merely to invert the scale for these evaluations.

$Max(Competitors)$ is the highest value in all the boroughs, in our case 17.

- **Pay inequalities**: The data have been collected by the UK Office for National Statistics through the Annual Survey of Hours and Earnings in 2007. As no data were available for the city of London, we have assumed that it has an equal hourly pay than the highest value in London (i.e. Kensington and Chelsea). As pay inequality has to be minimized, we use formula 2.6, where £38.00 is the highest hourly pay of top quartile among all boroughs.

- **Regeneration**: Areas to be regenerated are the deprived areas [Johnson, 2009]. Therefore, we use the index of multiple deprivations 2007, which combines a number of indicators, chosen to cover a range of economic, social and housing issues, into a single one. The index has been calculated by the Department for Communities and Local Government. The data have been found in the London Datastore.
• **Potential Employment:** The number of unemployed persons has been found in the London Datastore. We use the data of the Annual Population Survey 2007.

With this step, our model is complete and ready to be analyzed using the three multicriteria methods we have chosen.

### 2.4 Results

In this section, we present the results obtained by the three methods. For PROMETHEE, we have used linear preference functions where the indifference thresholds are all equal to zero and the preference thresholds are equal to the biggest difference between all evaluations of a given criterion. The normalization functions for the weighted sum have been designed to have a similar shape, but instead of the highest difference, we use the maximum evaluation. Figure 2.3 represents a ranking with the PROMETHEE method, when the top criterion *Number of customers* has a weight of 0.6, *Regional or social benefits* weighs 0.4 and all sub-criteria have an equal weight. These weights have been chosen arbitrarily.

This is a particular case. As we do not know the preferences of the stakeholders, we need to consider all the scenarios in a sensitivity analysis. Figure 2.4 is the sensitivity analysis of the weighted sum. Westminster is the preferred borough to implement a casino if the weight given to the *Number of customers* is higher than 0.505. The gap with the runners-up is extremely high. This result is not surprising because Westminster has a very high tourism spending. Indeed, Westminster has a large concentration of London’s historic and prestigious landmarks and visitor attractions, including Buckingham Palace and Westminster Abbey. In a compromise solution between *Number of customers* and *Regional or social benefits*, Newham would be recommended. If the *Regional or social benefits* criterion is the most weighted, then Hackney would be preferred (Table 2.6).

**Table 2.6: Results for the weighted sum method**

<table>
<thead>
<tr>
<th>Weight number of customers</th>
<th>Weight regional or social benefits</th>
<th>Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>From 1 to 0.506</td>
<td>From 0 to 0.495</td>
<td>Westminster</td>
</tr>
<tr>
<td>From 0.506 to 0.370</td>
<td>From 0.495 to 0.627</td>
<td>Newham</td>
</tr>
<tr>
<td>From 0.370 to 0</td>
<td>From 0.628 to 1</td>
<td>Hackney</td>
</tr>
</tbody>
</table>

(Neunham is second)

The results are very similar for the PROMETHEE method (Figure 2.5 and Table 2.7).

However, the results are different with TOPSIS (Figure 2.6 and Table 2.8). Westminster is by far the preferred borough for a large part of the sensitivity analysis. Only when the weight of the criterion *Number of customers* drops below 0.138, the recommended borough becomes Tower Hamlets and then Hackney. For these scenarios, Newham is the
Figure 2.3: PROMETHEE ranking of the boroughs with the weights Number of customers = 0.6 and Regional or social benefits = 0.4

second preferred borough with a very close score. These results are in agreement with [Opricovic and Tzeng, 2004], who observed that an extreme action would be preferred over a superior compromise.

As the ideal point on the criterion Tourism spending is set by Westminster, the distance to the other alternatives is very high and cannot be compensated by most of the scenarios.

PROMETHEE and the weighted sum method have a consensual recommendation, which corresponds to the real taken decisions:

- If the purpose of a casino is to maximize its financial profits, then the best location is Westminster.
- If a compromise between regeneration and social benefits and its financial profits is
Figure 2.4: Sensitivity analysis with the weighted sum method

Table 2.7: Results for the PROMETHEE method

<table>
<thead>
<tr>
<th>Weight number of customers</th>
<th>Weight regional or social benefits</th>
<th>Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>From 1 to 0.504</td>
<td>From 0 to 0.496</td>
<td>Westminster</td>
</tr>
<tr>
<td>From 0.504 to 0.371</td>
<td>From 0.497 to 0.630</td>
<td>Newham</td>
</tr>
<tr>
<td>From 0.371 to 0</td>
<td>From 0.631 to 1</td>
<td>Hackney</td>
</tr>
</tbody>
</table>

(Newham is second)

searched as in the Casino Advisory Panel [CAP, 2007], Newham is the best location. TOPSIS arrives at a different conclusion than the one taken by the CAP. In both configurations, it will recommend Westminster.

2.5 Conclusion

Since the Gambling Act 2005 creates a new framework for the regulation of all forms of gambling activities in United Kingdom, there has been a debate around the location of new casinos. This chapter has considered this complex problem in Greater London through three multi-criteria analysis. Three boroughs stand out: Westminster for its high number of possible customers, Hackney for its needs of regional and social benefits and Newham for a compromise between both. It is probably not surprising that only Westminster and Newham have submitted a proposal when the Gambling Act 2005 permitted the construction of new casinos. Even if Hackney is in need of regeneration, a casino would not be helpful as it would have attracted a low number of customers. Westminster had
Figure 2.5: Sensitivity analysis with PROMETHEE

Table 2.8: Results for TOPSIS

<table>
<thead>
<tr>
<th>Weight number of customers</th>
<th>Weight regional or social benefits</th>
<th>Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>From 1 to 0.135</td>
<td>From 0 to 0.862</td>
<td>Westminster</td>
</tr>
<tr>
<td>From 0.135 to 0.111</td>
<td>From 0.863 to 0.888</td>
<td>Tower Hamlets (Newham is second)</td>
</tr>
<tr>
<td>From 0.111 to 0</td>
<td>From 0.889 to 1</td>
<td>Hackney (Newham is second)</td>
</tr>
</tbody>
</table>

a proven history of successful casinos in its area. However these casinos were built in an era where the unique criterion considered was to maximize its financial profit. Recently, the modern gambling industry has evolved and casinos are now accepted as a driver for regeneration of deprived areas. The Casino Advisory Panel had clearly stated that the regeneration criteria would be used in awarding the licence for a new casino. In 2007, Newham was recommended by the Casino Advisory Panel. This choice is in agreement with the PROMETHEE and the Weighted Sum Method suggestions. TOPSIS suggests a different recommendation, where the extreme action is preferred over the superior compromise. Opricovic et al. have already observed this phenomenon on a simple theoretical example of three actions and two criteria [Opricovic and Tzeng, 2004].

The case study presented here is a real decision problem and must hence be taken very seriously. From the perspective of economists, decision making is almost always about making compromises. Trying to reach a better outcome in one dimension is often at the expense of achieving a worse outcome in another dimension. It is obvious to most
consumers that if one chooses a lower-priced product (superior in the price dimension), it is usually at a lower quality (the other dimension): one gets what one pays for. A good decision-maker will typically have to correctly trade off one dimension against another. In an approach like TOPSIS which favors extremes, we fail in that purpose.

The choice of the method is thus important, but so is it parametrization. When we do not have access to the preference information, or this data is uncertain, MCDA offers tools to compensate for that and reach reliable conclusions. In the next chapters we will therefore consider all the aspects of MCDA and not focus just on the aggregation procedures.
Chapter 3

GIS and Visual Tools for Decision Support

3.1 Introduction

This chapter presents a review of GIS and how they can be improved with the inclusion of MCDA tools, as well as the visual tools that will help us communicate the results we obtain after an analysis. In a first step, by comparing the functionalities of GIS to those we saw in the MCDA domain, we will start to understand how these systems can benefit from an integration with MCDA. We find it important to warn the reader that the following section does not cover all of the spatial analysis techniques. It only focuses on the aspects that will help us in later parts of our work. For further information on the principles of GIS and their applications, we invite the interested reader to examine the extensive literature on the subject [Cowen, 1988, Goodchild, 2000, Heywood et al., 2002, Longley et al., 2010].

In a second step, we will focus on tools for information visualization. As geographical information systems are visualization tools by nature, we will need to acquire some knowledge about the properties of these tools if we are to hope to use them to build and design our very own visual tools that make use of MCDA and GIS functions. Having a better understanding of how the many graphical devices work and interact will help us avoid mistakes that render our work useless for the purpose it was designed for.

The chapter is divided in two sections: Section 3.2 is focused on several aspects of GIS and their functionalities. In it we see the history behind these systems and the limitations they present. Section 3.3 describes some of the many tools for the visual display of information that we will be using in the later chapters.
3.2 Geographical Information Systems

Geographical Information Systems (GIS) could be considered as systems that allow a user to capture, store, manage, analyze, and present spatially referenced data. They are not limited to computer based tools and can include manual procedures. The acronym GIS can also refer to Geospatial Information Studies which is the discipline that makes use of and works along geographical information systems. In this thesis we will use the acronym to designate the systems and tools that offer support for spatial analysis.

Definition 1 (GIS). A geographic information system is a computer-based information system that enables capture, modeling, storage, retrieval, sharing, manipulation, analysis, and presentation of geographically referenced data [Worboys and Duckham, 2004].

To achieve these functions, GIS unify database technology, cartography, and statistical analysis. For the input of data in the system, GIS feature functionalities that involve the acquisition, reformatting, georeferencing, and compiling said data. The management of these data involve the functions usually available in database management systems for the storage and retrieval of information. Indeed most GIS are database oriented and feature the capability of performing an integrated analysis of both spatial and attribute data. Finally GIS provide ways to see the data in the form of maps, tables, graphs, diagrams...

Not all authors agree on what the functionalities of a GIS are or on what type of system it is. Some consider them to be a special case of information systems or decision support systems that have been extended to handle geographical data, while others think of them as systems that were developed with the purpose of manipulating both spatial and non-spatial data. As GIS are used in a plethora of disciplines and domains, this is not surprising [Goodchild, 2000, Longley et al., 2010, Worboys and Duckham, 2004].

3.2.1 Data Models

GIS use two types of data: spatial data and attribute data [Heywood et al., 2002]. The first type describes the locations of spatial entities (e.g. parcels of land, houses, streets, rivers, lakes, municipalities, provinces, countries...). Attribute data refers to the properties of these spatial entities. Also called tabular data, they are the quantitative and/or qualitative information one would find in regular database management systems.

Spatial entities can belong to either one of three groups: points, lines and polygons (or areas). Points are coordinates on a map that can be used to represent simple locations such as peaks, points of interest... Of the three types of spatial entities, points convey the least amount of information as no measurements are possible (i.e. size or volume of the entity). Points can however be used to represent areas at a smaller scale. For example, cities on a map of the world would be represented by points instead of polygons. Points
can also be connected to form lines, which are used to represent linear features such as roads, trails, rivers... Distance measures are possible on lines. Finally, closed lines or two dimensional polygons are used for geographical entities that cover a certain surface such as lakes, buildings... They can also represent the boundaries of various areas such as forests, cities, countries.

Geospatial data is usually divided into two classes: vector data and raster data. These two types of data models refer to the way the data are stored in the system.

The vector format for storing spatial data makes use of vectors, which are line segments defined by their endpoints. A vector of size 0 represents a point. A succession of interconnected vectors represents a line and chains of vectors that are connected back to their starting point can represent the boundaries of an area or polygon. Contrarily to the raster model where the space is explicitly discretized as a grid, the vector model uses a coordinatization of the space to locate the endpoints of each vector. The use of vector data simplifies several operations where there is a need to combine, stack, or scale data from different sources.

Raster data is structured as an array or a grid of cells called pixels. Rasters are
able to represent any of the three spatial entities mentioned previously (see Figure 3.1). As such, points can be represented by a single pixel, while lines are represented by a sequence of connected pixels and areas by a group of contiguous pixels. Rasters can make data handling or systematic operations easier because arrays are commonly supported by programming languages. They are however inefficient in terms of memory space used. Unless the entities represented feature complex shapes, the display of spatial entities using rasters will take more space than the same entities using the vector data model.

3.2.2 Examples of Spatial Analysis

GIS feature a large panel of spatial analysis tools. To demonstrate the several capabilities and functions supported by a GIS we will review a series of examples that illustrate these through different types of applications. These examples are inspired from the works of Worboys and Duckham [Worboys and Duckham, 2004].

**Resources inventory** The inventory of spatially referenced resources is possible using a GIS. Like any other information system, GIS are able to gather data from different sources, relate and combine them. Provided that the separate sources of data are compatible through the use of common spatial coordinate system, the system will be able to compare them. The GIS can therefore be used as a unifying tool to aggregate data from different inputs and manage them.

**Network analysis** Network analysis is one of most iconic functions of GIS. It allows the system to manage entire networks for transportation, communication... and apply a variety of algorithms on them. In doing so, the system is able to answer queries like finding the shortest path to a given destination, calculating an itinerary that links several points... The criteria considered here can be the distance, but also the time it takes to travel, the cost of following certain paths, or the capacity of certain branches in the network in a flow problem. Many of the algorithms used here come directly from the field of operations research like the shortest path problem, the traveling salesman problem, or the critical path problem.

**Distributed data** Similarly to the collection of data from several sources, a GIS can distribute its services on a communications network. Input data can be obtained from several web services, users can consult the generated information from a distance, calculated routes can be sent to mobile terminals, and results of spatial analyses can be put on maps to be printed. As such, GIS are often used when a participatory system is needed to bring together several inputs or make stakeholders work together.
Terrain analysis  Terrain analysis is based on the use of elevation data sets (see Figure 3.2) to derive information like degrees or directions of slope. These models which can be viewed in three dimensions can be used to perform more complex analyzes like watersheds or viewsheds. Watersheds are used in hydrography to predict the flow of water in case of rain or flood. Viewsheds are a case of visibility analysis that calculate all the areas from which a point is visible. These can be used in regional planning studies to position wind turbines out of the sight of the inhabitants.

Figure 3.2: Contour map where each line has a constant elevation

Layer-based analysis  Layer-based analysis is used for queries that involve selection under certain constraints. For example, finding all the hospitals that are in a 5km radius from an accident or finding all the mobile phone antennas that are within 0.5km of a major road and inside the boundaries of a city. In order to perform such an analysis, the GIS will present the user with two of its main functions: buffering and overlay. Buffering refers to the formation of areas that satisfy distance constraints, like the area within 0.5km of a major road. Buffers are usually circular or rectangular around points and corridors of constant width around lines. Figure 3.3 shows examples of buffers for these two cases. Boolean overlay refers to the union, intersection, difference, or other variants of boolean operators between several layers. It allows the GIS to combine several constraints to generate a single map of entities that satisfies them.

Location analysis  The previously mentioned terrain models and layer analyzes can be used to select locations. To complete them, GIS offer other functionalities that
serve to generate some of the layers to be used in those analyzes. One example is the generation of isochrone maps that shows the neighborhood of certain locations based on the distance or time it takes to reach them. Another example is the use of proximal polygons or Voronoi/Tissen diagrams. These separate the terrain in polygons that contain the area closest to a given point. An example of such a result is given in Figure 3.4.

**Figure 3.4:** Voronoi diagram generated for five points

**Spatiotemporal analysis** Fewer works exist regarding the use of three kinds of dimensions to reference data (i.e. space, time, and attribute). Even though current systems are able to collect and store temporal data, analyzing it and using it for predictions presents several hardships. This type of analysis could lead us to the understanding
of temporal evolution of spatial phenomena if handled correctly.

In conclusion, depending on the analytical processing an analysis requires, we could place them in one of three categories: (1) geometric or topological analyzes if the analysis requires boolean or membership operations (e.g. connectivity, overlay...) over spatial entities, (2) field-based analyzes if it involves the variations of attributes over a region (e.g. the elevation, the type of soil, the wind velocity...) which can be discrete or continuous, and (3) network analyzes which involve configurations of connections between nodes (e.g. road networks, public transportation networks, wired communications networks...).

3.2.3 History of GIS

Spatial analysis has always been a part of our culture in many different forms. The earliest documented examples of it can dated back to the 19th Century with works on epidemiology such as those of Snow on cholera outbreak [Snow, 1855]. The 20th Century then saw the possibility of printing maps with geographic locations, but it wasn’t until 1960 that the first computer system implementing a database for the management of spatially referenced data was developed. This first operational GIS was developed in Canada and used for the management of data for Canada Land Inventory. This system was called CGIS (Canada Geographic Information System). It applied several spatial analysis techniques such as scanning, overlay, and measuring features and lasted until the 1990’s but was never commercialized.

GIS as we know them started appearing in the 1980’s when teams of companies combined the experience from CGIS with their own developments to make products to be sold. While being quite young, they have seen several transformations in their ways of handling data, communicating, and even in their users. Initially designed to serve the needs of government agencies, research centers, or very big companies, GIS have recently become affordable due to their standardization which occurred in the late 1990’s. For the interested reader, an in-depth history of GIS can be found in a book by Foresman [Foresman, 1997] or an article by Goodchild [Goodchild, 2000].

However even after being standardized, GIS come in many forms and with many purposes. Indeed, Figure 3.5 shows us some examples of GIS that we can encounter in very different situations. some desktop software solution exist that allow us to manage spatial data. Some can only handle two dimensional maps, but other are also capable of displaying 3D models such as Digital Terrain Models (DTM). GRASS GIS (http://grass.osgeo.org/), gvSIG (http://www.gvsig.org/), Quantum GIS (http://www.qgis.org/), or ESRI ArcGIS (http://www.esri.com/) are some examples of desktop GIS that implement an impressive toolbox of features.
Then there are systems whose purpose is somewhat limited and are destined to be used in everyday life situations. GPS (Global Positioning Systems) can also be considered GIS as they manage network data and are able to analyze it to generate routes towards several destinations. Web services like Google Maps (http://maps.google.com/) and Bing Maps (http://www.bing.com/maps/) are another example of GIS that are dedicated to a specific task and only a limited number of analysis functions for the users.

\section*{3.2.4 Analytical Limitations of GIS}

Compared to the powerful data manipulation capabilities of GIS, their spatial analysis functions can be quite rudimentary. Indeed, most GIS are focused on the cartography aspect and the management of cartographic data rather than their analysis. This is a direct consequence of the needs of the market for such systems which are driven by the management of data rather than analysis or planning.

Up until the end of the 20th Century, the analysis methods in common GIS were limited to boolean operations on the attributes of spatial entities (points, lines, polygons), line intersections, \textit{point-in-polygon} inclusion operators, \textit{overlay} of layers in matrix or vector data formats, \textit{buffer} creations, statistical analyzes on attribute data, interpolation, and network analyzes \cite{Fischer, 2006, Andrienko and Andrienko, 2006, Andrienko et al., 2007}. 

Figure 3.5: Several uses of GIS: (from left to right, top to bottom) two desktop GIS that display a set of layers or a DTM, a Tomtom GPS used for navigation, the Google Maps web service.
These functionalities albeit powerful quickly reach their limits when used in complex situations.

An example of such limitations of GIS’s functionalities is given by Laaribi concerning the overlay function [Laaribi, 2000]:

- Results from the overlay function rapidly become inseparable when the number of factors concerned is higher than four or five;
- Most overlay procedure do not consider the fact that the variables used can be of different importance (i.e. that these could be weighted differently);
- The use of the overlay function involves the specification of a series of parameters like thresholds. These can lead to a loss of data and are sometimes defined arbitrarily.

3.2.5 Contributions of MCDA to GIS

From what we have seen of their functionalities, GIS are truly designed to support decision making. Through their capability to extract information from raw spatial data, they do offer a series of tools to help us make informed decisions. MCDA can push these strong points even further. It should be noted that GIS already offer support for cartographic modeling and map algebra operations. These can be used to build simple MCDA models such as the weighted summation. MCDA however is not limited to just aggregation models. There several support tools that can have a non-negligible impact on the decision process.

Decision problem structuring MCDA offers a wide variety of methods and tools that help structure a decision problem: aggregation methods, sensitivity analysis procedures, techniques for the generation or description of alternatives, elicitation of parameters, taking uncertainty into account, etc. Most spatial studies indeed rely on alternative-focused approaches where the solutions are at the center of the study. MCDA brings other ways of viewing a problem which a more value-based and place value judgements from the decision maker as the fundamental element of the analysis.

Value scaling Several MCDA procedure require the evaluations for the different criteria to be transformed to comparable units. Linear scale transformations are the simplest and most commonly used in GIS implementations. MCDA however has defined other types of transformations such as the value/utility functions and the preference functions which are designed to be a formal representation of a human’s judgement.

Criterion weighting In all spatial decision problems the criteria considered will likely be of different importance to the decision makers. The most popular weighting pro-
Decision rules  Decision rules refer to the most fundamental procedures that generate results (e.g. ordering, selection...). MCDA presents a variety of them, each method or technique being though for a specific type of problem or situation. As we have mentioned in Chapter 1, several classifications of these methods exist and the choice of a suitable technique is a delicate problem in itself. Applying the correct method however guarantees good results and offers a set of compatible tools to document the decision.

Sensitivity analysis  A unique trait of MCDA which does not have an equivalent in GIS-based decision support is the sensitivity analysis step. It determines the robustness of the obtained results based on variations of the method’s parameters. It is commonly applied on the criteria weights and can be done systematically. Few studies have used any form of sensitivity analysis when dealing with a spatial context. The inclusion of such tools in GIS would offer an additional contribution for the support of spatial decision problems.

The interactions and possible synergies between MCDA and GIS will be covered in more detail in the next chapter where we explore the history of MCDA-GIS integration. As we will see then, one of the aspects which is rarely dealt with, is the possibility of visualizing results and data using both GIS and MCDA tools. Since both fields provide representations to help the user understand the problem, we will explore those possibilities. The next section will thus focus on notions of information visualization and how these will help us display information on maps.

3.3 Visual Display of Information

Next to the spatial analysis capabilities, one of the main advantages of working with GIS is the possibility to visually represent geographic data. When removing the ‘geo-’ prefix in geographical information systems, we are faced with an image processing tool with a specific purpose... that purpose being to convey information to the user. As such, we find ourselves interested in the different tools/icons that could be used to encode nominal, ordinal, or quantitative information. Indeed a picture is worth a thousand words, and most of the time representing certain phenomena on a map can reveal the presence of spatial autocorrelation situations or help us understand the reasons behind some behavior [Cliff and Ord, 1981] Geary, 1954. Indeed according to several social studies the behavior
of people is often influenced by their demographic, economic, sociologic and geographic environments [Latour and Le Floc’h, 2001].

Figure 3.6: Influence spaces that lead to similar behaviors [Latour and Le Floc’h, 2001]

It is however important to be careful not to use the wrong tools for these representations. Throughout the years, several researchers have identified relations between the graphical tools we use and the information they allow us to extract. By taking a look at these empirical and cognitive studies we could select the most appropriate tools for specific tasks to be completed by the user [Bertin, 2010, Mackinlay, 1986, Nowell, 1997].

3.3.1 Basics of Information Visualization

Before focusing on the results that research on visual tools has produced we must first identify the different tasks that a user could be brought to complete when in the presence of a graphical representation:

- Visual search tasks: This first category of tasks consists in scanning a big number of icons while searching for one or several that meet certain criteria [Christ, 1984, Wickens, 1992].

- Identification tasks: This second category deals with situations where the user focuses on a given icon or symbol and tries to report semantic information about it [Christ, 1984].

Different types of icons, symbols, or tools can thus be evaluated for these tasks in terms of time, accuracy, and cognitive workload [Nowell, 1997]. In all cases one must consider several aspects for the successful application of these tools (data types, dimensions and item count, data transformations...).
When dealing with tables of values, we need to distinguish quantitative data, ordered data, and nominal or categorical data. Quantitative data refers to numerical values such as measures of size or distances. Ordered data denotes values that do not support arithmetic but still feature a well-defined ordering such as clothing sizes. Nominal or categorical data does not present an implicit ordering. On top of these data tables we can think of different contexts in which the applications or uses vary. Relational data, for example, is used when values are associated to links between nodes. They are often represented using several types of graphs or networks. Finally some data can be inherently spatial, such as field measurements, locations, routes... In these cases the location of the data in visual representations is naturally reserved for the implicit geographical location of the studied elements. This unfortunately limits the panel of usable visual tools to other possibilities.

The types of data are not the only aspect that influences the effectiveness of visual representations. Indeed the number of dimensions is also going to affect them. Graphical techniques that work with few dimensions are often useless when the number of dimensions increases too much. Most of the time in such cases, the data needs to be transformed in order to be represented or one needs to focus only on some parts of it.

Different problems can also occur when the number of elements to be displayed is too great. In these situations the amount of elements can introduce algorithmic problems with computations taking too long for complex representations on interactive displays. Visual clutter might also occur when the number of elements to be compared is too great to be taken in by a person.

### 3.3.1.1 Graphical Elements

There are several ways to display data, but all of them make use of graphical elements we refer to as marks. These can be represented as points, lines, areas, symbols... and have several properties. Bertin is one of the first to have identified a list of marks and their graphical properties in his works on semiotics of graphics [Bertin, 2010]. Figure 3.7 shows an excerpt of his graphical vocabulary with the list of marks and their application to different spatial entities (points, lines, areas).

Later extended by Mackinlay, these works have inspired numerous studies on graphics. However, having been published in the 80’s, they did not cover newer techniques for the visualization and exploration of data. One must therefore rely on the input of newer studies to assess the impact of these marks in the context of networks, diagrams, 3D visualizations, animations, photography... Furthermore these studies did not consider any interaction with the user which is often present in current information systems.

When thinking about how marks can be constructed or decomposed, Mackinlay’s more recent works can be of help [Mackinlay, 1986, Card and Mackinlay, 1997, Card et al., 2009].
Marks can indeed be of different types and convey information through diverse means. The way these marks stimulate cognitive responses can be decomposed in graphical devices or visual channels. Many graphical devices can be used for that purpose including position, color, size, orientation, shape... These devices or channels can be used to convey different dimensions of information to the user or be combined to convey less but more accurate information through redundancy.

Obviously all graphical devices have a different effectiveness depending on the nature of the data used. It is therefore important to adequately select graphical devices for each data type in order to make them distinguishable enough. One of the first studies we can mention regarding that matter is the work done by Mackinlay [Mackinlay, 1986]. Through several works and publications, Mackinlay has established a ranking of graphical devices based on their effectiveness to convey quantitative, ordinal or nominal information (see Figure 3.8).

It is interesting to see that spatial position dominates our visual perception and is the most accurate device for all three types of data. It is therefore unfortunate that in geographical representations this channel is monopolized by the spatial location of entities,
forcing us to settle with other means to convey additional information to the user. When working with geographical maps we are thus left with devices for which the effectiveness changes tremendously depending on the type of data to be displayed.

As for the others, we see that their effectiveness varies greatly depending on the type of data represented. The variants of size (length, area, volume) and orientation (angle, shape) are the most discriminant for quantitative data, but work poorly for other types. The three usual components of color (hue, lightness, and saturation) are the most distinguishable for ordinal data with a preference for lightness or saturation. Finally, only devices with a small number of states such as hue, textures... are adapted for nominal data.

These differences in perception are best shown with examples [Stolte et al., 2008]. Stolte et al. refer to graphical devices as retinal properties and define mappings for their use to best represent certain types of data. Figure 3.9 shows these mappings.

3.3.1.2 Color

According to several studies, color is one of the most effective graphical devices for reducing visual search time [Christ, 1984] [Carter, 1982] [Smallman and Boynton, 1990]. Some studies have tried to compare color to other graphical devices, while others have focussed on the best number of colors to use. In Mackinlay’s ranking, color is second only to position for the display of ordered or nominal data. This however is only true when its properties are understood and the device is used correctly [Munzner, 2009]. Indeed, region size strongly
affects our ability to sense color as small regions make it difficult to perceive light colors and big regions of bright color can blind the user.

There are several ways to separate color into its components. The additive coloring RGB system separates color in red (R), green (G), and blue (B) light components that when added together at different intensities gives the other colors of the spectrum (see Figure 3.10). The subtractive coloring system separates color in yellow (Y), magenta (M), and cyan (C) dyes.

However, for human perception, color can more easily be separated in terms of three independent graphical devices: hue, saturation and brightness. In the HSB system, hue refers to the actual color that can be named, like the colors of the rainbow. It is quite useful for representing nominal categories. The other two components are more suited for ordinal data because they present an implicit perceptual ordering. As for quantitative data, it should be represented using a colormap, a range of color values. For the best effects, the number of hues in the color map should be limited. Using a rainbow colormap has several defects: it uses hue to represent ordering and is not perceptually isolarinear.
The use of the HSB system is also encouraged by the fact that the conversion from other systems is very easy as is shown by Figure 3.11.

Figure 3.11: Conversion from the RGB to the HSB system

3.3.1.3 Glyphs

Glyphs are symbols that combine several graphical devices and possess an internal structure where each sub-region uses a different device to encode data. Figure 3.12 shows several examples of glyphs from different sources. Some of them resemble statistical graphs while others are more complex. We can even see the face glyphs from Chernoff [Chernoff, 1973]. Several studies exist to show the adequacy of these glyphs for certain purposes [Ward, 2002, Ward, 2008].

The conclusions that we can draw from these studies is that we should aim for glyphs that are simple or are easily understood. Glyphs that require too much cognitive work on the part of the user should be avoided as well as complex glyphs that take too much space. The use of graphical devices that make use of our emotional response can be very effective: we can think about the use of the color red for 'bad' and green for 'good', or some parts of the Chernoff face glyphs like the eyebrow orientation or the size of the mouth to which we are more attuned in comparison to the size of the nose or shape of the face.

3.4 Conclusion

In this chapter, we covered several of the notions that will be used to design the tools in the later parts of this thesis. We saw that geographical information systems are quite valuable as decision support systems. They are indeed quite complete for the management
of spatial data and its automatic representation on geographical maps. However as they still lack some analytical capabilities, they could benefit from an integration with MCDA methods or support tools.

We also had an overview of several information visualization techniques that will help us generate maps in the last chapters of this thesis.
Chapter 4

MCDA & GIS: Recent Developments

4.1 Introduction

The aim of this chapter will be to present some of the previous works that have given birth to this research field. We will then focus on the main advances that have been made during the last years. And finally, we will try to identify the path that lies before us and the objectives that still have to be reached.

Examples of MCDA applied to spatial problems can be traced back to the very beginnings of MCDA. Indeed, we can think of an early study made by Roy and Bertier involving a highway project in the west region of Paris (described in [Bertier and de Montgolfier, 1978]). It is not surprising to see that spatial problems have been treated using MCDA so early since these issues usually tackled existing situations: situations that involve several stakeholders and many aspects to be taken into account. Not only were the direct costs and results important, but the consequences on inhabitants or the environment had to be considered too.

However, those problems were usually formulated the way typical MCDA problems were written. The spatial aspect was in fact reduced to mere quantitative or qualitative criteria. At that time, the spatial properties and relations between the alternatives could hardly be taken into account because GIS as we know them did not yet exist.

4.2 State of the Art

As the subject of integrating MCDA methods in GIS is rather young and most publications and works essentially rely on applications instead of conceptual work, this section will merely focus on the most significant advances in the field. For an in depth review of the
publications the readers can refer to the works done recently by Chakhar or Malczewski [Chakhar, 2006, Malczewski, 2006].

All of the publications focusing on specific applications can give precious guidance as to the difficulties encountered and the ways to avoid them. Unfortunately, the many years that have passed have made it pointless to rely on previous works to ease the development process of new projects. Indeed out of the multitude of solutions that have been coded to apply integrated MCDA-GIS approaches to problems, hardly any have survived. In most cases, once the problem had been solved the developments just ceased, letting any programs or modules to get outdated and no longer work with newer systems. In other cases, the developments were never made public, be it for commercial reasons or just unable to share the work done. Many publications thus only focused on presenting results without explaining in detail how these were obtained.

As such, this section will not describe all the works done with clear applications as targets. Instead we will focus on the few works that have tried to gain a better understanding of the integration itself and the hardships it encounters as it is applied.

4.2.1 The Early 1990’s: First MCDA and GIS Applications

The earliest works that rely on both GIS and MCDA appeared as early as 1988 but were still quite scarce [Diamond and Wright, 1988]. Several of these works studied land-use management or land suitability [Carver, 1991, Langevin et al., 1991]. Carver in particular successfully implemented three multicriteria techniques (ideal point analysis, hierarchical optimization, and concordance-discordance analysis) using the programming language FORTRAN 77 and then linked the methods to the Arc Info GIS software to execute them on his problem’s data. The application of three methods allowed him to obtain robust results.

One of the very first book chapters on the subject was written by Janssen and Rietveld [Janssen and Rietveld, 1990]. It covered a land-use management problem for agriculture purposes in the Netherlands. They developed a multicriteria procedure that they applied to data from the National Physical Planning Agency of the Netherlands using the Arc Info GIS software.

This sudden appearance of GIS and multicriteria themed publications is linked to the massive appearance of user-oriented GIS. During this period, the importance of combining multicriteria decision support tools and the newly available geographical information systems started getting recognized. The analyses however did not involve any integration step as they were treated as spatially referenced problems like several other studies done earlier [Roy, 1985, Scharlig, 1985]. In these studies, spatial analysis and multicriteria evaluation were done separately using separate tools. An example of such a study can
be found in Scharlig’s book [Schärlig, 1985]. It involves the choice of a course for a new highway among seven possibilities. The technical study that led to this list of seven alternatives was of course obtained through spatial analyses carried out by specialists of the road network. Their task also involved measuring the advantages and inconveniences of each option. Once this starting set was established, the selection among the seven viable routes was then done though decision aid methods.

In several recent studies, similar approaches are still applied. However, with the extension of GIS to integrate more and more analysis tools, analysts started to wonder if integrated tools would not make the process easier. One of the most significant achievements that started the next trend of works was the implementation of the IDRISI GIS software. Eastman’s input on this was to make the addition of an analytic decision support tool to version 4.1 of the GIS [Eastman et al., 1993]. That tool contained multicriteria evaluation modules such as the weighted summation and the weights elicitation method by Saaty [Saaty, 1980].

4.2.2 The Late 1990’s and Early 2000’s: First Models and Taxonomy

In the late 1990’s, the number of publications involving MCDA and GIS has started showing signs of an accelerated growth. A survey done by Malczewski on the literature from 1990 till 2004 has identified 319 articles covering both fields [Malczewski, 2006]. It revealed that 79% of the articles considered have been published during the last five years only (see Figure 4.1). It also showed the increasing number of application fields in which these tools were being applied.

In a more recent work, Greene et al. published a similar count that shows that this tendency continued the same way until 2009 [Greene et al., 2011]. The associated graph is displayed in Figure 4.2.

This need for a higher level of integration led several researchers to consider the structuring of MCDA-GIS problems [Jankowski, 1995, Malczewski, 1999, Laaribi, 2000, Chakhar and Martel, 2003, Malczewski, 2006].

4.2.2.1 Classification of MCDA-GIS Problems

Malczewski proposed to classify MCDA-GIS problems using two viewpoints: geo-information components and MCDA components [Malczewski, 1999, Malczewski, 2006]. When considering geo-information components, problems can be distinguished based on three dichotomies:

- **Raster-based or vector-based data model** This denotes the type of data used to represent the geographic entities. This distinction is important as the methods used on
Figure 4.1: Total number of GIS-MCDM articles per year for the period 1990-2004 [Malczewski, 2006]

each type will differ greatly (as will be seen below). Raster-based data models will require the use of continuous methods, whereas vector-based data models will make use of discrete methods.

Explicitly or implicitly spatial criteria The nature of the criteria can sometimes be identified as being explicitly spatial when they are spatial characteristics of the alternatives such as shape, size, orientation, compactness, contiguity... On the other hand some criteria can be considered as implicitly spatial when they involve spatial attributes such as distance, accessibility, slope, proximity, elevation... for the computation of the alternatives’ evaluations.

Explicitly or implicitly spatial alternatives This distinction is used to differentiate cases where the spatial component of an alternative is explicitly or implicitly defined. Explicitly spatial alternatives appear in decision problems such as location selection [Ishizaka et al., 2013], location association, land use suitability, partitioning... Typically, in these problems the alternatives are spatial entities or combinations of those. In contrast, implicitly spatial alternatives refer to decision problems where the solutions implemented can have a series of positive or negative influences on certain locations. For this type of problem we could think of a series of measures to reduce
the risk of erosion [Macary et al., 2010]. The solution implemented can indeed have a positive impact locally but have other consequences in other parts of the watershed.

For the MCDA components, Malczewski once again distinguished three dichotomies to categorize GIS-MCDA decision problems:

**Multiattribute and multiobjective decision analysis** In multiattribute problems, we assume that we deal with a predetermined and limited set of alternatives which will go through a selection process. These methods can be referred to as discrete methods. On the other side, multiobjective problems are usually continuous in that we deal with a set of feasible solutions usually defined implicitly through the use of constraints. Malczewski goes further in dividing MCDA approaches based on the types of decision rules they use [Hwang and Yoon, 1981]. As such, multiattribute approaches are separated into the following five categories: (1) Weighted summation/Boolean overlay, (2) Ideal/reference point (TOPSIS, MOLA), (3) Analytical Hierarchy Process (AHP), (4) outranking methods (ELECTRE, PROMETHEE), (5) Other. Multiobjective approaches are separated in four categories: (1) Multiobjectives programming algorithms (linear-integer programming), (2) Heuristic search/evolutionary/genetic algorithms, (3) Goal programming/reference point algorithms, (4) Other. In his survey Malczewski also indicated the amount of articles belonging to each category which indicates their popularity [Malczewski, 2006]. Those results
are represented in Table 4.1. In the case of outranking methods, we see that only 4.7% of all papers use them for their studies. An even lower number of papers make use of the PROMETHEE methods, that is why we will propose an extension of PROMETHEE-GAIA to the ranking of spatial alternatives and their visualization.

**Individual and group decision making** Most problems involve only one person or goal-preference structure as the decision maker. However others may have several individuals or interest groups represented by different goal-preference structures (i.e. different preference parameters or weights associated to the different criteria of the problem). In the later case the problem is referred to as a group decision making and will rely on participatory decision making approaches [Jankowski and Nyerges, 2003].

**Decisions under certainty and uncertainty** This separation depends on the amount of information available to the decision maker or the analyst. When we have perfect knowledge about the problem, we are working under condition of certainty and the approach is referred to as deterministic. When some aspects of the problem are unknowable or very difficult to predict we refer to it as working under condition of uncertainty. There are then two types of approaches that can be used based on the types of uncertainty: (1) Probabilistic (or stochastic) when we work with limited information and (2) fuzzy decision making when our information presents fuzziness (i.e. imprecision).

<table>
<thead>
<tr>
<th>Combination Rules</th>
<th># of articles</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MADA Weighted summation/Boolean overlay</td>
<td>143</td>
<td>39.3</td>
</tr>
<tr>
<td>Ideal/reference point (TOPSIS, MOLA)</td>
<td>35</td>
<td>9.6</td>
</tr>
<tr>
<td>Analytical Hierarchy Process (AHP)</td>
<td>34</td>
<td>9.4</td>
</tr>
<tr>
<td>Outranking methods (ELECTRE, PROMETHEE)</td>
<td>17</td>
<td>4.7</td>
</tr>
<tr>
<td>Other</td>
<td>30</td>
<td>8.3</td>
</tr>
<tr>
<td>MODA Multi-objectives programming algorithms</td>
<td>57</td>
<td>15.7</td>
</tr>
<tr>
<td>Heuristic search/evolutionary/genetic algorithms</td>
<td>29</td>
<td>8.0</td>
</tr>
<tr>
<td>Goal programming/reference point algorithms</td>
<td>9</td>
<td>2.5</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>2.5</td>
</tr>
<tr>
<td>Total</td>
<td>363</td>
<td>100</td>
</tr>
</tbody>
</table>

Several other classifications of MCDA methods have been attempted by other researchers [Laaribi et al., 1996] [Guitouni and Martel, 1998] [Greene et al., 2011]. Moffett et al., for example, proposed a decision making process that is based on the categories...
also used by Malczewski [Moffett and Sarkar, 2006]. The result is a flowchart that is represented in Figure 4.3.

![Flowchart of decision making scenarios](image)

Figure 4.3: Flowchart of decision making scenarios [Moffett and Sarkar, 2006]

Others like Laaribi or Guitouni also tackled this problem with the aim of helping with the choice of a method to apply on a problem [Guitouni and Martel, 1998, Laaribi, 2000].

### 4.2.2.2 Integration Levels

Several researchers have also attempted to classify integration works in terms of the extent of integration [Malczewski, 2006, Chakhar, 2006]:

**No integration** This refers to works where both MCDA and GIS were used independently. Transfers of data or results from one tool to the other were thus done by the analysts or had very little influence on one another.

**Loose coupling** Loose coupling or weak coupling involves two systems accessing data through shared exchange files. This means that the output data from one system is formatted such that the other system can use it as input.

**Tight coupling** In tight coupling not do the systems share a common data source, they also have a common user interface.

**Full integration** A more complete integration can be attained by making it possible to implement user-specified methods or routines using generic programming languages.
These levels of integration, as seen by several researchers, are often the same and seem to translate an implicit objective of reaching the most complete integration. However if we look at all the MCDA-GIS works that focus on practical cases, we see that most of them do not make use of very high integration levels. Even that some integrated solutions exist, are they hardly used in analyzes. This hints at the fact that maybe technical integration is not the most important aspect of the conjoint use of MCDA and GIS to solve problems. This does not mean that integration is not necessary, but rather that the aim is slightly different than we expected. We will cover this aspect in detail in the next chapter.

4.2.2.3 Limitations of Integration Works

In his contributions, Chakhar identified a list of difficulties that have hindered a complete integration up to this point and still do [Chakhar and Martel, 2003, Chakhar, 2006].

**Loose or tight coupling** when considering the works that have been done we see that very few have considered a complete integration. Malczewski’s survey [Malczewski, 2006] indeed shows us that there are only 11% of the articles considered use the full integration compared to 33.2% for loose coupling and 29.8% for tight coupling (the rest of the articles making no mention of the type of implementation used). This situation is due to several advantages of the less than complete integrations: they are easier to use and easier to implement from the start; software already exists for some MCDA methodologies; the complete integration had not yet convinced all researchers that the considerable effort it needed would be equally rewarded.

**Integration of a single or limited number of methods** In most practical studies, the researchers only went to the trouble of implementing few methods that were deemed useful for the case. Other used only methods that were already available in the systems used. It is however common knowledge that each and every method has its advantages and weaknesses [Carver, 1991, Heywood et al., 1995, Guitouni and Martel, 1998; Chakhar and Martel, 2003]. They all have limited sets of problems for which they are suitable. As we witnessed in Chapter 2, the three methods applied to the same location selection problem did not give us compatible results in all cases.

**Choice of a MCDA method** The choice of MCDA method for a given decision problem is often arbitrary. Analysts will usually select a method they are familiar with or which is easily available. However the choice of a method is a crucial part of the decision support process and has been documented by several authors [Hobbs and Meier, 1994, Guitouni and Martel, 1998, Guitouni et al., 1999, Laaribi, 2000, Opricovic and Tzeng, 2007]. In most works, the authors propose the use of a classification

Integration of the weighted sum and boolean overlay techniques Without any surprise, we discover through Malczewski’s survey [Malczewski, 2006] that most articles about MCDA-GIS rely on the weighted summation (39.3% according to Table 4.1). Even though outranking methods are often found to be more reliable on the field [Malczewski, 1999], these are harder to implement due to their relative complexity when compared to the weighted summation and they are subject to computing limitations. This can be quite bothersome when we realize that GIS implementations of the weighted summation are often used without full understanding of the assumptions underlying the approach. The method is also often used without full insight into the meaning of the weights assigned to attribute maps and the procedures for deriving commensurate attribute maps. Several articles report and discuss this aspect in the literature [Heywood et al., 1995, Hobbs, 1980, Lai and Hopkins, 1989, Malczewski, 2000].

In-depth knowledge of GIS and MCDA Regardless of the number of works on MCDA-GIS integration, their use stays somewhat limited to practitioners of these fields or for academic research. This is due to the fact that the tools require a good understanding of both GIS and MCDA. Furthermore an implementation that worked for a given problem might not be directly usable on another due to the specifications of each decision problem.

4.2.2.4 Necessity to Integrate

Obviously one would not have started such difficult work without a purpose. In the year 2000, a book in French by Laaribi on GIS and Multicriteria Analysis was published [Laaribi, 2000]. This book is one of the very first to consider the integration problem on a conceptual level and establish some reasons behind the necessity of integrating MCDA and GIS. He also proposed several approaches on how to counter the several difficulties including the ever so delicate choice of a multicriteria aggregation method to apply on a problem.

1. GIS are valuable decision support systems: GIS have become mature as data management systems and are able to handle more and more spatial decision problems.
Their limitations in terms of analytical functionalities can be taken care of by adding MCDA functions and tools to them. Spatial decision problems now often present several criteria to be taken into account, which further justifies an integration.

2. GIS can help in defining parts of the problem: GIS with their spatial functionalities can help define the feasible solutions in a problem, to establish the constraints and criteria and even evaluate the alternatives automatically for certain spatial aspects.

3. GIS can be used to identify the feasible solutions: once the feasibility constraints have been set, a GIS can generate the set of feasible solutions. Its overlay functions can apply several constraints and delimit the solutions that verify them. This reduced set of solutions could then be evaluated using a MCDA method.

4. MCDA brings tools to solve different types of problems: while GIS have functionalities that help with the design step of the problem, MCDA can help with the selection step. Both tools thus present a complementarity when considered in a greater decision support process.

4.2.3 The Late 2000’s: A Look at the Future

4.2.3.1 New Methodological Approaches

The first works that associated MCDA and GIS have covered a high number of application domains and help solve many types of decision or evaluation problems [Malczewski, 2006]:

- Land suitability
- Plan or scenario evaluation
- Site search or location
- Resources allocation
- Transportation, vehicle routing or scheduling
- Impact assessment
- Location-allocation

The late 2000’s however have seen the birth of new types of decision problems. Among the new types of problems that have been tackled recently we can mention the approach proposed by Tavares-Pereira et al. to partition a territory while taking several criteria into account [Tavares-Pereira et al., 2007a]. This approach used a genetic algorithm to group several small areas in sectors according to several criteria and constraints such
as contiguity, integrity, compactness, homogenization, similarity... It allows us to generate several partitions which can then be subject to an evaluation or selection process.

Other works involved the comparison of maps according to several criteria in order to identify the positive or negative effects of certain decisions. Works such as the thesis of Metchebon Takougang have but laid the foundations of what could be achieved if this were to be studied further [Metchebon Takougang, 2010].

Works that involve evaluation through time are also very scarce. However when considering what is attempted in the domain of information visualization [Andrienko et al., 2007], it is but a matter of time before spatio-temporal decision support methods become clearly defined.

4.2.3.2 Elementary MCDA Functions

One of the works by Chakhar to make the MCDA-GIS integration progress further is to decompose MCDA decision support processes into their elementary functions [Chakhar, 2006]. The list of functions is given in Table 4.2. Figure 4.4 gives us a view of these same functions in the two approaches that are multiattribute (discrete) and multiobjective (continuous) decision making. This work is based on the fact that all methodologies follow a similar decision support process beginning with the collection of data, the design of the decision model, the application of the decision rules and finally the recommendation.

<table>
<thead>
<tr>
<th>#</th>
<th>MCE function</th>
<th>Used in</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Definition / generation of actions</td>
<td>Discrete</td>
</tr>
<tr>
<td>2</td>
<td>Construction of criteria maps</td>
<td>Discrete</td>
</tr>
<tr>
<td>3</td>
<td>Construction of attribute maps</td>
<td>Discrete</td>
</tr>
<tr>
<td>4</td>
<td>Definition of mathematical program</td>
<td>Continuous</td>
</tr>
<tr>
<td>5</td>
<td>Resolution of mathematical program</td>
<td>Continuous</td>
</tr>
<tr>
<td>6</td>
<td>Generation of performance table</td>
<td>Discrete</td>
</tr>
<tr>
<td>7</td>
<td>Quantification</td>
<td>Discrete</td>
</tr>
<tr>
<td>8</td>
<td>Normalization</td>
<td>Discrete</td>
</tr>
<tr>
<td>9</td>
<td>Pre-analysis of dominance</td>
<td>Discrete</td>
</tr>
<tr>
<td>10</td>
<td>Generation of feasible solutions</td>
<td>Discrete / Continuous</td>
</tr>
<tr>
<td>11</td>
<td>Elicitation of preferences</td>
<td>Discrete / Continuous</td>
</tr>
<tr>
<td>12</td>
<td>Weighting of evaluation criteria</td>
<td>Discrete / Continuous</td>
</tr>
<tr>
<td>13</td>
<td>Sensitivity / robustness analysis</td>
<td>Discrete / Continuous</td>
</tr>
<tr>
<td>14</td>
<td>Aggregation</td>
<td>Discrete</td>
</tr>
<tr>
<td>15</td>
<td>Proposition</td>
<td>Discrete</td>
</tr>
</tbody>
</table>

If these steps are separated into their most basic components, and implemented into a GIS, they would allow the user to construct his own MCDA routine by using the functions...
Figure 4.4: Flowcharts for the elementary MCDA functions

as building blocks. And indeed almost all MCDA aggregation methods can be decomposed in such a way. Chakhar then developed an implementation of this notion and also defined a language to allow its use. Although this implementation has not seen other uses yet, a similar approach has been implemented in the Decision Deck framework for its diviz initiative [Bigaret et al., 2010]. Diviz is a tool that allows the user to select among several web services to build workflows that execute MCDA methods.

4.3 Conclusion

We can clearly see that there is still a lot of research possibilities that have not yet been fully explored. Just like the MCDA field has not yet reached its maturity yet, the MCDA-GIS integration still has a long way to go. Indeed if we compare these fields to the more spread statistics or even operations research domains, we can see that they are used by a minority of analysts. Statistics and operations research tools can be found in most data analysis softwares or toolboxes and even appear in more popularized spreadsheet softwares like Microsoft Excel or OpenOffice Calc. Of course the place for MCDA and spatial MCDA is not the same. However in the future we might reach a point where these fields become
part of of our daily lives without us even noticing.

In the next chapters we will first explore the reason that drive us to try and integrate these two fields. Then we will present visual tools that we have developed to ease the use of both tools and set a purpose for the integration. We will indeed show through examples and case studies that a correctly applied integration can have several positive contributions for decision support.
Chapter 5

Reasons and Objective for the Integration

5.1 Introduction

This chapter is a summary of several conclusions drawn from discussions with some of the speakers and participants of the MCDA’74 Meeting (October 2011, Yverdon-les-Bains, Switzerland). In it, we explore the reasons for integrating MCDA and GIS and give an overview of the concrete works that have been achieved, the tools that are available and their use.

The reflections presented here are an analysis of the results that have been accomplished and how these position themselves in the general objective that is the MCDA-GIS integration. From a technical point of view to the usability that is expected from integrated systems, we will see what drives this work and the types of results that one can or should expect.

The results of these thoughts were published in a special issue of the International Journal of Multicriteria Decision Making [Lidouh, 2013]:


So why should we integrate? Since the start of the 90’s, research works have been studying the conjoint use of spatial analysis methods and multi-criteria methods [Chakhar and Martel, 2003, Chakhar, 2006, Eastman et al., 1993, Greene et al., 2011, Laaribi, 2000, Malczewski, 1999, Malczewski, 2006]. The reason behind it being that an increasing number of problems presented characteristics from both fields. Indeed, most spatial decision problems apply to existing situations with several stakeholders and several factors that need to be taken into account. Geographical Information Systems (GIS) have always been considered as decision support systems because of their data management, analysis, and visualization.
functionalities. However, on their own, they lack the analytical functionalities to take several factors into account simultaneously and produce multivariate results. Introducing multi-criteria aggregation functions, sensitivity tools, and methodologies [Bouyssou et al., 2006] in general can therefore help improve the analyses that are made of spatial decision problems.

During the last twenty years, the integration attempts made remarkable progress. Programs have been developed to allow the conjoint use of the two disciplines, new methods were designed to take care of specific geographic and multi-criteria questions, some methods have been adapted and used on problems that weren’t initially targeted, and the world of spatial information systems has opened up to multi-criteria analysis as a new way to aggregate sources of data.

We could stop at this, but we wouldn’t be entirely persuaded that there really is a need to integrate these two fields. Certainly, an integrated system is always useful, but there are many ways to solve these types of problems without resorting to an integrated system. Up till now the systems that have been developed can be placed into several technical levels of integration ranging from weak coupling to fully integrated systems [Chakhar and Martel, 2003, Laaribi, 2000]. Surprisingly however, the solutions with the greater success are the ones that rely on weak coupling [Malczewski, 2006].

This makes us wonder about the real usefulness of this integration and there is now a real discussion as to whether it is necessary to continue the work even further than the solutions that already currently exist. For now, we don’t know exactly what this integration is exactly needed for and if it really is interesting to invest several months of work in the development of such a solution. To find an answer to this problematic, it is necessary to focus on what has been accomplished already, on the types of problems that have been studied, and on what could be improved within the process of these analyses and the way they are conducted.

Section 2 of this chapter describes the integration of MCDA and GIS, the types of solutions that can be developed, as well as their advantages and inconveniences. The aim is not give an entire state of the art of the subject but well to raise the different factors that will affect the conjoint usage of MCDA and GIS. The interested reader can refer to Malczewski’s exhaustive review [Malczewski, 2006] or several other works [Chakhar, 2006, Greene et al., 2011, Laaribi, 2000] to get an idea of the current situation and conjoint uses of MCDA and GIS. In Section 3 we explore the technical aspects and difficulties of the integration and illustrate it with examples of existing tools and software. We also see some development options that are available for people wanting to program a custom integrated solution. Finally, Section 4 is a study proposal on the usability of integrated systems, the target audience, the needs that should be met ... This future study should
be the most important contribution one could bring to this domain. It would help better orient the efforts that are being made in order to quickly attain significant results.

5.2 The Integration of MCDA and GIS

In this section, we will cover the characteristics and advantages of the different levels of MCDA-GIS integration. Our aim is to understand when each of these solutions is used and with what goal. Of course the following classification is not crisp as there can be solutions that share similarities with different levels of integration.

5.2.1 Weak Coupling

Loose coupling is the first usage that has been made of GIS and MCDA-capable systems. It consists in using specific programs or tools for the different steps of the analysis. The communication between these programs, which can be complex at times, is handled by the analyst. A common decision process could resemble the following steps:

- defining the alternatives using the GIS’s data management functionalities (e.g. binary or overlay operations),
- evaluating the alternatives by entering their characteristics in an MCDA application,
- taking knowledge of the results and producing a map to display them using a GIS.

The direct advantage of this procedure was its low development cost. However this came at a heavy price as the concrete separation between these steps made it hard to fully comprehend the spatial aspect of the problem. Spatial relationships between alternatives were therefore not considered explicitly during the MCDA evaluation step. Furthermore since the transfers of data were done by human interventions, the risk of errors was high. And in the cases where interfaces were developed to link different systems and transfer data automatically from one another, the works achieved could not be used with other systems because of technical differences. A practical example of weak coupling would be any study that would have been done using two software packages for the individual parts of the process: a GIS for the gathering of the problem’s data and a MCDA program for the evaluation step.

5.2.2 Tight Coupling

In the last ten to fifteen years, the need for automatic communications between systems has given birth to works that coupled MCDA tools and GIS under a single interface. To achieve such a result the methods of one field (usually MCDA) were implemented as a
module, routine or script into the other system which is used as a base. Even though the sources of data were not always the same and the data transfers from one system to another were not entirely transparent to the user, this type of tight coupling has been an important step towards the integration of both fields. By making both tools known to the analysts and easier to use, these systems showed the importance of combining them. A few drawbacks were still present though. Flexibility and interactivity were the main concerns of these systems. Researchers began therefore to search for ways to break multi-criteria decision processes in order to give analysts the same freedom they experience with spatial analysis functions Chakhar and Martel, 2003.

The IDRISI GIS system that will be described later is probably the first attempt at tight coupling of MCDA methods within a GIS Eastman et al., 1993. In this GIS, MCDA methods were implemented as modules that gave access to MCDA analyzes.

5.2.3 Full Integration

When looking at the evolution of Information Technologies, we see that through the Internet, several applications and services have opened themselves to the masses. Sometimes simplified, they were made available as web services to all individuals and thereby formed a new generation of participatory systems. These are characterized by their accessibility but also by the way they interact. Web-services nowadays are standardized so that mash-ups (i.e. hybrid combinations of web applications) have been made possible. Their evolution into linked services is reaching a new stage in the type of experience they offer to the user Pedrinaci and Domingue, 2010. The number of websites that use Google Maps as a base and add services on top (e.g. weather, navigation, traffic, parking...) has recently experienced very strong growth. Of course one could expect the same type of user experience from multi-criteria spatial decision support systems. This however would require a good formalization of all data structures used in these kinds of problems. Standards have started being defined for geographical information, yet there is, apart from the XMCDA encoding standard elaborated by the Decision Deck Consortium Dec, b, hardly anything similar available for multi-criteria techniques.

Ideally a fully integrated MCDA-GIS system should offer multi-criteria functionalities and spatial analysis functionalities and allow the user to access any of those at any given time during an analysis. The user should also be able to interact with the different processes by changing the parameters of the methods and visualize the results or the different spatial elements of the problem directly on a map.
5.2.4 Technical Aspect vs. Operational Aspect

The separation that we have just seen in three levels of integration has been made while keeping the technical aspect in mind as that has been the main concern for these past years. We should however be careful not to forget the operational aspect of such systems. Indeed, the distinction between levels of integration can be made based on the technology used for the system, but it could also be used to measure the usability offered by it.

When we consider these two aspects, we can easily understand why the higher levels of integration suffer from a lack of interest while the lower levels thrive. Indeed, even nowadays, for most novel studies that mix MCDA and spatial analysis, the researchers prefer to resort to separate programs instead of using integrated solutions. It could be due to a lack of available systems or knowledge thereof (most integrated solutions are rarely known to the analysts that need them). However the most plausible explanation could be related to the usability of these systems and the place they can take in an analysis. When considering an integrated system, unless the solution perfectly matches the problem at hand, the user will always try to couple it with separate tools that better fit the tasks to be done.

When an analyst takes on a new problem, its characteristics are rarely explicitly defined. The analyst must therefore proceed by exploring the properties of the decision problem before modeling it. Oftentimes, no system will be able to replace the analyst’s expertise on the matter. Furthermore a fully integrated system that has been designed to follow general decision support procedures will not always have the flexibility to allow small changes in the process to adapt to a very specific problem. We could therefore summarize the situation as follows (‘+’ signs indicate advantages, ‘−’ signs indicate inconveniences):

**Weak coupling:**
Uses separate tools or programs for the different parts of the analysis that are able to export-import data from one another.

− Technically the weakest integration.
− Error-prone.
− Not iteration friendly.

+ Offers the most freedom to the user (i.e. analyst) and relies on his/her experience.

**Full integration:**
Uses a single system that proposes several preselected functionalities to deal with most commonly encountered problems.

+ Better communication of data.
+ Makes application of processes easier.
+ Allows interactions during the analysis.

− Forces the user to work with predetermined methods or tools.
Now that we have identified the strong and weak points of the two different approaches, we can now take interest in the existing solutions for such implementations. Furthermore, by taking interest in the way that integrated systems are used we should be able to identify the shortcomings on which to work to fulfill the needs of the users.

5.3 Technical Aspects of the Integration

In the following section, we will illustrate these technical levels of integration while trying to identify the associated operational level of integration. These examples will show us that unfortunately, even though impressive technical advances have been made, the operational aspect has up till now been neglected.

5.3.1 Existing MCDA-GIS Software

Since the start of the integration works, several options of MCDA-GIS systems have been proposed. But out of those, very few have survived the test of time. Indeed, most developments by researchers have either been forgotten and never updated or were never made public. That is the case for most solutions that were developed in the framework of private projects. In the end, the solutions that stayed in use and are still available today are the one that were backed up by powerful commercial platforms or active communities.

Initially developed for research and scientific purposes, IDRISI GIS has rapidly grown to become an important actor in the field of geographical information systems [Idr, ]. As such, an impressive number of modules have been added to it since its debut and it was the first commercial GIS to ever integrate routines for the SMART methodology and for the determination of the weights using Saaty’s method. These routines were well documented [Eastman et al., 1993] and the authors also proposed examples of usage on raster as well as vector data. Later extended, the current version of the software includes a complete MCDA module with support for the Ordered Weighted Average (OWA), MOLA heuristic, and Analytical Hierarchy Process (AHP).

Of course, we couldn’t talk about commercial GIS without mentioning the most widely used software suite. ESRI’s ArcGIS software suite is now one of the best known alternatives on the market. Proposing systems for virtually every type of application requiring the management of geographic information, ESRI has put a great effort into making its systems very accessible. The ArcScripts section of their website has now for a long time allowed developers to propose their own functions and additional packages to complete the existing systems [Esr, ]. Among those scripts, one can find that several methods are available: AHP, OWA, SAW, TOPSIS ... ArcScript is unfortunately not open anymore for new submissions, but ESRI has added several other services in replacement. This behavior is
typical of most software platforms, which offer the possibility of extending their functions to attract user communities.

A third system that deserves praise is the DECERNS SDSS, a project being developed at the IATE (Obninsk University, Russia) and integrating MCDA and GIS in a single web-based architecture [Sullivan et al., 2009]. Since they share the same database, this system allows the user to rapidly switch to any tool at any moment of the analysis. They can thereby engage in a more iterative process of testing the robustness of a model and changing the parameters before giving final recommendations. The system implements an impressive number of MCDA methods (e.g. MAVT, MAUT, AHP, TOPSIS, PROMETHEE, SMAA, FlowSort ...) and support tools (e.g. walking weights, line weights, sensitivity analysis ...). However, even though both tools have been developed together, they were kept separate as two independent subsystems that are able to interact. This system is therefore technically a complete integration, but behaves like a tight coupling. Indeed, from the user’s view point, the two tools seem separate as he/she has to change interfaces to use either one. Yet, some improvements are planned which means that DECERNS might very well become one of the first complete systems to provide a full integration experience for the users. Figures 5.2 and 5.3 show screenshots of the
Decision-Making Support (DMS) subsystem and the GIS subsystem.

The last three systems were examples of integrated systems, but there are other options to solve a decision problem using MCDA and a GIS. As we said earlier, most analysts resort to loose coupling solutions to do preliminary analyses of the problems if not entire studies. For problems that do not require too complex analytical functions, there is indeed the possibility of coupling separate software for the different steps of the analysis. Analytical software such as R, Mathworks Matlab, and Microsoft Excel have since long been included in spatial analyses with some of the major GIS. For that purpose, most of them are able to export their attribute data in some of the most popular formats such as Excel or CSV.

5.3.2 Existing Spatial or MCDA Tools

Aside from the already integrated MCDA-GIS solutions, there are several tools, libraries, and software parts that could be used to produce a working MCDA-GIS solution for a specific purpose. As we have just seen earlier, most GIS are able to extend their functionalities using scripts, but that is not the only way.

Systems like GRASS GIS, an open source desktop software, offer the possibility to the
users to develop modules or plug-ins to enhance the already impressive offer of functionalities. After getting used to its complex interface, one can discover the most complete open source GIS with support for a large panel of formats. Other examples of extensible GIs are ESRI’s ArcGIS, QuantumGIS, SAGA GIS ...

Some commercial GIS such as ESRI’s ArcGIS or Microsoft’s MapPoint are also available as APIs (i.e. Application Programming Interfaces) which are components to be used to add GIS functionalities to almost any program one could have done on its own. It is thereby possible to add map visualization possibilities or management of spatial data or even access to some spatial analysis functions. Finally these APIs can also be independent from any complete GIS software. GeoTools and GeotoolKit are examples of libraries that can be used to add geographic functionalities to existing development projects.

On the other hand, when we consider MCDA components for development. There are almost none. The only available libraries are now part of Decision Deck [Dec, b], a group of initiatives acting to offer an open source platform for developing, designing, and sharing multi-criteria tools.
5.3.3 Technical Difficulties Encountered

Even though so many tools exist to help us with the integration, there are still some pieces missing. Indeed, there are little to no MCDA open source libraries or software parts. Also the semantic differences between the two fields or within each field make it difficult to conceive a unique system with both types of functionalities. It is indeed necessary to agree on a common language to be used in this type of integration.

Also, since the MCDA functionalities are more likely to be integrated in a GIS than the opposite [Malczewski, 2006], there is a need to select some specific methods from the multi-criteria field and find ways to help a user decide on which one to use and how to do that. Several works have focused on the choice of an MCDA method depending on several characteristics of a problem (decision problematic, type of alternatives, ...) but unfortunately there haven’t been any implementations that were made public.

5.4 Usability Aspects of the Integration

As already stressed by Malczewski [Malczewski, 2006] in all the works that have been realized on MCDA-GIS integration, there is a lack of usability studies aimed at expressing the needs these systems should fulfill. In this section, we aim to clear the ground and lay the foundations for such usability studies. Among others, we will ask ourselves the following questions: who will be the users of such systems and what will be their goals which such tools?

5.4.1 The Users

In order to better assess the needs that integrated systems have to fulfill, it is necessary to ask ourselves who will be their users. And the question is not that easy to answer either because there can be several stakeholders with different profiles in a decision problem. It seems obvious that the ones that will most benefit from an integrated MCDA-GIS system are of course the analysts that are familiar with MCDA and spatial analysis. But let us not forget all the other roles in a decision problem. We have the experts that can offer their help in evaluating alternatives or parameters for the methods. We also have the decision makers that have to indicate their preferences and give an idea of where their interest lies. These roles can in some cases be assumed by a single person.

These are thus the main users we can encounter in a typical decision problem:

- The analyst: has extensive knowledge of the MCDA and spatial methods used and is in charge of modeling the problem as well as giving the final recommendations to the decision makers.
• The experts: are in charge of providing the data for the effective comparison of the problems alternatives.
• The decision makers: give information on their way of thinking and the aim they pursue and expect an understandable report on the decision they need to make and the reasons behind the obtained results.

Naturally a complete study would not stop at this, but would involve characterizing user groups for each of the previous roles by analyzing:
• who the end-users are (profiles, knowledge of the methods used...),
• which tasks they perform with GIS and MCDA in order to achieve specified goals, when and where they use the different functionalities of such a system,
• the way they react and what they expect from the usage of a MCDA-GIS platform.

In order to collect data, one could then make use of the usability checklist developed by Johnson et al. [Johnson et al., 1989] or some of the later works by Davies et al. [Davies and Medyckyj-Scott, 1996]. It would allow us to evaluate MCDA-GIS user interface mockups in terms of effectiveness, efficiency, and user satisfaction. On a more amusing note, one can notice that Johnson’s studies tend to evaluate the users’ perception of the system based on several criteria. The aim however is absolutely not to choose, rank or sort anything (i.e. the typical objectives of MCDA aggregation methods), but merely to evaluate the different aspects of a given user interface as well as the system behind it.

5.4.2 The MCDA-GIS Problem Solving Process

Now that we know which roles we have to pay attention to, taking a deeper look into the entire decision process for a problem may help us identify the needs of all these persons as well as the kinds of systems that will be able to cover them.

Several researchers have proposed concepts of flowcharts for decision processes, starting with Malczewski [Malczewski, 1999] in an attempt to identify the tools that were most needed for each step of the analysis. Figure 5.4 shows a flowchart in three phases that are composed of the following steps:

1. Intelligence Phase: The analyst gathers data on the problem in order to determine the type of method to use and stores the information in the GIS.
   • Problem Definition: This step defines the objective of the process. It consists in choosing a particular decision problematic (choice, ranking, sorting, description).
Evaluation Criteria: Since some of the criteria might be spatial, the analyst can use the GIS to define them with the help of experts.

Constraints: Since the alternatives for these problems will be spatially-referenced, the analyst can use the GIS to define a number of constraints that delimit them.

2. Design Phase: After gathering all the needed data, the analyst establishes the model for the decision problem, in a second phase that will mostly rely on multi-criteria functionalities.

Alternatives: Once the constraints have been entered if any, the analyst can generate a list of alternatives or define the set of feasible solutions.

Decision Matrix: When both the criteria and the alternatives have been defined, the analyst can start filling the evaluation table with the help of experts if necessary. In some cases, if all the necessary data had been gathered before, this step could be automatic.

Decision Maker’s Preferences: At this point the decision maker can indicate his view of the problem and help the analyst enter the parameters for the chosen method.

Decision Rules: Depending on the method used, there might be a need to resort to decision rules too.

3. Solution Phase: Finally, once the model has been completed, can it be used to explore the set of feasible solutions, usually in company of the decision maker.

Sensitivity Analysis: This step will serve to verify the robustness of the obtained solution or to adjust the model according to the decision maker’s wishes.

Recommendation: At the end of the analysis, the analyst is expected to provide a robust conclusion and its justification to the decision maker concerning the problem and explain how to implement it.

There are of course other flowcharts that have been proposed more recently by other researchers [Chakhar and Martel, 2003], but this one already tells us a lot on what needs to be done by each person involved.

5.4.3 The Importance of an Integrated System

From the previously analyzed flowchart, we can tell that up till now only the first two phases have been correctly implemented in the existing systems. Indeed those two phases only serve to establish the decision model and obtain an initial solution. That can be done easily in tightly coupled solutions such as a MCDA module in a GIS.
To further explore the possible solutions with the decision maker would require an integrated interface that allows for changes in the parameters while keeping the map displayed and constantly updated. Interactivity is what could be the most valuable function in an integrated system. Some works on visualization could help with this aspect [Andrienko and Andrienko, 2006, Bertin, 2010, Lidouh et al., 2011]. Unfortunately with the few existing systems that currently allow it, one must settle for separate interfaces that create a transition between one scenario and the next with some modifications. It is currently not possible for MCDA modules to interactively change the displayed results on a map by modifying the parameters of a method.

Finally, let us not forget that the usual MCDA problematics are not the only ones that matter in this field. There are also several analysis techniques that couple MCDA and spatial aspects to produce results that support a decision. For example some methods allow you to segment a territory based on several criteria [Tavares-Pereira et al., 2007b], to study the evolution of decision maps [Metchebon Takougang, 2010], or even to compare different maps according to several criteria. All of these can be used to help the analyst at several steps of the analysis. Unfortunately though, there isn’t yet to our knowledge a complete list of these types of works since they are quite recent.
5.5 Conclusion

In this chapter, we took an interest in two fundamental aspects of the MCDA-GIS integration: the technical aspect and the operational aspect.

We have explored the current offer in the MCDA-GIS field. There have been numerous works but very few have lasted until now and continue to be used by geographers or spatial analysts. There is however a lack of multi-criteria tools made for the purpose of helping with developments. Indeed almost no projects are available on the key distribution channels for open source developments.

Integrating the two fields obviously brings together their advantages but also their drawbacks. Where two separate systems offer more flexibility, an integrated system can offer robustness and the possibility for higher interactivity with the decision processes. It is therefore quite difficult to predict which should be used when ignoring the characteristics of the problem to be solved.

A usability analysis should be done to assess the expectations of all the actors in a decision problem and to help them communicate and work together. The brief look we have directed to this aspect has shown us that the analysis and communication of the results can truly benefit from the integration in a single system.

In the end, it seems the question we should ask ourselves is not whether we need to integrate, but how we should do it. In this world where data is becoming much easier to obtain, bringing together different disciplines and finding new synergies are what we should aim for. Thus we need to focus on the usability of the systems and tools that we develop and determine who will be the users and what are their needs.
Chapter 6

Extensions and Variants of Existing MCDA Methods

6.1 Introduction

As we have seen in Chapter 1, the GAIA method is a descriptive and graphical extension of the PROMETHEE method. It permits to represent in two dimensions the actions and criteria involved in the MCDA problem [Brans and Mareschal, 1994]. This type of approach is quite rare in MCDA and it will be interesting to adapt it for cartographic purposes. The GAIA methodology is based on a principle components analysis (PCA) of the PROMETHEE unicriterion flows but it may thus suffer from loss of information in the two dimensional view. This loss of information may be expressed by inconsistencies with the PROMETHEE rankings. Since our objective will be to integrate information extracted from the GAIA plane on a geographical view of a studied decision problem, loss of information is something we will want to avoid at all costs.

Recently, Hayez et al. presented some new extensions of the GAIA method in order reduce these inconsistencies [Hayez et al., 2009]. They proposed for this purpose two GAIA-type visual representations: the GAIA-stick and GAIA-criterion representation.

In this chapter we analyze another way of decreasing the loss of information in the cases of these three techniques. The GAIA method does not take the weights into account given by the decision maker although these weights are used for the computation of the PROMETHEE ranking. For this reason, we have slightly adapted the initial GAIA method and taken the weights into account of the different criteria when projecting the actions and criteria into the GAIA plane. The purpose of this chapter is thus to analyze the impact of taking the weights into account in the different GAIA projections. This includes measuring the induced gain or loss of information.

These approaches were initially published in an article of the International Journal of
Information and Decision Sciences in 2011 [Nemery et al., 2011]:


To achieve this analysis some illustrative examples are given as well as numerical simulations. From these experiments we may conclude that it may be worthy to take the weights of the decision maker into account when representing the decision problem.

The chapter is organized as follows: in Section 6.2 we describe the different variants of the GAIA method. In Section 6.3 we present the weighted GAIA methods and their properties. The numerical analyzes which inform us about the gain or loss are given in Section 6.4. In the last section of this chapter, we present our conclusions and give some further research directions.

### 6.2 The GAIA Methods

In this section, we will briefly describe the variants that the GAIA method has inspired. We take this opportunity to introduce two particular GAIA planes: the GAIA-stick plane and the GAIA-criterion plane. These planes, as we will see, avoid some shortcomings from the classical GAIA plane.

As already explained in Section 1.4.2, the projection from a k-dimensional space to a 2D plane leads often to loss of information (when $k > 2$). There can thus be some distortions between the raw data and their GAIA representation for low values in the unicriterion net flows matrix $\Phi$ [Hayez et al., 2009]. At first, the ranking of the actions in the GAIA plane, obtained by their projections on the decision stick, is not always completely consistent with the PROMETHEE II ranking. Besides, the ranking of the actions on a particular criterion is not always coherent with the unicriterion net flows. An illustration of these assertions can be found in the article by Hayez et al. [Hayez et al., 2009].

In order to address these problems, Hayez et al. proposed two different representations: the GAIA-Stick and the GAIA-Criterion planes. For both planes, the eigenvectors are computed differently and lead to two different representations.
6.2.1 GAIA-Stick Plane

In the GAIA-Stick plane, the first (horizontal) dimension of the plane is defined by the unit \( L_2 \) norm weight vector:

\[
1_w = \left( \frac{w_1}{\sqrt{\sum_{j=1}^{k} w_j^2}}, \ldots, \frac{w_k}{\sqrt{\sum_{j=1}^{k} w_j^2}} \right)
\]  

(6.1)

Given this first dimension, the projections of the actions on this axis are necessarily proportional to the net flow of the PROMETHEE II ranking. The second orthogonal axis results from the first component of the PCA applied to the projections of the actions on the subspace orthogonal to the weight vector. This ensures that the GAIA-Stick plane gathers as much information as possible. The coordinates of action \( a_i \) in this subspace, noted \( b_i \), are obtained as follows:

\[
b_i = \alpha_i - \left( \alpha_i \frac{w_i}{\sqrt{\sum_{j=1}^{k} w_j^2}} \right) 1_w
\]

(6.2)

where \( \alpha_i \) is the \( i \)th row in the \( \Phi \) matrix.

The quantity of information contained in the GAIA-Stick plane, noted \( \delta_S \), will obviously be lower than \( \delta \). It can be computed as follows:

\[
\delta_S = \frac{\sum_{i=1}^{n} u_i^2 + \sum_{i=1}^{n} v_i^2}{\sum_{i=1}^{n} \sum_{j=1}^{k} \phi_j(a_i)^2}
\]

(6.3)

where \( u_i \) and \( v_i \) are the coordinates of the actions’ projection on the plane and \( \phi_j(a_i) \) is the unicriterion net flow of the action \( a_i \) for criterion \( f_j \).

Contrarily to what was proposed by Hayez et al. for the computation of the \( \delta_S \) value [Hayez et al., 2009], we will use the projection of the actions \( a_i \) on the decision stick instead of the net flow values. That is because even though the projections are proportional to the net flows, they are not equal to them and thus lead to an incorrect \( \delta_S \) value. Additionally, let us note that Equation (6.3) also works on the regular GAIA projection and gives us the same result as Equation (1.32).

Working in the GAIA-Stick plane ensures us that the projections of the actions on the
decision stick are completely coherent with the PROMETHEE II ranking. Nevertheless, there still might be a loss of information. This may lead to the fact that the projections of each criterion axis in the plane are not coherent with the unicriterion net flows.

6.2.2 GAIA-Criterion Plane

On the other hand, in order to coherently represent the unicriterion net flows (for a chosen criterion), the decision maker may use the GAIA-Criterion plane. In this plane, the first (horizontal) dimension of the plane corresponds to the selected criterion. The second vector is orthogonal to the first one and corresponds to the first PCA applied to the \( k - 1 \) dimensional subspace orthogonal to the chosen criterion.

Obviously, there may exist \( k \) different GAIA-Criterion planes and the ranking of the actions according to the projection on the first vector is completely coherent with the uni-criterion net flows.

The quantity of information in the GAIA-criterion plane is given by:

\[
\delta f_k = \frac{\mu_1^k + \sum_{i=1}^{k} \phi_k(a_i)^2}{\sum_{j=1}^{k} \lambda_j} \tag{6.4}
\]

where \( \mu_1^k \) is the largest eigenvalue from the subspace PCA.

Working in the GAIA-Criterion plane ensures us that the projections of the actions on the criterion axis are completely coherent with the unicriterion net flows. However, nothing may be said about the consistency of the projections on the decision stick in regards to the PROMETHEE II ranking.

6.2.3 Role of the GAIA-Stick and GAIA-Criterion Planes

To summarize this section, the GAIA plane is thus a method which permits to represent the decision problem in a two-dimensional view. Therefore, it uses a PCA analysis. However, due to the inevitable information loss, the GAIA plane may present some inconsistencies. To avoid these, the decision maker may use the GAIA-criterion and GAIA-stick planes in order to avoid some rank inconsistencies or unicriterion flow projections' inconsistencies. These two additional types of projections can help a decision maker explore the criterion space. As they focus on particular dimensions, they can be used to obtain a graphical representation of the PROMETHEE II ranking or for comparisons of alternatives on a given criterion.

This means that a decision maker interested in all aspects of the problem should use
these variants alongside the regular GAIA plane. Indeed, the GAIA-Stick and GAIA-Criterion planes bring a different insight on the problem but do not replace the GAIA projection.

6.3 The Weighted GAIA Methods

In this section we describe a new approach to visualize the decision problem. The aim will be to represent even more information than the regular GAIA method by taking the criteria weights into account. Indeed, the regular GAIA method uses only the preference structure information (i.e. the intra-criteria information) from the decision maker and then adds a decision stick that represents the objective (i.e. the inter-criteria information). However, the view remains the same whatever weight distribution is chosen. In our approach, we will adapt the view to the weight distribution at all times.

6.3.1 Taking the Weights into Account

The idea behind the weighted GAIA approach will be to integrate the weights’ information to the cloud of points representing the actions before proceeding with the principal component analysis. We will conduct this process with the aim of increasing the amount of information displayed on the plane.

Let us consider the Φ matrix we defined in Equation 1.27. It is the unicriterion matrix used to position the alternatives in the criteria space. In order to take the weights into account, we will compute a matrix $\Phi^w$ as follows:

$$
\Phi^w = \Phi \times w^T = (\phi_j(a_i)\omega_j)
$$

where $w^T$ is a column matrix containing the weights of all criteria. In our new weighted matrix, all rows of unicriterion net flows have been multiplied by the weights thereby adding the information of relative importance to the values. Each net flow of an alternative $i$ for the criterion $j$ is thus equal to $\phi_j^w(a_i) = \phi_j(a_i)w_j$. This leads to the conclusion that the sum of each row will give us the global net flow of an alternative as $\phi(a_i) = \sum_{j=1}^{k} \phi_j^w(a_i)$.

We then apply the Principal Component Analysis on our newly obtained matrix $\Phi^w$ instead of the initial $\Phi$. As a consequence the set of points $a_i$ will be rearranged so that the evaluations of an important criterion are spread out on the axis while the evaluations of a less important criterion are pulled together. The application of the usual steps for the GAIA method will thus lead to different results. We first compute the modified matrix $C'$:

$$
nC' = \Phi^wT\Phi^w
$$
Then we find the two eigenvectors $u'$ and $v'$ of matrix $C'$ with the highest eigenvalues $\lambda_1'$ and $\lambda_2'$. Finally we use the obtained eigenvectors to project the coordinates in matrix $\Phi^w$ on the new weighted GAIA plane.

As each dimension of our criteria space has undergone a different scaling (i.e. anisotropic scaling), we will have to apply the same transformation to the criteria axes before displaying them. Each criterion axis will therefore have to be multiplied by its respective weight.

The decision stick $\pi$ will also have to be transformed. It was initially displaying the vector of weights in order to give us insight on the objective of the decision maker. As the weights have already been taken into account, the decision stick will only serve to point at the direction that simultaneously maximizes (and respectively minimizes) all criteria. However we also have to note that one of the decision stick’s initial properties was also to give us the global ranking of the alternatives when these are projected onto it in the criteria space. Thus the following relation still needs to be verified:

$$\alpha_i \pi_w = \phi(a_i)$$ (6.7)

Since the new coordinates of the $\alpha_i$ points already take the weights into account, their sum is equal to the net flow. We must therefore choose a decision stick on which the projection will have the same effect as a sum of all coordinates. Let us define the new decision stick as:

$$\pi_w : (1/k, 1/k, ..., 1/k)$$ (6.8)

where $k$ is the number of criteria.

The usual formula for computing the amount of information preserved on the GAIA plane is still valid. We can thus compute the $\delta^w$ value as:

$$\delta^w = \frac{\lambda_1' + \lambda_2'}{\sum_{j=1}^{k} \lambda_j'}$$ (6.9)

As the shape of the cloud of points has been changed and the dimensions of lesser importance have been greatly reduced, the projection obtained by the weighted approach will focus on displaying the most important dimensions of the problem only. This will give us better results when we evaluate the amount of information on the plane.

### 6.3.2 GAIA-Stick and GAIA-Criterion

Just as we adapted the original GAIA method, we can apply this process to the already existing variants of GAIA introduced by Hayez et al. [Hayez et al., 2009]. In both cases, we will set one of the dimensions of the plane as either the new decision stick $\pi_w$ or one
of the criteria. The second dimension of the plane will be determined by applying a PCA on the subspace orthogonal to the first dimension and obtained as follows (in the case of the GAIA Stick variant):

$$\alpha_i - (\alpha_i \pi_w) \pi_w$$

(6.10)

In the weighted GAIA-Criterion approach, even less adaptations are necessary. Once the weights have been applied to the $\alpha_i$ points, the plane for the projection is determined by choosing one of the criteria as the first dimension, then applying a PCA on the subspace that does not contain the selected first dimension.

The amount of information in the plane can be computed by using the same corrected formulas that we defined in Equations 6.3 and 6.4. In the GAIA stick variant, as we change the shape of the cloud of points to better reflect the weights, the weighted GAIA stick approach will almost always gives us better results than its regular equivalent because the decision stick $\pi_w$ will be pointing in the same direction as the cloud of points. As for the GAIA criterion variant, in cases where we focus on the most important criteria, the $\delta_w$ value will be higher than the usual $\delta$. When trying to display criteria of lesser importance however, the weighted approach will struggle to produce results as good as the regular one.

6.3.3 Differences in Use

When applying the weights to the cloud of alternatives in the criteria space, we reduce the dimensions of this cloud making them very small when the weight of a particular criterion is weak. As a consequence, the new cloud of points is more representative of the decision maker’s true preferences and is easier visualized using a two-dimensional projection.

Compared to the regular GAIA method, it means that we will obtain a new cloud of points and a new projection for each set of weights that we choose. The regular approach only proposes a single projection for any set of weights. This makes it so that even if the quantity of information conserved by the plane is important, there will be weight configurations where a small loss of information will greatly affect the interpretations that we make using it.

The only case where this approach does not help us any further is when all the weights are equal. In such a situation the cloud of points is not deformed and the projection obtained is the same as the one from the GAIA method.

In Figure 6.1, we display the $\delta_w$ value for a weight configuration whose distance from the equally weighted case is computed using:

$$\text{dist} = \sum_{i=1}^{k} |w_i - 1/k|$$

(6.11)
Figure 6.1: Quantity of information preserved due to a diversification of the weights

To obtain this figure, we have generated a random set of data comprised of evaluations uniformly chosen between 0 and 100, thresholds selected the same way, and a set of equal weights as reference. Then, by choosing 500 sets of weights with uniformly distributed values that were normalized, we obtained the cloud of points in Figure 6.1.

The horizontal line indicates the quantity of information preserved using the regular GAIA method, i.e., $\delta$. That quantity is of course constant whatever the weights used. The dots indicate several observations of $\delta_w$ at a given $\text{dist}$. Dots under the line indicate a loss of information, while dots above the line indicate a gain of information when applying the weighted GAIA method.

We can see that the more the weights are diversified (i.e., the higher $\text{dist}$), the more often we have an increase of the quantity of information preserved. Furthermore, when the distance between a given weight configuration and the equally weighted configuration becomes greater, the gain in information is also significantly affected. One downside to the weighted approach is that applying the weights will reduce all distances and force us to zoom on the projection to be able to use it. This, however, can be solved by implementing an automatic zoom feature in the software used to apply the method.

It will also be interesting to ask ourselves the following question: where on the GAIA plane does the loss of information occur? To answer it we would need to quantify the loss of information on each alternative projected on the plane. We can do so by transforming
Equation 6.3 and then dissociate it into a $\delta_{a_i}$ value for each alternative:

$$
\delta = \frac{\sum_{i=1}^{n} u_i^2 + \sum_{i=1}^{n} v_i^2}{\sum_{i=1}^{n} \sum_{j=1}^{k} \phi_j(a_i)^2} = \frac{\sum_{i=1}^{n} (u_i^2 + v_i^2)}{\sum_{i=1}^{n} \sum_{j=1}^{k} \phi_j(a_i)^2} = \sum_{i=1}^{n} \delta_{a_i} \tag{6.12}
$$

where:

$$
\delta_{a_i} = \frac{u_i^2 + v_i^2}{\sum_{j=1}^{k} \phi_j(a_i)^2} \tag{6.13}
$$

Naturally, the sum of all $\delta_{a_i}$ values is equal to the $\delta$ for the entire plane. We can also display this information on the plane using colors, for example. The illustrations in the next section will feature this element.

### 6.3.4 Conclusion

This new instance of the GAIA method produces some interesting results but is not without drawbacks. Indeed, it is destined to focus on the most important aspects of the problem, giving more importance to the criteria that are relevant for the decision. It also produces a static view which is well adapted for the display of results on other media such as maps, etc.

However, it is not well suited for an exploration of the criteria space as even the slightest change in weights can affect the entire projection and will lead to a re-computation of the representation. Another flaw is the fact that the PCA is now even more affected by extreme values than it was with the regular GAIA method. Yet this issue is of less importance as the decision maker might precisely be interested in knowing about those characteristics of the problem.

In the next section we will illustrate the gain of the approach qualitatively by considering some applications.

### 6.4 Numerical Examples and Simulations

The initial article to present these techniques [Nemery et al., 2011] used several examples of decision problems taken from the PROMETHEE-GAIA literature [Mareschal and Brans, 2002]. In this chapter we will take two different cases that analyze spatially referenced alternatives. Both cases use the 27 member states of the European Union (EU). They study two wellknown indices:
• The Human Development Index (HDI) as defined and computed on the UNDP website \(\text{http://hdr.undp.org}\).

• The Environmental Performance Index (EPI) as presented by the Yale Center for Environmental Law & Policy \(\text{http://epi.yale.edu}\).

### 6.4.1 The Human Development Index in European Countries

This index uses four criteria (as defined in 2009):

1. Life expectancy at birth (in years)
2. Adult literacy rate (in %)
3. Combined gross enrolment ratio in education (in %)
4. GDP per capita (in purchasing power parity US$)

<table>
<thead>
<tr>
<th>Country</th>
<th>HDI 100%</th>
<th>Life expectancy 50%</th>
<th>Adult literacy 20%</th>
<th>Enrolment ratio 30%</th>
<th>GDP per capita 0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0.9506</td>
<td>79.60</td>
<td>99.00</td>
<td>90.50</td>
<td>35523.00</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.9484</td>
<td>79.10</td>
<td>99.00</td>
<td>94.30</td>
<td>32434.00</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>0.8339</td>
<td>72.90</td>
<td>98.30</td>
<td>82.40</td>
<td>10295.00</td>
</tr>
<tr>
<td>Cyprus</td>
<td>0.9121</td>
<td>79.00</td>
<td>97.60</td>
<td>77.60</td>
<td>25837.00</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>0.8972</td>
<td>76.20</td>
<td>99.00</td>
<td>83.40</td>
<td>22004.00</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.9537</td>
<td>78.10</td>
<td>99.00</td>
<td>101.30</td>
<td>35125.00</td>
</tr>
<tr>
<td>Estonia</td>
<td>0.8727</td>
<td>71.30</td>
<td>99.80</td>
<td>91.20</td>
<td>19155.00</td>
</tr>
<tr>
<td>Finland</td>
<td>0.9557</td>
<td>79.10</td>
<td>99.00</td>
<td>101.40</td>
<td>32903.00</td>
</tr>
<tr>
<td>France</td>
<td>0.9547</td>
<td>80.40</td>
<td>99.00</td>
<td>95.40</td>
<td>31980.00</td>
</tr>
<tr>
<td>Germany</td>
<td>0.9401</td>
<td>79.30</td>
<td>99.00</td>
<td>88.10</td>
<td>31766.00</td>
</tr>
<tr>
<td>Greece</td>
<td>0.9487</td>
<td>79.10</td>
<td>97.00</td>
<td>101.60</td>
<td>31290.00</td>
</tr>
<tr>
<td>Hungary</td>
<td>0.8766</td>
<td>73.10</td>
<td>98.90</td>
<td>90.20</td>
<td>18154.00</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.9607</td>
<td>78.60</td>
<td>99.00</td>
<td>97.60</td>
<td>40823.00</td>
</tr>
<tr>
<td>Italy</td>
<td>0.9444</td>
<td>80.40</td>
<td>98.80</td>
<td>91.80</td>
<td>28828.00</td>
</tr>
<tr>
<td>Latvia</td>
<td>0.8650</td>
<td>72.30</td>
<td>99.80</td>
<td>90.20</td>
<td>15389.00</td>
</tr>
<tr>
<td>Lithuania</td>
<td>0.8706</td>
<td>72.70</td>
<td>99.70</td>
<td>92.30</td>
<td>15739.00</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0.9927</td>
<td>78.60</td>
<td>99.00</td>
<td>94.60</td>
<td>77089.00</td>
</tr>
<tr>
<td>Malta</td>
<td>0.8939</td>
<td>79.20</td>
<td>91.40</td>
<td>81.30</td>
<td>21715.00</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.9582</td>
<td>79.40</td>
<td>99.00</td>
<td>97.50</td>
<td>36099.00</td>
</tr>
<tr>
<td>Poland</td>
<td>0.8751</td>
<td>75.30</td>
<td>99.30</td>
<td>87.70</td>
<td>14675.00</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.8999</td>
<td>77.90</td>
<td>94.60</td>
<td>88.80</td>
<td>20845.00</td>
</tr>
<tr>
<td>Romania</td>
<td>0.8257</td>
<td>72.20</td>
<td>97.60</td>
<td>79.20</td>
<td>10433.00</td>
</tr>
<tr>
<td>Slovakia</td>
<td>0.8723</td>
<td>74.40</td>
<td>99.00</td>
<td>80.50</td>
<td>17837.00</td>
</tr>
<tr>
<td>Slovenia</td>
<td>0.9247</td>
<td>77.70</td>
<td>99.70</td>
<td>92.80</td>
<td>25021.00</td>
</tr>
<tr>
<td>Spain</td>
<td>0.9490</td>
<td>80.70</td>
<td>97.40</td>
<td>96.50</td>
<td>29208.00</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.9566</td>
<td>80.70</td>
<td>99.00</td>
<td>94.30</td>
<td>34056.00</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.9423</td>
<td>79.20</td>
<td>99.00</td>
<td>89.20</td>
<td>32654.00</td>
</tr>
</tbody>
</table>

Table 6.1: Performance matrix for the HDI problem
For the sake of comparing the two approaches we have used several weight sets and parameter values. For the preference functions we used usual functions where the indifference and preference thresholds are both equal to zero. With just four criteria considered in this problem, the amount of information available in the GAIA plane is quite high ($\delta = 85.24\%$) as can be seen in Figure 6.2.

![Figure 6.2: GAIA plane for the HDI Index problem](image)

However even with such a high $\delta$ value, the weighted GAIA method gives us an improved projection for sets of weights that diversify the importance of each criterion. One of the results, obtained for the set $w_1 = 0.5, w_2 = 0.2, w_3 = 0.3, w_4 = 0.0$ is represented in Figure 6.3 and displays a $\delta$ value of $93.23\%$.

Furthermore, we have displayed the $delta_{a_i}$ values of each alternative on both planes to locate losses of data. Alternatives in blue present a high $delta_{a_i}$ value and are thus close to the projection plane. Alternatives in red are far from the plane and their $delta_{a_i}$ value is low. Alternatives with intermediate values are displayed in purple. As we might expect from the properties of a principal component analysis, the alternatives on the border of the cloud of points end up close to the projection plane and have a high $delta_{a_i}$.

On the GAIA plane in Figure 6.2 we can see some alternatives in red and purple near the center of the cloud of points. However on the weighted GAIA plane in Figure 6.3 there are only two purple alternatives which indicates a smaller loss of data.

The next case will feature more criteria. The great number of dimensions to be taken into account should lead us to even greater differences between the two approaches.
6.4.2 The Environmental Performance Index in European Countries

For this second case study we chose an index with a higher number of criteria with the purpose of triggering a significant loss of data in the process of applying the Principal Component Analysis. The EPI uses an entire hierarchy of criteria with as much as 37 nodes. The first level separates the criteria into two categories: (1) Environmental health and (2) Ecosystem vitality. For the sake of this exercise we will consider only the 10 criteria present at the second level of the hierarchy (as defined in 2010):

**Environmental Burden of Disease** measures child mortality due to diseases.

**Air Pollution (Effects on Humans)** measures indoor air pollution and particulate matter.

**Water (Effects on Humans)** measures access to drinking water and sanitation.

**Air Pollution (Effects on Ecosystem)** measures $SO_2$ (sulfur dioxide) emissions per capita and per GDP.

**Water (Effects on Ecosystem)** measures change in water quantity.

**Biodiversity and Habitat** measures biome, marine and critical habitat protections.

**Forestry** measures forest loss, cover change and growing stock change.

**Fisheries** measures coastal shelf fishing pressure and fish stocks overexploited.
Agriculture measures agricultural subsidies and pesticide regulation.

Climate Change measures CO₂ emissions and renewable electricity.

All of these criteria are indices collected from several sources or aggregated by the Yale Center for Environmental Law & Policy with values ranging from 0 to 100.

Once again we tested several parameter values for this problem and settled for the following ones: for the preference functions we used linear functions with indifference thresholds equal to zero and preference thresholds equal to the highest difference for each criterion; for the weights we chose the same as in the actual computation of the EPI index (as described later in this section).

With such a complex problem we are bound to lose a high amount of information when applying the GAIA method and indeed, we obtain a GAIA plane with a δ value of 47.06% (see Figure 6.4). This means that even if some positions on the plane might seem good or bad, the ranking that we infer will most likely be different from the complete ranking obtained by PROMETHEE II.

Figure 6.4: Weighted GAIA plane for the EPI Index problem
<table>
<thead>
<tr>
<th>Country</th>
<th>Environmental health</th>
<th>Ecosystem vitality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPI 100%</td>
<td>Disease 25%</td>
</tr>
<tr>
<td>Austria</td>
<td>78.1</td>
<td>86.86</td>
</tr>
<tr>
<td>Belgium</td>
<td>58.1</td>
<td>80.96</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>62.5</td>
<td>65.50</td>
</tr>
<tr>
<td>Cyprus</td>
<td>56.3</td>
<td>86.86</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>71.6</td>
<td>75.96</td>
</tr>
<tr>
<td>Denmark</td>
<td>69.2</td>
<td>80.96</td>
</tr>
<tr>
<td>Estonia</td>
<td>63.8</td>
<td>59.41</td>
</tr>
<tr>
<td>Finland</td>
<td>74.7</td>
<td>82.81</td>
</tr>
<tr>
<td>France</td>
<td>78.2</td>
<td>82.81</td>
</tr>
<tr>
<td>Germany</td>
<td>73.2</td>
<td>82.81</td>
</tr>
<tr>
<td>Greece</td>
<td>60.9</td>
<td>82.81</td>
</tr>
<tr>
<td>Hungary</td>
<td>69.1</td>
<td>66.64</td>
</tr>
<tr>
<td>Ireland</td>
<td>67.1</td>
<td>84.77</td>
</tr>
<tr>
<td>Italy</td>
<td>73.1</td>
<td>86.86</td>
</tr>
<tr>
<td>Latvia</td>
<td>72.5</td>
<td>59.41</td>
</tr>
<tr>
<td>Lithuania</td>
<td>68.3</td>
<td>60.35</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>67.8</td>
<td>82.81</td>
</tr>
<tr>
<td>Malta</td>
<td>76.3</td>
<td>86.86</td>
</tr>
<tr>
<td>Netherlands</td>
<td>66.4</td>
<td>86.86</td>
</tr>
<tr>
<td>Poland</td>
<td>63.1</td>
<td>70.31</td>
</tr>
<tr>
<td>Portugal</td>
<td>73.0</td>
<td>79.20</td>
</tr>
<tr>
<td>Romania</td>
<td>67.0</td>
<td>64.40</td>
</tr>
<tr>
<td>Slovakia</td>
<td>74.5</td>
<td>70.31</td>
</tr>
<tr>
<td>Slovenia</td>
<td>65.0</td>
<td>77.54</td>
</tr>
<tr>
<td>Spain</td>
<td>70.6</td>
<td>84.77</td>
</tr>
<tr>
<td>Sweden</td>
<td>86.0</td>
<td>86.86</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>74.2</td>
<td>80.96</td>
</tr>
</tbody>
</table>

Table 6.2: Performance matrix for the EPI problem
The colors for each alternative also show us that there is a great loss of data on most alternatives and only few alternatives on the border are accurately represented by their projection. In a case like this one, where we have 10 dimensions to be analyzed, the GAIA projection unfortunately has a very limited use as far as alternative comparisons are concerned. The next step is to apply the weighted GAIA variant on this case to see if our new approach give us any better results. Among the several weight distributions that we could have used, we chose the following:

\[
\{ w_1 = 25, w_2 = 12.5, w_3 = 12.5, w_4 = 4.2, w_5 = 4.2, w_6 = 4.2, w_7 = 4.2, w_8 = 4.2, w_9 = 4.2, w_{10} = 25 \}
\]

This weight distribution is the same as the one used by the Yale Center for Environmental Law & Policy to compute their index. The resulting projection is given in Figure 6.5. Figure 6.6 shows the same projection, where we have zoomed to focus on the alternatives instead of the criteria.

Figure 6.5: Weighted GAIA plane for the EPI Index problem

The result of the weighted GAIA method is a projection with a significant increase in information preserved (\( \delta = 84.39\% \)). We also see that most points are now accurately represented with only a few central alternatives that have suffered a loss. This is of course a welcome improvement for the geographical representations that we will be designing in the next chapter. As we start from a representation that features such a great improvement compared to the initial projection, we will allowed several transformations while keeping
meaningful interpretations of our results.

These two examples have shown us that using the weighted GAIA projection can help us to better represent the problem, especially in situations where the weights are modified and have very different values. While the regular GAIA plane is a compromise for all weight distributions, the weighted GAIA approach adapts itself and takes advantage of the inequality of the weights to display only the essential parts of the problem.

This same behavior is observed with the GAIA-stick and GAIA-criterion variants as demonstrated in the associated article [Nemery et al., 2011]. This constant adaptation leads to an increase in the quantity of information displayed for all the cases studied.

### 6.4.3 Empirical Tests

Our article also features a series of simulations destined to quantify the gain due to taking the weights into account and its number of occurrences. In the simulations, we have considered two main parameters. At first, the difference of the quantity of information preserved in the weighted case and the non-weighted case is computed. This is measured by the value $\delta^w - \delta$. A positive value indicates that there is a gain of information preserved by this new plane.

Secondly, we have measured the gain or loss of the coherency between the stick projections and the actual net flow values (and thus the ranking of the actions). To do so, we
have used the Pearson coefficients \( \rho \) and \( \rho^w \) defined as follows:

\[
\rho = \frac{\text{covariance}(D_p, \Phi)}{\sigma_{D_p} \ast \sigma_\Phi}
\]

where \( D_p \) is the vector of the projections of the actions on the decision stick and \( \sigma_{D_p} \), \( \sigma_\Phi \) respectively the standard deviations of the stick projections and the net flows. The Pearson coefficient \( \rho \) thus measures the coherency between the stick projections and the net flows in the GAIA plane. Analogously, we define:

\[
\rho^w = \frac{\text{covariance}(D^w_p, \Phi)}{\sigma_{D^w_p} \ast \sigma_\Phi}
\]

which represents the Pearson coefficient for the stick projections in the weighted GAIA plane \( (D^w_p) \) and the net flows.

Clearly, if \( \rho^w - \rho > 0 \) there is a higher coherency between the net flow ranking and the deduced ranking in the weighted GAIA plane. This is indeed an advantage for the decision maker when analyzing his decision problem since the stick projections in the plane are closer to closer the actual net flows (and thus the ranking).

To have an idea about the different values of \( \delta^w - \delta \) and \( \rho^w - \rho \) and their distributions, we have generated random decision problems. For each decision problem, we have drawn the GAIA plane and the weighted GAIA plane and computed the \( \delta^w - \delta \) and \( \rho^w - \rho \).

Each decision problem consisted of a set of \( n \) actions \( (n = 8, 20, 60) \) evaluated on \( k \) criteria \( (k = 3, 6, 9) \). In the simulations we have chosen the linear preference function since this function covers most of the other preference functions. Furthermore, the evaluations of the actions, the weights, and the indifference and preference thresholds are generated randomly using a uniform distribution. For each number of actions and alternatives, we generated 10000 problems for which we calculated the two values. The results of these simulations are represented in a series of figures that are described in the article [Nemery et al., 2011].

Figure [6.7] shows us two samples of those frequency distributions where we see that the values are seldom negative. This indicates that the weighted variants of the GAIA method almost always lead us to an increase in the amount of information represented.

### 6.5 Conclusion

In this chapter, we have proposed a slightly modified version of GAIA, called weighted GAIA, where the weights are explicitly taken into account in the principal component analysis. This leads to different representations as in the usual GAIA plane. However, as
presented in some illustrative examples of the chapter, it may lead to a better representation of the decision problem. In order to quantitatively evaluate the possible advantages, two measures have been defined. At first, we computed the gain in information preserved by the weighted projection. Secondly, we measured the coherency between the stick projections and the rankings.

From the illustrative examples and from the simulations we can conclude that in most of the cases, the decision maker has a better view of his decision problem. He thus gains insights in his decision problems. Moreover, taking the weights into account in the PCA enables the decision maker to perform a sensitivity analysis of his decision problems while considering different settings of weights.

A further research direction could be the analysis of a three-dimensional projection. Adding a dimension should increase the quantity of information retained even more and thus the quality of the representation. However, the understanding of this view may be somewhat more difficult for the decision maker as it would require to move the view to actually be able to perceive all three dimensions. Some works have already tried implement such a view [Macharis et al., 2013].

Finally these extensions will prove valuable once we tackle the cartographic representation of the information from the GAIA plane in the next chapter. As the two main axes on maps are already used to spatially position the alternatives, we will have to resort to other means to represent the alternatives’ profiles. Using a GAIA projection with as much
information as possible will naturally give us more precise results.
Chapter 7

Adaptations of MCDA Methods for Cartography

7.1 Introduction

In Chapter [1], we saw an overview of several visual tools from the MCDA field. Chapter [3] showed us spatial analysis tools used in GIS. In Chapter [4], we presented the concept of integration of these two fields and its evolution up till now. Then, Chapter [5] described the challenges that lie behind such a project and the paths for future research.

We now arrive at the last chapter. This will cover some developments that were made in order to ease the integration of MCDA tools in GIS in general. We will focus on the GAIA method and how it can be adapted both methodologically and practically to be used in a spatial context.

The two graphical tools that will be described in this chapter were first studied in two refereed conference papers. On the one hand, we have a paper that presents the decision clocks: a glyph that focuses on displaying the compatibility of alternatives with a set objective. It was presented at the International Conference on Information Visualisation in July 2009 [Lidouh et al., 2009b]:


On the other hand, a second paper explored how to display profile data without losing too much information in the process. In proposed the use of colored circles associated to a color chart to allow the comparison of multiple profiles. This tool was presented at the Symposium Series in Computational Intelligence in April 2011 [Lidouh et al., 2011]:
In Section 7.2, we go over the elements from the GAIA projection that are displayable on a geographical map. In Section 7.3, we describe the use of decision clocks to show the alternatives’ adequacy with regard to the global objective of the problem. Section 7.4 goes further by showing how colored circles can be used to identify an alternative’s profile and compare several across the map. In Section 7.5, we see how the Weighted GAIA method (developed in Chapter 6) can improve the quality of the results displayed and lead to better interpretations of the problem’s data.

7.2 Extraction of Information to be Displayed from GAIA

As we have seen in the description of the GAIA method in Section 1.4.2, the GAIA plane gives us a lot of visual information on the decision problem. In addition to the geometric data given to us from the projection, we can draw several conclusions from the relations between its components. As a tool that relies heavily on the visual interpretation of its results, it makes a good implement for use in a GIS.

The following is a short list of elements that could be of interest when constructing a geographical view of a region under study.

First we have all the elements that are representable using a single dimension or graphical device:

Net flows of the alternatives The final score for each alternative, obtained through PROMETHEE II, can easily be displayed and help us identify areas of alternatives suitable for the decision problem.

Position of the alternatives in the ranking In some cases, the quantitative aspect of the PROMETHEE II ranking can be left aside to focus solely on the ordinal aspect of the ranking.

Graphical devices for the representation of these pieces of information are numerous and can include any of the color components, sizes of symbols, angles or slopes... Even labels could be considered but they would induce a higher cognitive burden on the user to quickly identify or locate elements of interest.

The second set of elements can only be represented through a combination of several graphical devices or dimensions:
Positions of the alternatives (according to the criteria) The location of an alternative on the plane will give us an indication on the type of profile it has. It will point out the strongest and weakest features of a solution.

Relative positions of the alternatives Groups of alternatives on the plane will represent solutions with similar profiles. The comparison of such an information with spatial proximity is sure to give some interesting insights.

Positions of the alternatives (according to the decision stick) When projected on the decision stick the alternatives take their positions from the PROMETHEE II ranking. Even though the ranking inferred from a projection could present differences due to loss of data, it still is an interesting use of the tool when more precise information is not available.

As all of these elements require at least two dimensions to be represented, they will need to be displayed using glyphs for each alternative. By choosing the appropriate graphical devices and combining them we will design symbols that will allow us to easily draw the same conclusions as on the GAIA projection but with the added knowledge of the geographical location of each alternative.

Finally, some information will not necessarily need to be repeated on each alternative:

Positions of the criteria The orientation of the axes will indicate which criteria are concordant and which ones are discordant. Furthermore their size will point out the discriminant criteria within the problem.

Position of the decision stick In this multivariate view, the indication of an objective is of high importance.

δ value These results wouldn’t be complete without an indication on their reliability. The δ value (i.e. the amount of information preserved by the plane) will give us a confidence level for the results and will have to be indicated alongside them.

These elements would work well as part of the map legend.

As there are several graphical devices that work well for these purposes, displaying each of these elements separately should not be very difficult except for multi-variate information (such as coordinates or profiles). The following sections will focus on displaying combinations of these elements while trying to overcome the challenge of not losing too much information or precision in the process. Each tool defined in this manner will have a particular purpose.
7.3 Decision Clocks: Displaying Proximity to Objective

The decision clocks are symbols that will display the approximate position of the alternatives in comparison with the decision stick. We chose to use symbols to make all information visible, even on a base map. In addition to the positioning and the decision stick both represented by vectors, the symbol will be completed by adding the net flow of the alternative as an information unaffected by the loss of information in GAIA. Finally, as a frame of reference, the criteria axes will be displayed to help locate the alternatives in the criteria space, although this purpose is secondary.

7.3.1 Definition of the Components Used

Positioning of the alternatives This information will be represented by a vector whose origin will be the center of the symbol. As the vectors can go in all directions, the symbol will have a circular shape. Since in most cases, the symbols will have a fairly small size, using a constant vector norm will ensure that the symbols keep a good visibility. In such circumstances, only the direction of the positioning will be of interest. Of course, the user could always be given the freedom of displaying vector norms proportional to the actual distances between the alternatives and the origin of the GAIA plane. The vector representing the positioning of the alternative will be computed as follows in the default view:

\[ \beta_i = \frac{a_i}{||a_i||} \]

Decision stick In order to compare the positioning of the alternatives to the objective pursued by the decision maker, the decision stick will be displayed in each symbol as a normalized vector indicating the direction of the objective. This vector whose origin will also be the center of the symbol will be obtained in a similar way:

\[ \rho = \frac{\pi}{||\pi||} \]

Net flow of the alternatives As a final piece of information, the background of the symbol can be colored to give us the net flow of the alternative (i.e. its score in the PROMETHEE II method). A simple colormap will be used to convert the net flow values to colors as displayed in Figure 7.1. In that particular case, positive values are represented by a green color and negative values by a red color. The RGB components for this colormap can be obtained by using the following formulas. If
the net flow is positive:

\[ R(a_i) = B(a_i) = 255 - 255\phi(a_i) \]

\[ G(a_i) = 255 \]

If the net flow is negative:

\[ R(a_i) = 255 \]

\[ G(a_i) = B(a_i) = 255 + 255\phi(a_i) \]

Figure 7.1: Colormap for the net flow

**Criteria axes** Vectors indicating the direction of the different criteria axes of the problem will then be added separately in the legend of the map. These will serve to understand what differentiates two alternatives that might have obtained a similar score but for different reasons. Just like the vectors inside the symbols, these axes can be normalized to have comparable lengths or stay proportional to their original versions to give us the same information. This should also be left at the discretion of the user.

Figure 7.2 show us some examples of decision clocks that combine all of the aforementioned elements.

Figure 7.2: Examples of decision clocks

**7.3.2 Illustrative Example: City of Brussels**

The following example shows the decision clocks used on a small set of data. The alternatives are the municipalities of the city of Brussels. The criteria are different measures of the inhabitants wealth such as their mean income and the realty prices. The aggregation
of these characteristics should therefore indicate the wealthiest municipalities of the city. Our aim will be to see if there is a relation between the geographic location and the score obtained. This example was described in a conference paper by Lidouh et al. [Lidouh et al., 2009b].

Statistical data on the municipalities of Brussels was available on on the website of Statbel a Belgian institution that offers different types of data (http://www.statbel.fgov.be/). For our example, we will select the following criteria:

- Realty price: the mean buying price of a house in 2005 in euros.
- Mean income: interval containing the mean income of the inhabitants for the year 2003 in euros. For the sake of simplicity, we will use the average of each interval.

The preference functions will all be linear with 0 as the indifference threshold. The preference threshold will be set as the maximum difference between the available evaluations for each criterion.

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Realty price</th>
<th>Mean income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderlecht</td>
<td>205751</td>
<td>19610</td>
</tr>
<tr>
<td>Berchem Ste Agathe</td>
<td>208218</td>
<td>27155</td>
</tr>
<tr>
<td>Molenbeek</td>
<td>190791</td>
<td>19610</td>
</tr>
<tr>
<td>Forest</td>
<td>263945</td>
<td>19610</td>
</tr>
<tr>
<td>Uccle</td>
<td>328204</td>
<td>34630</td>
</tr>
<tr>
<td>Ganshoren</td>
<td>215890</td>
<td>24005</td>
</tr>
<tr>
<td>Jette</td>
<td>216041</td>
<td>25280</td>
</tr>
<tr>
<td>Koekelberg</td>
<td>253522</td>
<td>22765</td>
</tr>
<tr>
<td>Bruxelles</td>
<td>265565</td>
<td>22765</td>
</tr>
<tr>
<td>St Gilles</td>
<td>254598</td>
<td>19610</td>
</tr>
<tr>
<td>Ixelles</td>
<td>390384</td>
<td>25280</td>
</tr>
<tr>
<td>St Josse</td>
<td>229277</td>
<td>19610</td>
</tr>
<tr>
<td>Scharbeek</td>
<td>246045</td>
<td>19610</td>
</tr>
<tr>
<td>Etterbeek</td>
<td>336573</td>
<td>25280</td>
</tr>
<tr>
<td>Watermael Boitsfort</td>
<td>317519</td>
<td>34630</td>
</tr>
<tr>
<td>Auderghem</td>
<td>298571</td>
<td>27155</td>
</tr>
<tr>
<td>Evere</td>
<td>209609</td>
<td>22765</td>
</tr>
<tr>
<td>Woluwe St Lambert</td>
<td>345145</td>
<td>34630</td>
</tr>
<tr>
<td>Woluwe St Pierre</td>
<td>395588</td>
<td>34630</td>
</tr>
</tbody>
</table>

Table 7.1: Performance matrix for the city of Brussels

The first information to strike us in Figure 7.3 is the background color of the decision clocks. As it indicates the global score of the alternatives, we are able to rapidly identify areas where the municipalities have similar scores. These results indicate that most of the richest municipalities are located in the Southern part of the city. Also it becomes now possible to compare profiles of neighboring municipalities.

The second analysis we can do will focus on the reasons why such scores were obtained. Figure 7.4 shows an extract of the previous map where we have zoomed in to focus on
some of the glyphs. By looking at the directions of the arrows, we first observe that even though the alternatives have similar scores, their profiles do not always seem to match. By looking at some specific alternatives and comparing their arrow to the criteria referential, it becomes possible to understand why some municipalities obtain their respective scores.

The example given above is voluntarily quite simple and does not have any strong purpose. However, it shows how intuitive the interpretation of results becomes using this tool. Thanks to the visual aspect of this method, we can easily detect alternatives that are
incomparable even though their ranking is similar, or alternatives that are similar because
of their location.

Since the values used were mean values, this does not necessarily reflect the reality
of the situation in every neighborhood of each municipality. Extending the study to the
neighborhoods of the city would give us a more realistic view of the problem but would
also involve other problems such as aggregation problems or identifying the alternatives
to be considered. The first problem would arise when some of the neighborhoods are
part of multiple municipalities. Then the second one would appear when focussing on a
single municipality: a choice would have to be made on whether or not to consider all the
alternatives in our evaluation.

From this simple case, it becomes clear that the use of the decision clock is limited to
the visualization of the scores and that it only gives us an intuition on the reasons that
lead to the results obtained. In order to appraise other parts of the information provided
by the GAIA plane, we need another type of tool.

In fact, this attempt was not approved by the several communities involved. In par-
ticular, geographers found the tool to be too complex to be interpreted easily. By taking
into account some of the properties seen in Chapter 3, we shifted our attention to other
types of graphical devices.

7.4 Colored circles: Comparing multivariate profiles

For this second tool our aim will be to focus on the exact positioning of the alternatives
on the GAIA plane and to transfer that knowledge (i.e. the profile of the alternative) to a
symbol on a geographical map. From our first attempt, we saw that keeping only a part
of an already incomplete set of data could make us loose a lot of information and lead us
to mistakes in our interpretations. This time, we will therefore focus on maintaining the
exact coordinates of the GAIA projection.

We will begin by describing the design of the tool. We will then illustrate it using two
study cases. Finally we will end this section by describing the weaknesses that this type
of glyph can have.

7.4.1 Definition of the Components Used

Before selecting all the elements we are going to display and choosing their place in the
following symbol, we need to find a way to display 2-dimensional data without modifying
it too much. Our objective will be to define a transformation that we will apply to the
coordinates from the GAIA plane to obtain something displayable on the symbol. Of
course this transformation would need to be reversible in order to allow the user to infer
the initial coordinates from its result. Furthermore this process should be intuitive to let
the user focus on the meaning behind these coordinates.

When considering works on the effectiveness of graphical devices [Bertin, 2010, Mackin-
lay, 1986, Nowell, 1997] one can’t help but notice the contradictions in the different results
they propose. That is due to the hypotheses and different conditions in which the studies
took place. What is thus important is to consider the usage that is intended for these
devices and the type of data to be represented.

In general, color, size, and shape seem to produce the best results for the different
visualization tasks studied. And since our aim will be to represent multivariate profiles
which can contain quantitative or ordinal data, we are going to focus here on color and
size.

Color can usually be decomposed in three components:

- Color hue: which represents the pure color that can be named (e.g. red, blue, green,
yellow, orange...).

- Color saturation: refers to the relative darkness or lightness relative to a grayscale
for light emitting objects. Typical saturation scales go from gray to the bright and
easily distinguishable color.

- Color brightness (or value): refers to the perceived intensity of reflecting objects.
Typical brightness scales go from black to white with the chosen color in the middle.

Figures 7.5, 7.6, and 7.7 show us the three components for color.

![Figure 7.5: Hue scale](image1)

![Figure 7.6: Saturation scale](image2)

![Figure 7.7: Lightness scale](image3)

To represent the positions of the alternatives on a geographical map, and thereby
indicate the type of profile of the alternatives, we will make use of the HSV color model
This model is used to represent the coordinates of color points under the form of a solid cylinder. It stands for hue, saturation, and value, which are the three coordinates used to define all colors as can be seen in Figure 7.8.

This model presents the advantage of being more intuitive for human interaction. This characteristic will thus allow the decision maker to compare the profiles more easily and to quickly identify their position in the GAIA referential.

By forcing the value component to always have its maximum value, we reduce the cylinder to a 2-dimensional circle which we will use to represent the GAIA plane. The positions of the alternatives will thus be represented by the angle at which they are positioned (i.e. the hue) and their distance from the center of the plane (i.e. the saturation). The corresponding color chart is given in Figure 7.9.
By superposing the GAIA plane on the color chart given in Figure 7.9, we can thus identify the color associated with each alternative. Alternatives that are close on the GAIA plane, and therefore have a similar profile, will end up having a similar color.

Once these colors have been chosen for each alternative, they will have to be displayed on the geographical map. But instead of using it to color entire areas, we chose to display the colors within circles of variable diameters. This allowed us to add another piece of information to the representation which is the net flow obtained with PROMETHEE II. The result is a symbol that adequately combines a 2-dimensional result with the global score of the analysis.

The next parts of this section will serve to illustrate this new tool on two ranking problems.

7.4.2 Illustrative Example: The Human Development Index of European Countries

Let us take once again the Human Development Index problem we described in the previous chapter, Section 6.4, and use it to generate a geographic map of Europe with the results. The obtained map is given in Figure 7.10.

In this representation, we have chosen to rotate the referential in order to give a particular meaning to the colors. In this case, the color green indicates the best alternatives as it lies in the direction of the decision axis. This allows us to verify that countries with well-positioned profiles obtain a high score (identifiable by the bigger size of their circle):

- Sweden, France, Netherlands, Luxembourg, Finland, Ireland, and Denmark all obtain a green color and a circle with reasonable size.
- Some countries such as Italy, Germany, and Spain obtain a blue color but still have a high position in the ranking.
- The lowest positions are taken by the countries in purple and pink such as Romania, Bulgaria, and Hungary whose symbols have the smallest sizes.

We see that the profile indeed confirms the rank that is obtained by the countries. As can be expected, there are color clusters of similar profiles on the map which indicates a link between geographic proximity and profile similarity.

7.4.3 Illustrative Example: The Environmental Performance Index of European Countries

This time let us try the Environmental Performance Index. As we had seen in Section 6.4, this problem suffered from a great loss of information due to its high number of criteria.
Figure 7.10: Geographic map of colored circles for the HDI problem

Indeed, its projection only had a $\delta$ value of 47%. We will see what conclusions can be drawn from a map displaying those results.

In this case we chose to orient the referential so that the red color indicates bad alternatives (see Figure 7.11).

We can see that the worst alternatives in the ranking appear in red, orange or purple, such as Romania, Bulgaria, Estonia, Poland, etc. However, red does not indicate the worst alternative which in this case is Estonia. Instead, because of the loss of data on the GAIA plane, Romania’s seems to be the lowest score. The global score is actually given to us by the size of their circle. Even though it might be misleading, the user should see the colors as an indication of a country’s strong or weak points instead of an indication of how good or bad they are:

- Bulgaria, for example, has one of the highest evaluations in fisheries and the highest
Figure 7.11: Geographic map of colored circles for the EPI problem in agriculture.

- Estonia has the lowest evaluations in environmental health and forestry combined with one of the highest in agriculture and air pollution effects on humans.

- Romania has one of the worst scores in environmental health and the worst in water effects on humans combined with one of the best in agriculture and forestry.

In general, we can notice that the high number of criteria has not completely hindered our ability to identify good or bad alternatives on the map since the result from GAIA has been completed with information from the complete ranking. The loss of information in such a case means that the apparent similarity between profiles will not always be as accurate as before. And most importantly, the position of the criteria might not always give us all the information on strong or weak points of the compared alternatives. Nevertheless, the colors still allow us to make interesting observations provided that these are compared to the initial data set.
7.4.4 Weaknesses of Colored Circles

Even though this glyph can be considered an improvement compared to the previous one regarding certain aspects, it is not exempt of weaknesses.

The first one is the high number of colors (hues) present on the colormap. Studies show that the use of colors to convey information gives the best results when the number of colors is within the cognitive limit (i.e. the magical number seven plus/minus two [Miller, 1956], [Wickens and Carswell, 1997], [Cowan et al., 2003]). However, for the purpose intended, we should not focus on the colormap but on the glyphs themselves. Indeed the user will only have to process the colors present on the map, which means that when orienting the colormap correctly, we can reduce the number of hues represented and thereby reduce the cognitive workload for the user.

The second weakness is linked to the varying sensitivity of the human eye to different color components which induces a slight chance of error from the user even though the colormap is displayed next to the geographical view. We can refer to several works by Post et al. or Kaufmann et al. who have published chromaticity diagrams to show that humans do not perceive all colors similarly [Post and Greene, 1985], [Post and Greene, 1986], [Kaufmann and O’Neill, 1993]. Figure 7.12 shows one such chromaticity diagrams obtained by Gage et al. using the CIE color space [Gage and Hewlett-Packard Company, 1977].
In order to avoid such problems, one could use a customizable interface that allows the user to select the best representatives of each named color. A study by Smallman et al. has shown that such an approach can lead to very good results [Smallman and Boynton, 1993]. Furthermore previous works have shown that under some circumstances larger numbers of colors could work well for some applications [Smallman and Boynton, 1990]. Specific studies centered around cartographic representations would therefore be interesting to guide us in such choices.

### 7.5 Weighted GAIA Variant: Increasing the Amount of Information Displayed

As we have seen in earlier sections, the display of GAIA information on a geographical map often comes with loss of data. We have tried reducing this loss by modifying the glyphs used, but this section adopts a new strategy which is to increase the amount of information to be displayed. In this section, we will apply the weighted GAIA method we described in Chapter 6 to increase the amount of information available in the projection used.

When studying the GAIA method we realized that the weights information held a very small role in the construction of the projection. Indeed that information was only added after the actions and criteria had all found their relative positions and only served to explore or highlight parts of the projection. In this variant of GAIA we took the weights into account while constructing the projection thereby obtaining a plane that completely reflects the decision maker’s vision of the problem.

Since only the projection changes and not the process of constructing the glyphs, we will immediately illustrate this new approach on the last case we have used: the Environmental Performance Index of European countries. The projection using the weighted approach had seen a very big increase in conserved information as the $\delta$ value had gone from 47% to 84%. The result of using the colored circles for this new projection is given in Figure 7.13.

With a projection of this quality, the interpretations we make are much closer to the actual data of the problem. We can now identify similar profiles more accurately:

- Sweden seems to stand out with some of the highest scores.
- Denmark*, Finland, France, Germany, Malta, Portugal*, and the United Kingdom have overall good scores with just slight differences in their profiles.
Austria, Ireland, Italy, and Spain have some very high or average evaluations with the exception of climate change and agriculture.

Belgium*, Cyprus, Greece, Luxembourg*, and Netherlands seem to have similar profiles with overall good scores except in air pollution, climate change, and agriculture.

Hungary, Latvia, Lithuania, Poland, Romania, and Slovakia have very good evaluations for water and air pollution. However, their burden of disease is among the lowest in Europe and they have average or low scores for the other aspects considered.

Bulgaria and Estonia share the lowest ranks among the considered countries.

We left out Czech Republic and Slovenia as these alternatives had the highest loss of information according to the $\delta_{\alpha}$ values displayed on the GAIA plane back in Figure 6.6. In fact when we compare their data, it turns out that all the alternatives with a moderate
loss are also the ones that differ the most from the others in the groups we just considered. These countries are marked with an asterisk (*) in the previous list and include Belgium, Denmark, Luxembourg, and Portugal.

7.6 Conclusion

In this chapter, we have presented new ways of displaying multi-dimensional information using a single glyph. The decision clocks allowed us to quickly identify the scores of alternatives and their adequacy regarding the objective. And the colored circles allowed us to display multivariate information such as profiles and make comparisons of alternatives. To do so, the HSV color system allowed us to convert a set of coordinates from the GAIA plane into a color which can then easily be associated to certain characteristics by the user.

We apply both these tools to the GAIA visualization method for ranking problems and thereby displayed the results of a MCDA analysis on a geographical map. These glyphs have several uses. By comparing the alternatives to each other, the user can identify sets of alternatives with similar profiles or detect the best alternatives by appreciating their global score.

Even though we have limited their application to problems with a finite set of alternatives, we might consider applying these glyphs to classification problems in order to deal with higher numbers of alternatives to be compared. These can indeed also be solved using outranking methods such as FlowSort [Nemery and Lamboray, 2008] or ELECTRE Tri [Bouyssou et al., 2006]. The use of other types of methods would indeed be encouraged for cases with very high numbers of comparisons to be made. This tool could also easily be adapted to multi-attribute utility problems.
Conclusion and Perspectives

Conclusion

Spatial decision problems in real life are complex and present several conflicting aspects. For years, researchers and analysts have combined the techniques from spatial analysis and multicriteria decision aiding to help solve these kinds of problems. The solutions proposed were very diverse. Some used the tools sequentially while other sought to implement dedicated systems for these cases. When Geographical Information Systems (GIS) became standardized and started evolving to answer the needs of their users, researchers saw in them an ideal basis to develop complete Spatial Decision Support Systems (SDSS). GIS had become indeed very powerful data management systems capable of collecting, analyzing, modeling and displaying spatially referenced data. The main limitations these systems presented were a lack of analytical capabilities to handle several decision criteria with all the properties that they entail [Laaribi, 2000, Malczewski, 2010, Chakhar, 2006].

On the other hand Multicriteria Decision Aid (MCDA), a sub-domain of Operations Research, has developed entire theories of decision support based on the concept of trying to find a suitable compromise to a problem rather than an optimal solution. MCDA brings several contributions that can be of great use in GIS: different ways to structure a decision problem to consider different aspects, value scaling that allows all the criteria scales to be transformed to comparable units, criterion weighting that allows us to give a different importance to each criterion considered, decision rules that dictate how the choice/sorting/ranking is executed, and sensitivity analysis that evaluates the robustness of the obtained results.

Naturally, both of these fields are quite complex and require a certain level of expertise to be handled correctly. This is one of the reasons that has slowed the integration of MCDA and GIS and restricted its use mostly to the academic and scientific community. The technical difficulties encountered in extending GIS and the sheer variety of existing MCDA methods ensured that most integration works only yielded incomplete results. The differences in semantics of both fields did not help the process either. Through a critical review of the existing works and systems we demonstrated that we still, to this day, have
to see a completely integrated system that implements a decent panel of MCDA tools with results that can be harnessed easily. Furthermore, we have shown that technical integration means nothing if it is not visible to the user on an operational level.

For these reasons we decided to tackle one of the least addressed issues in MCDA-GIS integration: the usability of MCDA methods in a geographic context. Since GIS are essentially graphical tools for the visualization of spatial data, we explored several visual tools from the MCDA field, searching for one that would give us interesting results. Among the works that had tried to integrate both fields, most used the weighted sum approach and very few considered the outranking methods. Among those, even less used the PROMETHEE family of methods and its tools. The visual tool GAIA used in the PROMETHEE methodology to give a two-dimensional view of the decision problem seemed like a good candidate for an adaptation in geographic maps.

In order to transform the result obtained from the GAIA process into something displayable on a map, we decided to break down the GAIA representation into all of its components and carefully select which ones to try and display. By using some notions from information visualization we selected the most appropriate graphical devices to display each type of data and built two symbols to be added on top of each alternative on the map.

The first symbol, called Decision Clock, allows us to easily compare the scores of the alternatives and their adequacy to the global objective. The second symbol, is a colored circle of variable size. It allows us to display the profile of each alternative and compare them to find clusters of similar alternatives on the map. Using these tools we see that several synergies are possible as they allow us to make interpretations about the problem that take into account both the multicriteria evaluations of the alternatives and their geographic position.

An additional contribution aims at improving the quality of the results and the conclusions we draw from them by obtaining a GAIA projection with a higher amount of information. By taking all the subjective information of the decision maker into account (weights of the criteria on top of the preferences), we obtain a representation of the problem that is more accurate and is subject to a smaller loss of data. In problems with a significant amount of criteria to be considered, this can drastically increase the amount of information correctly displayed. It goes without saying that by using this weighted GAIA projection as a starting point, the cartographic representation is also improved.

By using tools like these that speak to the user we should start seeing a trend to adapt other MCDA methods in a similar manner. This would ultimately lead to an integration of MCDA and GIS that would really benefit the user and set a goal for future developments.
Perspectives

In this section we will consider several tracks for further research in this domain. Among the many possibilities we will describe the following:

- Improvement of usability
- Development of new methodologies
- Taking Temporal Aspects into Account
- Extension of MCDA to applications that do not actively involve the decision maker

Improvement of usability

- One of the components with the highest added value for the user is the program’s interface and its usability. That is the part that can benefit the most from an integration and should attract most of our attention. Depending on the type of application and the intended use, we will have numerous possibilities for constructing an interface that will be appealing for the user and present all the necessary tools in an accessible way. With a little work and by exploring some of the possibilities, it should be possible to propose guidelines to help researchers understand how their methods and tools could interact with the rest of a geographic interface.

- MCDA offers support to the decision maker during the entire decision process and to take advantage of it, we should include those other tools as well. The first category of tools is the one for parameter elicitation [Dias et al., 2002, Mousseau et al., 2003, Eppe and De Smet, 2012]. These tools can come in several forms depending on the methodology used:
  - There are techniques that help set the weights for the different criteria by asking questions to the decision maker. Others require a limited amount of information and use it to deduce the rest.
  - Some methodologies even have tools to help set the inter-criteria preference information (see Section 1.3.2).

One possibility that has not yet been explored would be the elicitation of parameter values based on spatial information: instead of asking random questions to the user, the software could allow the user to compare alternatives that are visible on a map.

- The second category of tools comes in play once the results have been computed: sensitivity analysis tools [Saltelli et al., 1999]. Some of the following will introduce changes to the results that are displayed on the map:
– Sensitivity analysis tools can help predict the results depending on the values of the weights, while others allow the user to change the weights and observe the changes on the results dynamically.
– Finally some tools can help predict the results depending on the values of the weights or determine the robustness of the obtained solution by computing stability intervals.

As these elements deal with the results directly, they need to be available next to the geographic view. Furthermore, for the integration to be complete, the results displayed using glyphs on the map should also be affected by the changes that the user tries.

Development of New Methodologies

• A way to favor the integration of MCDA and GIS would be to develop new methods for specific spatial decision problems such as territory partitioning. This subject has already known a few developments like an article by Tavares-Pereira et al. [Tavares-Pereira et al., 2007a] and more recently a masters thesis by Harold Waterkeyn which focused on improving the method [Waterkeyn, 2013]. A possible development would be to apply a sorting or clustering method based on PROMETHEE to complete the process that was described by these works.

• Another spatial problem that deserves attention is the aggregation of spatial entities and what it implicates for the results characteristics. Evaluations for sets of alternatives are indeed not that easy to compute as they depend on the criteria considered.

• We have already mentioned rank reversal when describing multicriteria methods [Belton and Gear, 1985, Mareschal et al., 2008, Saaty and Vargas, 1984b]: they are the changes that occur in the ranking positions of some alternatives when changes are made to other seemingly unrelated alternatives for example. This phenomenon can also occur when the context or size of a problem changes. Indeed the rankings of alternatives in a region could be in conflict when a greater or smaller region is considered. When dealing with interfaces that can change the focus from one scale to another, we can expect such behaviors to happen and study them.

Taking Temporal Aspects into Account

• There have already been several works proposed to take temporal aspects into account within spatial databases [Du Mouza and Rigaux, 2005, Parent et al., 2006].
In the case of MCDA modeling however, there are hardly any works that consider the changes over time of evaluations or parameter values. A recently started thesis by Issam Banamar focuses on extending the PROMETHEE method to these aspects. In the case of spatial problems, one could also consider the changes of spatial characteristics and observe their impact on the results of evaluations.

Extension of MCDA to Applications that do not Actively Involve the Decision Maker

- Finally one last aspect which has not been of interest to date in the field of MCDA is the study of applications where the user is not expected to give his preference information knowingly. There are several examples in our daily lives where this happens: our navigation and search history in the Google search engine is used to present us with more adapted results for each of our newer queries, commercial websites like Amazon analyze our purchases to present us with new items that could arouse our interest, the Genius feature in the iTunes software analyzes the music we listen to frequently to provide us with other musical pieces corresponding to our tastes... This could be applied in the context of GIS by having the system learn the user’s preferences and use them for future suggestions. A simple example would be to develop a GPS system that learns the characteristics of the preferred routes that are taken, speed, areas visited before the destination, stops made, ...

In a world where technology is playing an even more important role, such developments would not be surprising.
Bibliography

[Dec, a] DECERNS SDSS. http://www.decerns.com


[Brans, 2004] Brans, J. (2004). The management of the future ethics in or: Respect, multi-
467.


GAIA decision support system for multicriteria decision aid. *Decision Support Systems*, 
12(4-5):297–310.


new family of outranking methods in multicriteria analysis. pages 477–490.


24(2):228–238.


sualization. *Human-Computer Interaction: Design Issues, Solutions, and Applications*, 
181.


[Harrah’s, 2000] Harrah’s (2000). *We know what our customers like*.


