

This is a PDF file of an unedited manuscript that has been accepted for publication in *International Transactions in Operational Research*. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain. The final version will be available at: <http://dx.doi.org/10.1111/j.1475-3995.2012.00874.x>

## Multi-Criteria Decision Analysis for Sustainable Data Centers Location

Miguel Trigueiros Covas<sup>a,b,e,\*</sup>, Carlos A. Silva<sup>b,e</sup>, Luis C. Dias<sup>c,d,e</sup>

<sup>a</sup> Portugal Telecom

<sup>b</sup> IDMEC, Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal

<sup>c</sup> Faculty of Economics, University of Coimbra, Portugal.

<sup>d</sup> INESC Coimbra, Coimbra, Portugal.

<sup>e</sup> MIT- Portugal Program, Sustainable Energy Systems

\*Corresponding address: Instituto Superior Técnico (IST), MIT Portugal Program, Av. Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal.

Email address: miguel.a.covas@telecom.pt; carlosaugusto.santossilva@mitportugal.org; lmcaldas@fe.uc.pt

### Abstract

Data Centers (DCs) are complex organizational and technical infrastructures that assure the performance and reliability of modern information and communication systems. The high installation and operations costs of DCs and the stringent requirements regarding reliability and safety require a close attention to the location of this type of facility. This paper proposes a Multi-Criteria Decision Analysis (MCDA) approach for identifying the most interesting locations to install sustainable DCs, taking into account technical, social, economic and environmental dimensions. For each of these main dimensions, the evaluation was formulated as a multi-criteria sorting problem. These problems were analyzed using the outranking MCDA method ELECTRE TRI through the IRIS software, allowing for uncertainty about the criteria weights. The results are summarized in graphical form, without attempting to reduce such incommensurable dimensions to a single value.

*Keywords:* Multiple Criteria Analysis; Environment; Location Problem, ELECTRE TRI; Data Centers

## 1. Introduction

The Information Technologies (IT) has a vital importance in today's economy. The increasing use of the internet, telecommunications services and internal IT networks in the organizations has led to a rapid growth of the number of computers over the world. Therefore, Data Centers (DCs) constitute a critical backbone of the worldwide IT infrastructure, as they are the infrastructures that operate mail servers' services, e-commerce services, data warehousing and search engines. Further, in the context of the new paradigm of collaborative networked society, DCs will also support the increasing use of cloud computing over the next coming years (Buyya et al., 2009) and thus it is expected that the number and size of this type of infrastructure will increase significantly over the next decade worldwide.

DCs are high energy intensive facilities due to the high power requirements of the computing equipment and the infrastructure needed to support it, in particular cooling systems (Fig.1). The average DC consumes as much energy as 25.000 households (Kapan et al., 2008). The impact of the increasing usage of DCs in the energy sector is significant, as shown by the fact that energy consumption for DC operation is doubling every five years in the U.S (EPA, 2007) and is expected to increase in Europe by 85% (E.U.C., 2008) over the next 10 years.

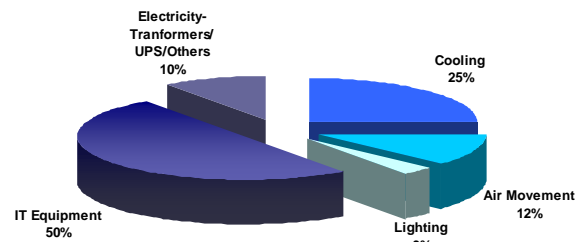


Fig.1. Energy Usage in a Data Center (Emerson, 2007)

Due to the modest cost of optical fiber networks, DCs can be located far from the ultimate users. The selection of DC locations is based therefore on a number of other factors. When a company decides to build a new DC, one of the main variables of the decision process is the energy cost, since it represents most of the DC's operational costs. But this is not the only important variable. The security of data as well as the reliability of the DC operation is closely related to its location, since the location influences for example the security of the facility, the probability of occurrence of natural disasters, the availability and reliability of electrical power and telecommunications networks, the transportation accessibilities, and labor availability. Only if all these aspects are taken into account, it is possible to guarantee the design, construction and operation of a DC that is environmentally as friendly as possible, with the lowest cost possible and the highest social and economic benefits: in conclusion a sustainable DC. Therefore, the location of a DC should follow a Multi-Criteria Decision Analysis (MCDA) approach (Belton and Stewart, 2002; Bouyssou et al., 2006) that considers different dimensions of sustainability.

The problem of locating a DC is of interest to the main Portuguese telecommunications operator, Portugal Telecom, who sought to identify the most interesting locations to install a sustainable DC. Two main managers were involved in this process: the principal IT director of the company and the head of the DC department. They were aided by the authors and other experts working for the company. In this text we will refer to this group of managers and experts as the Decision Makers Team (DMT).

In prior work, Covas et al. (2010) had established a methodology to find a subset of parishes in Portugal that are suitable for the construction of a sustainable DC, based on search analysis using geographic information systems. From a total of 4050 parishes that exist in Portugal mainland, the study highlighted 62 parishes as potentially good locations to install a DC. However, this set of alternatives was still deemed too large for a more detailed analysis. Indeed, since there are important costs associated with searching for a particular piece of land and negotiating with local authorities (e.g., for tax incentives), the identification of a smaller number of parishes that deserve a detailed study would be already an important step with high economic impact.

This paper proposes a tool using MCDA to analyze and support the decision of reducing the 62 pre-determined parishes to a smaller set of the most promising alternatives. The assessment is made at the parishes/region level and not in terms of a specific place or site. This paper describes how the set of criteria was developed, the reasons behind the choice of an appropriate MCDA method (ELECTRE TRI, a classification method proposed by Wei Yu under the supervision of Bernard Roy, described in (Yu, 1992)), and the application of the method using the IRIS software (Dias et al., 2002; Dias and Mousseau, 2003), which does not require setting specific values for the criteria weights. The final stage of this decision process consists of negotiations between the company and the local authorities in the parishes selected and is outside the scope of this paper.

Although the analysts in this problem (the authors) used a well-known MCDA method, an appropriate overall methodology had to be devised. The DMT was not willing to make judgments concerning trade-offs between different sustainability dimensions (economic, environmental, social and risk), and did not wish to commit to precise criteria weights even within each dimension. The analysts used therefore IRIS in an original way by evaluating separately different sustainability dimensions, using a common qualitative scale. This step was inspired by Neves et al. (2008) approach to account multiple stakeholders. Each parish was subsequently evaluated to assess the best and worst level it could attain with and without considering the elicited opinions about the relative importance of the criteria. Therefore, by presenting this application, this paper intends to contribute not only to the literature on locating a DC, but also to the literature of sustainability assessment, addressing the particular case where no decision maker (or group of stakeholders) provides judgments about the importance of the different sustainability dimensions.

This paper proceeds as follows. In the next section, we provide a brief introduction to DC site selection, the criteria identification and a detailed criteria characterization. Section 3 briefly presents the multi-criteria method chosen to provide the decision support. Section 4 shows the results of applying the proposed methodology for identifying interesting regions for the development of Sustainable DC. Finally, section 5 draws the main conclusions of this study.

## **2. Site Selection for the Development of Sustainable Data Centers**

The site selection is a fundamental step in the design of a DC. A poor site choice may affect the sustainability of the facility during the complete lifetime of the project, since there is no flexibility to change or modify the site once it is selected.

There are many variables involved in the problem of finding the best areas to locate a DC. Sun Microsystems (Snevely, 2001) considers that site location should take into consideration natural hazards (e.g. seismic activity, hurricanes), man-made hazards (e.g. industrial pollution, vibration), and the availability of utilities services (e.g. power supply, telecommunications) or even emergency services and vehicle access. The Code of Conduct on Data Centers published by the European Commission (E.U.C., 2010) established some criteria for DC location to promote the energy efficiency of the facility (e.g. locations where waste heat of the DC can be reused, locations with a source of free cooling). In addition to the mentioned criteria the Uptime Institute (TurnerIV et al., 2008) also considers economic criteria like utility consumption rates, financial incentives (property and sales taxes, rebates, reduced permit fees), etc. These are the main criteria that should to be assessed during the decision process of building a DC, which are followed by the largest IT companies in the world, like Cisco (Alger, 2005), Google (2010), HP (2008; Middleton, 2009), IBM (2009), and Microsoft (Dwyer, 2009).

Most of the literature on location problems of critical infrastructures uses optimization models in which the Euclidean distances, or other types of distances over a network, between the facilities to be located and the customers of the facility, the suppliers of the facility, or those negatively affected by the facility (e.g. nuclear power plants) are the most important aspects of the problem (e.g., see (Eiselt and Marianov, 2011)). However, in the case of locating a DC, the distance-related aspects of the problem are not the most important and coexist with many other site-specific characteristics. Moreover, the aim of this study was to identify areas that were highly adequate to locate a DC at the parish level, rather than choosing the best one, hence an optimization framework (even if multi-objective optimization is considered) was not the most adequate. For these reasons, an MCDA framework for evaluating a list of predefined alternatives was preferred.

The literature presents several examples of the use of MCDA to evaluate a list of potential locations. For instance, Merkhofer and Keeney (1987) used multi attribute utility analysis to evaluate 3 alternative sites for the disposal of nuclear waste. Siddiqui et al. (1996) used the Analytic Hierarchy Process (AHP) to identify and rank potential landfill areas for sitting municipal solid wastes, taking into account regulatory restrictions, area attributes, and site assessment criteria. Barda et al. (1990) used the ELECTRE III method for the location of thermal power plants, and Pavic (1991) used the PROMETHEE method to choose the location of a production system, taking into account criteria such as building costs, infrastructure costs, labor costs, weather conditions among others criteria.

Despite the list of criteria previously described and the references on the use of MCDA methods in facility location, it was not found in the literature a methodology for DC location that considers simultaneously all the criteria used by the industry. The available literature is in fact very scarce and the authors could find only two works that answer this question specifically. Abbasov et al. (2009) presented an optimal location model to locate DCs taking into account the risks (natural, political and economic risks) while minimizing two objectives (cost and risk). Chang et al. (2007) presented a model, based on the classical capacitated p-median problems, to optimize the locations of the US army DCs, such that the total data demand-weighted distance traveled is minimized and the IT load on any DC is small as possible.

These papers did not consider however that the location criteria should be included in a holistic approach for the development of a sustainable DC. Following the sustainability definition proposed by United Nations in the Brundtland Report (WCED, 1987), which states that “*Sustainable development (...) meets the needs of the present without compromising the ability of future generations to meet their own needs*”, for Portugal Telecom (2010), sustainability means to increase the shareholders return, to improve the customers satisfaction, increase the welfare of the employees and family, as well as, to contribute to the reduction of the greenhouse gases. This definition is also aligned with the broader corporate sustainability definition proposed in the Global Information Technology Report (Dutta and Mia, 2010) as “*increasing short and long-term profitability by holistically managing social, environmental, and economic risks and opportunities*”.

Independently of which sustainability concept is chosen (more examples can be seen in Hall (2006)), there seems to be a consensus that sustainable development is a multidimensional concept, and its promotion involves the integration of environmental, social and economic issues in the decision making processes.

## 2.1. Site Selection Criteria

Criteria are tools for assessing alternatives, reflecting all the concerns and relevant issues to the decision problem. We established four main evaluation dimensions for the site selection of a sustainable DC suggested by the literature, with agreement of the DMT: Risk, Social, Environmental and Economic (Fig.2). The tools of value-focused thinking proposed by Keeney (1992) were then used to guide the dialogue with the DMT and structuring their inputs at this stage.



Fig.2. Main Criteria – Site Selection for a Sustainable Data Center

Each main criterion comprehends several subcriteria according to the hierarchies of concerns described in Figs.3 to 6. These hierarchies were built as a structuring aid for the purpose of obtaining a list of criteria. This provided the DMT an easy means to visualize and organize their concerns facilitating the detection of potential omissions. The output of this phase of the process was a simple list of elementary criteria, as an evaluation of the hierarchical type (at all the nodes of a hierarchy) was not considered necessary. In the next section the criteria are presented with more detail.

It is noteworthy that some of the criteria are important for assessing in more than one main dimension. For example, energy savings are relevant both from an environmental perspective (less harm to the environment) and from an economic perspective (savings on operational costs). If one is required to perform an environmental assessment of the locations, energy savings are an important part of it. If one is required to perform an economic assessment of the locations, energy savings are again important. This means that, as one would expect, a location allowing significant energy savings will be in a better position to have a good evaluation under the environmental dimension as well as a good evaluation under the environmental dimension. The four dimensions will not be aggregated, thus avoiding any double-counting effects.

Let us also note also that the same concerns can be transformed into a criterion in different ways depending on the dimension being evaluated. Under the environmental evaluation, energy savings are measured in terms of CO2 (or other emissions); under the economic evaluation, energy savings are measured in monetary units.

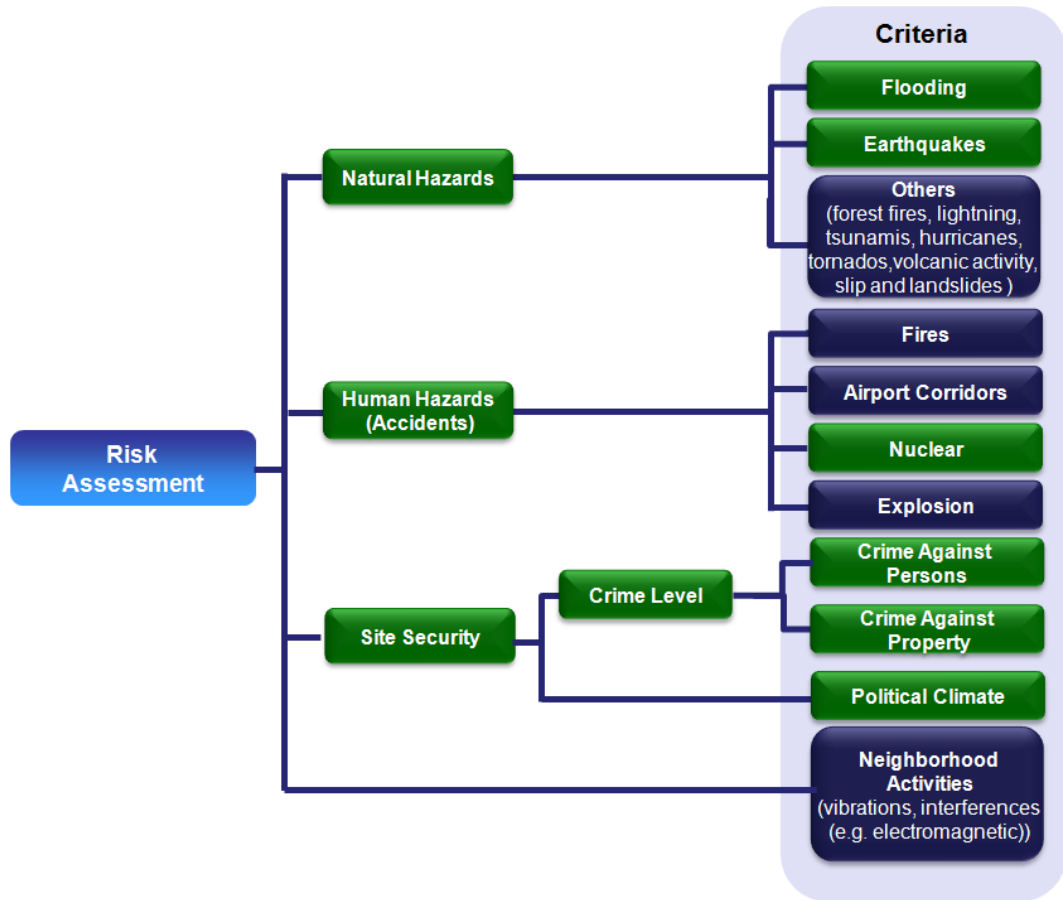


Fig.3. Risk Assessment Criteria

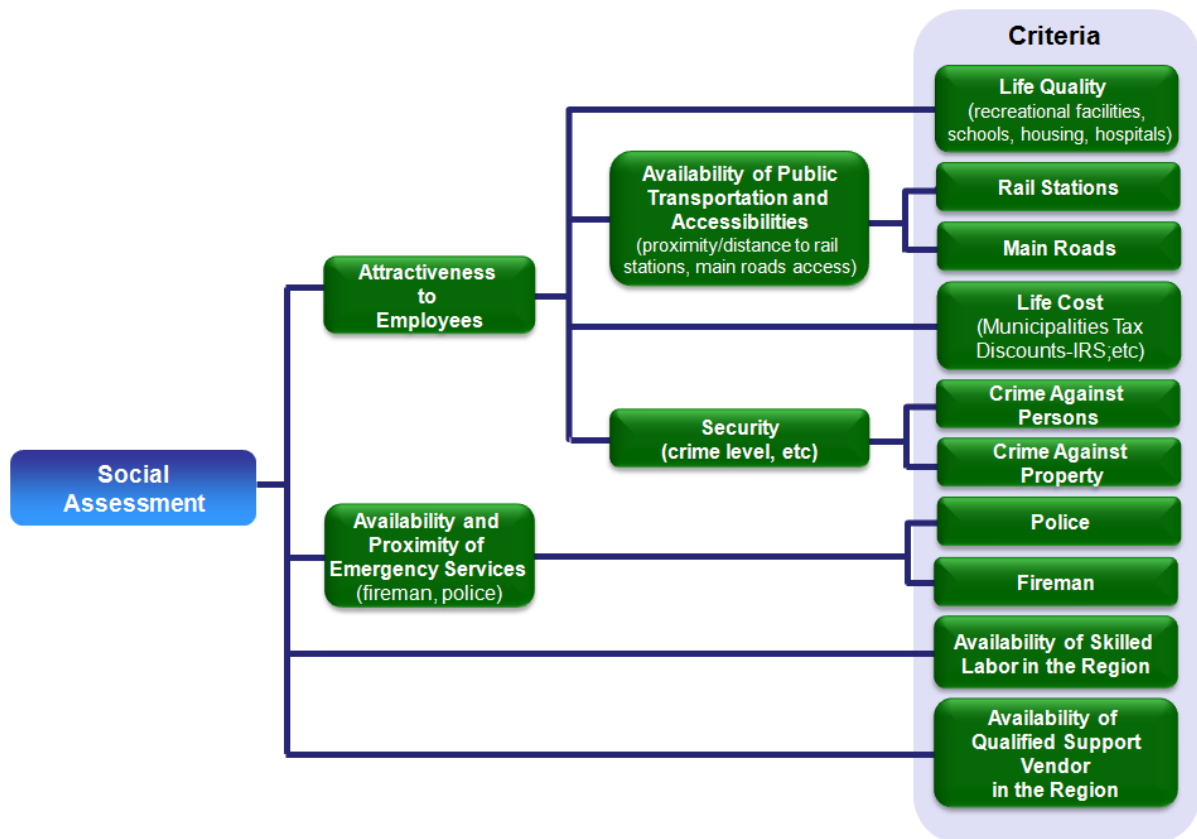


Fig.4. Social Assessment Criteria

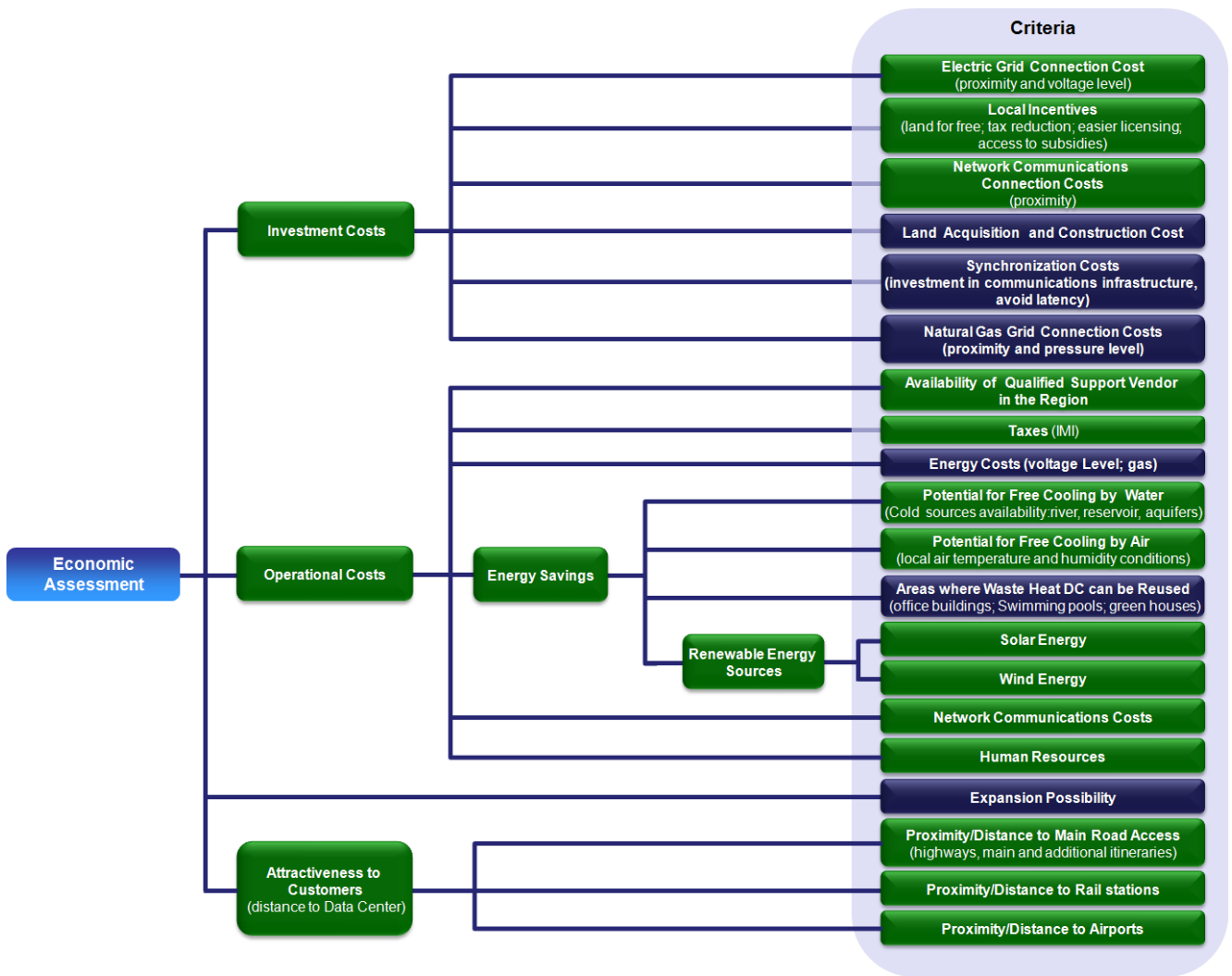


Fig.5. Economic Assessment Criteria

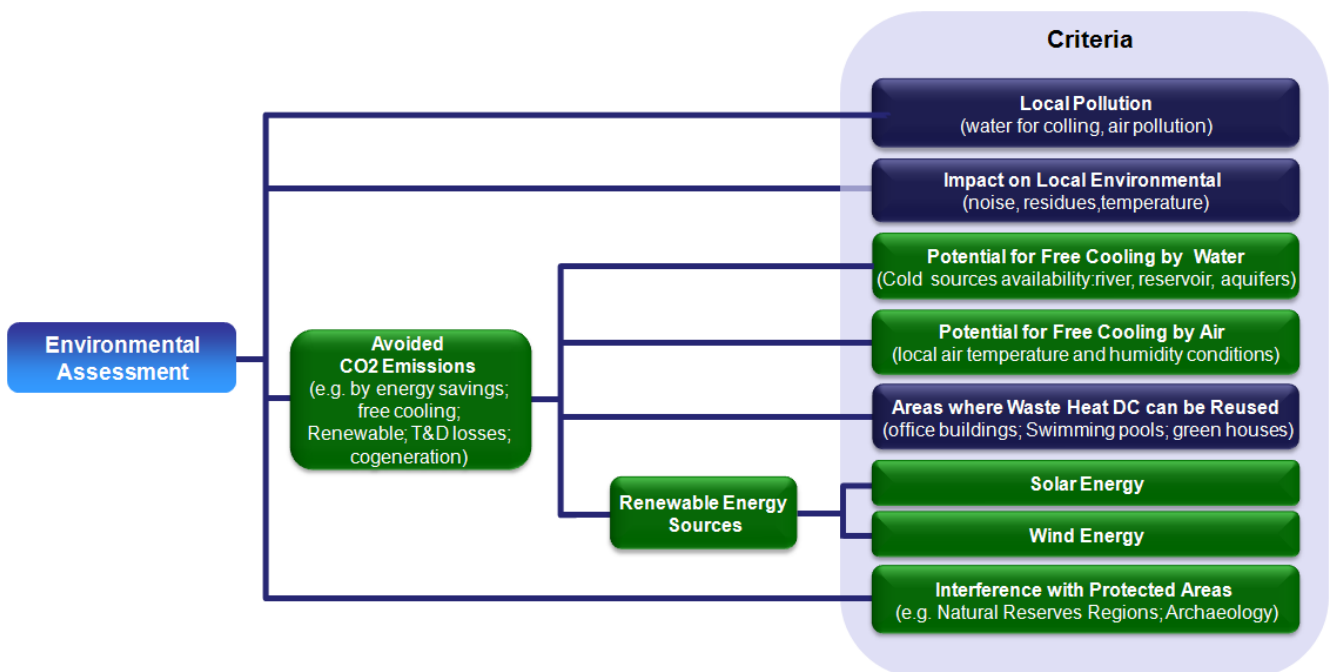


Fig.6. Environmental Assessment Criteria



## 2.2. Criteria Evaluation

In this work, the analysis of the sites is made at the parishes' level in Portugal, which means we will assess only the criteria that apply to this level of research, and we will consider only criteria with that are relevant in the Portuguese territory. The criteria painted with a darker shade in Figs.3 to 6 were not included in the analysis since they cannot be assessed at the parishes' level (the evaluation of such property-related attributes is possible only after identifying a precise piece of land). However, they could be taken into consideration in studies with very high spatial resolutions.

### 2.2.1 Risk Assessment Criteria

The risk assessment criteria define the requirements that are essential to maintaining the integrity, safety, and availability of the DC.

- a) **Natural Hazards Criteria:** Identifying areas with natural threats can help to form an effective risk avoidance strategy:

Flooding: Boolean criterion. Floods can destroy buildings and equipments, cause water damage to structure and contents, result in power failures, damage roadways, and cause human injury or death. Parishes within or intersected by a zone of flooding risk assume the value 0, otherwise assume 1.

Earthquakes: Earthquakes can destroy buildings and equipments or cause power failures, damage roadways, human injury or death. To evaluate this criterion we used the Modified Mercalli Scale (Vulcanologia, 2010), which defines intensity levels from 1 to 12.

- b) **Human Hazards (Accidents) Criteria:** Identifying areas with potential damage caused by human accidents:

Nuclear: Boolean criterion. The release of radioactivity due to a nuclear accident can cause the contamination of water and air and thereby stop the DC operation. To evaluate this criterion we can use the security perimeter for a nuclear accident established by the ANPC (2010) and the IAEA (International Atomic Energy Agency) Safety Standards (2007). In our particular study all parishes are outside evacuation zones (more than 30 Km from a nuclear power plant) as there are no such plants in Portugal. Hence this attribute was also discarded.

- c) **Site Security Criteria:** The security of the region of the DC is important. Areas subject to physical attacks, people threats, robbery and burglary, can put the DC and people related to it (e.g. workers, customers) in danger.

Crime Level: Crimes recorded by the police in the municipality (a municipality comprises several parishes), namely the number of crimes against persons per 1.000 habitants and crimes against property per 1.000 habitants provided by the Portuguese National Institute of Statistics (INE, 2009a, 2009b, 2009c).

### 2.2.2 Social Assessment Criteria

The social criteria are concerned with the ways in which people (services) are essential to run a DC.

- a) **Attractiveness to Employees Criteria:** Attractive aspects to host people in the region:

Life Quality: Recreational facilities, schools, housing, hospitals, among others subcriteria, at the municipality level. To evaluate this criterion we used a study about quality of life in Portugal (Manso and Simões, 2009) that considers these type of subcriteria and provides a quality life index which ranges from 36.72 to 111.63.

Availability of Public Transportation and Accessibilities: The availability and proximity (kilometers) of public transportation (e.g. train), and proximity (kilometers) to main roads access within or near the parish.

Life Cost: To evaluate this criterion we used the average value of the urban properties in Euros transacted in the municipality (INE, 2009a, 2009b, 2009c).

Security: This is the same criterion described in section 2.2.1 c) (Site Security Criteria), which can be also evaluated from an employee’s point of view.

- b) **Availability and Proximity of Emergency Services Criteria**: Proximity (distance in kilometers) to police and fireman stations within or near the parish.
- c) **Availability of Skilled Labor in the Region**: Number of students in high school technology courses by municipality (INE, 2009a, 2009b, 2009c).
- d) **Availability of Qualified Support Vendor in the Region**: Qualitative assessment made by the DMT, see Table 1.

Level	Description (Availability of Qualified Support Vendor)
1	Unsatisfactory
2	Reasonable
3	Good
4	Very Good

Table 1. Availability of Qualified Support Vendor Criterion Evaluation

### 2.2.3 Economic Assessment Criteria

Costs and benefits that a DC may have if it is located in a certain region:

- a) **Investment Costs**: Costs and benefits with impact on investment costs, related/dependent of region, that can be applied in order to develop a DC:

Electric Grid Connection Cost: Distance (kilometers) to the electrical grid (High voltage/Medium Voltage power substation) within or near the parish.

Local Incentives (e.g. land for free; tax reduction; easier licensing; access to subsidies etc): Boolean criterion: regions (at the municipality level) that are eligible for the European Cohesion Police co-funding (NSRF-National Strategic Reference Framework) get value 1, whereas other regions get value 0.

Network Communications Connection Cost: The communication cost depends on the distance (kilometers) to the telecommunications network. For distances over 4 Km (ANACOM, 2010) the use of DSL technologies is very difficult to ensure broadband.

Land Acquisition and Construction Cost: Not considered in the study since it can vary a lot within the same parish.

Synchronization Cost: Synchronization cost to connect two or more DCs. Not considered in the study (the DC under consideration will operate as a stand alone DC).

Natural Gas Grid Connection Cost: Not considered in the study, because Combined Heat and Power (CHP) possible solutions are not economically viable when it is possible to use free cooling solutions over 7.500 hours/year (Covas et al., 2010).

b) **Operational Costs**: operational costs related/dependent of the DC location:

Availability of Qualified Support Vendor in the Region: the same as described in the section 2.2.2. d) Inexistence of local support increases support costs.

Taxes (IMI): Annual property municipality tax.

Energy Costs (due to the choice of voltage level): Not considered in the study because only connection to high voltage was considered. A higher voltage level connection represents lower costs of energy transport and delivery.

Energy Savings: Energy strategies that can reduce operational costs:

- Potential for Free Cooling by Water: In regions where is not possible to achieve 100% of air free cooling throughout the year, the existence of water resources is also very important. The ample availability of water for adiabatic cooling can increase the air free cooling potential. Furthermore, the chillers cooling systems energy efficiency depends on the temperature of the cold source (the part of the system that extracts heat from the DC). This criterion is assessed by the aquifer productivity (m<sup>3</sup>/day km<sup>2</sup>) in the region.
- Potential for Free Cooling by Air: Number of hours per year with average air temperature below 21°C. According to the ASHRAE (2008) and the Best Practices for the EU Code Of Conduct On Data Centers (E.U.C., 2010) recommendations, a DC should be designed to operate at "inlet" temperatures between 18 °C and 27 °C, to maximize energy efficiency. Following these recommendations and considering the usual temperature gains in using free cooling technology, we established as an evaluation criteria the number of hours per year where average air temperature is equal or below 21°C as an indication of number of hours per year that the DC could operate without conventional cooling systems.
- Renewable Energy Sources: Potential of renewable energy sources: solar and wind. Regarding the solar source, the potential is measured by the sunshine hours per year (Ambiente, 1975). For the wind source, the potential is measured by the expected maximum number of hours per year at a nominal power (Esteves, 2004).

Network Communications Cost: The communications circuits fees also depends on the distance to the network.

Human Resources: The average base salary of a technical engineer in Euros (Social, 2010) was considered as a proxy for staff salary costs.

- c) **Attractiveness to Customers (distance to the DC)** : Convenience of the DC location to customers, namely considering the offer of hosting services to outside clients :

Proximity/Distance to Main Road Access (Highways; Main and Additional Itineraries): Please see the description on section 2.2.2 a) (Attractiveness to Employees Criteria)

Proximity/Distance to Rail Stations: Please see the description on section 2.2.2 a) (Attractiveness to Employees Criteria).

Proximity/Distance to Airports: Distance to airports in km.

#### 2.2.4 Environmental Assessment Criteria

The environmental criteria are concerned with the interactions between the DC and the environment:

- a) **Avoided CO2 Emissions by Energy Savings (free cooling, T&D losses and cogeneration)**: Energy efficiency measures that can reduce/avoid CO2 emissions:

Potential for Free Cooling by Water, Potential for Free Cooling by Air, and Renewable Energy Sources: Please see the description on section 2.2.3. b) (*Operational Costs\Energy Savings*).

Areas where Waste Heat of DC can be Reused: Not considered in the study. Difficult to assess at the parishes level, this criterion is specific/appropriate to be applied at the installation site (field) of the DC.

- b) **Interference with Protected Areas**: Boolean criterion. Reserves and natural parks must be avoided to locate the DC. Parishes within or intersected by reserves or natural parks assume the value 0, otherwise assume 1.

### 3. Parishes Assessment Using MCDA for the Development of a Sustainable DC

MCDA approaches (Belton and Stewart, 2002; Bouyssou et al., 2006) acknowledge the existence of multiple dimensions for evaluating alternatives without attempting to reduce them to a monetary scale. There is not a single criterion to optimize: MCDA allows decision makers to recognize that choices reflect a compromise among different conflicting objectives.

The problem of selecting the highly adequate parishes for building a sustainable DC presents a few important specificities. First, it would be difficult to ask the DMT to quantify trade-offs among criteria (how much exactly would they sacrifice on one criterion to obtain a given improvement on another criterion). It was also deemed that it should not be allowed that very good performances on one criterion would be able to compensate very poor performances in another criterion. This suggested the use of a non-compensatory method, namely an outranking method from the ELECTRE family (Roy, 1991) (for a description of ELECTRE and other outranking methods such as PROMETHEE, see (Figueira et al., 2005)).

In addition, the level of analysis and the nature of the data imply that the parishes cannot be assessed in a very precise way. Contrary to a specific site, when evaluating a region is not possible to accurately evaluate distances to the main infrastructures, or the expansion possibility of the site, among other factors. This suggests to the analysts that they should be parsimonious in what they require the evaluation model to do. Namely, building a model that would assign a precise value or utility to each parish, or that would rank the parishes, would provide to the DMT an erroneous illusion of precision that the analysts wish to avoid. Taking this

concern into account, it was proposed to use a method that would partition of the set of alternatives into several categories defined with respect to these norms (called by Roy (1996) the sorting problem formulation). There are several MCDA sorting methods, most of them surveyed by Doumpos and Zopounidis (2002) and Zopounidis and Doumpos (2002). For later examples we can cite, inter alia, Al-Obeidat and Belacel (2011), Fernandez et al. (2009), Figueira et al. (2004), or Rocha and Dias (2008).

The method applied is ELECTRE TRI (Yu, 1992) which meets adequately all the specificities of this case: it is used for sorting problems, it is a non-compensatory method, it allows an evaluation in absolute terms (i.e., alternatives are not compared against each other but to predefined norms), it avoids asking for trade-offs, and does not require converting the performance criteria into a uniform scale range (it allows the inclusion of criteria measured in different units and even measured in qualitative terms)..

A version of the ELECTRE TRI method was applied using the decision support software IRIS (v2.0) (Dias et al., 2002; Dias and Mousseau, 2003), which is designed to address the problem of assigning a set of alternatives to predefined ordered categories, according to their evaluations (performances) at multiple criteria. Among other characteristics, the software accepts uncertainty regarding the criterion weights and the cutting level, through the definition of intervals for each parameter, or the definition of linear constraints (e.g., stating that one criterion has more weight than another one). Thus, decision makers do not have to agree or to commit to specific values for the criteria weights.

In this work, the assessment of the regions using ELECTRE TRI method was carried out separately for each one of different dimensions identified in Section 2.1 (Risk, Social, Economic and Environmental), without aggregating the results obtained for these dimensions.

### 3.1. ELECTRE TRI Method - Overview

The ELECTRE-TRI method (Yu, 1992) belongs to the ELECTRE family of multi-criteria methods developed by Bernard Roy and his co-workers (Roy, 1991, 1996). This method was specifically designed for multi-criteria sorting problems, i.e., to assign each alternative to one of a set of predefined ordered categories according to a set of evaluation criteria. Let us introduce the following notation:

- $m$  denotes the number of alternatives;
- $n$  denotes the number of criteria;
- $t$  denotes the number of categories;
- $A = \{a_1, \dots, a_m\}$  denotes the set of alternatives;
- $G = \{g_1(\cdot), \dots, g_n(\cdot)\}$  denotes the set of criteria (valued functions on  $A$ );
- $C = \{C_1, \dots, C_t\}$  denotes the set of categories by preference order, in which  $C_1$  is the least preferred (worst category) and  $C_t$  is the most preferred (best category);
- $B = \{b_0, \dots, b_t\}$  denotes the set of reference profiles that separate the consecutive categories, where  $b_0$  is the worst profile and  $b_t$  is the best profile.

Each category  $C_h$  ( $h=1 \dots t$ ) is limited by two profiles:  $b_h$  is its upper limit and  $b_{h-1}$  is its lower limit.

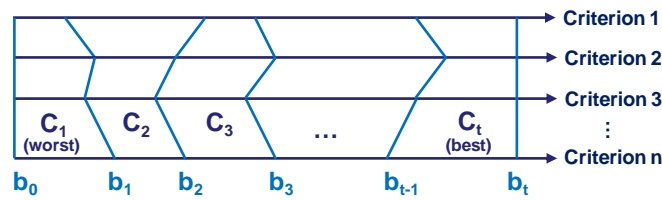


Fig.7. Definition of categories using profiles

ELECTRE TRI has two possible procedures to assign alternatives to categories: the pessimistic procedure and the optimistic procedure. In this work the more demanding pessimistic variant was used, which is the one implemented in IRIS (the name “pessimistic” comes from the fact that it places the alternatives in a category that is at most as good as the result of the optimistic procedure).

The assignment of alternatives to categories is based on the concept of outranking relation (a binary relation meaning “not worse than”). An alternative  $a_i \in A$  is said to outrank a profile  $b_h \in B$  (denoted by  $a_i S b_h$ ) if it can be considered that  $a_i$  is at least as good as  $b_h$ , given the performances of  $a_i$  and  $b_h$  at  $n$  criteria. The credibility of  $a_i S b_h$ , denoted by  $s(a_i, b_h)$ , is computed taking into account the weights of the criteria that are in favor of  $a_i$  (considering possibly indifference and preference thresholds) as well as the possibility of veto from a criterion in which  $a_i$  is much worse than  $b_h$  (considering veto thresholds). The outranking relation is decided by comparing  $s(a_i, b_h)$  with a cutting level  $\lambda$  defined by the decision maker(s), which defines the “majority requirement”. The assignment of each alternative to one category is done by comparing its performance in each criterion to the reference profiles. The pessimistic procedure assigns each alternative to the highest category such that its lower bound is outranked by  $a$ . For details about ELECTRE TRI see (Yu, 1992; Dias and Mousseau, 2003).

### 3.2. Evaluation Model Parameters

The use of ELECTRE TRI requires a set of parameters that represent the preferences of the decision makers. In this case, the analysts used the inputs of the DMT.. The DMT intervened at various stages of the parameterization process (definition of the categories, criteria thresholds, and constraints of criteria weights). The interaction with the DMT was conducted through one-to-one meetings and small workshops.

As explained in section 2.1, some criteria are common to several dimensions (e.g. the Crime Against Persons criterion exists in the Risk and Social dimensions, meaning that a low level of crime will have a positive impact on the evaluations under these two dimensions). Let us emphasize that these criteria are not “double counted”, because the four dimensions are analyzed independently and will not be aggregated into a synthesis value.

With this approach of making separate assessments of parishes under each of the four dimensions in analysis, the weighting of each dimension is avoided. The group agreed that due to the imprecise nature of the data and for the study’s purposes it would suffice to sort the alternatives into four categories (levels): Weak, Good, Very Good and Excellent (recall that the set of 62 alternatives already resulted from a GIS-based analysis used to exclude poorer options (Covas et al., 2010).

#### 3.2.1 Risk Assessment

### 3.2.1.1 Range Value Limits

The Risk Assessment criteria take into account the maintenance of the integrity, safety, and availability of the DC. Based on the criteria description presented in section 2.2.1, the categories range values for each criterion were established. Four categories were proposed by the analysts and approved by the DMT, where C1 is the worst, described as *Very High*, and C4 the best category described as *Low*, see Fig.8. Table 2 indicates the values considered for the reference profiles.

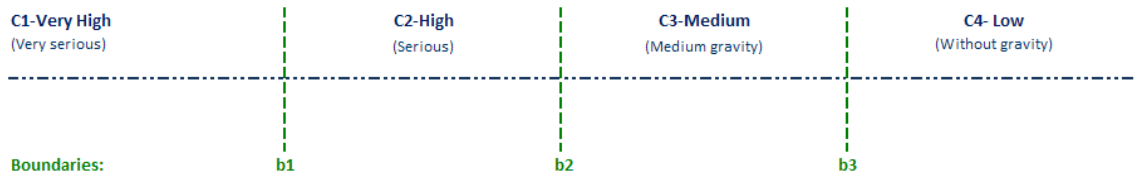


Fig.8. Relation between the Risk Categories and the category profiles

Risk Categories	Flooding Level	Earthquake Level	No.Crimes Against Persons/1.000 Municipality Habitants	No.Crimes Against Property/1.000 Municipality Habitants
Low	1	<= 5	< 5	<15
Medium	0	6-7	5-10	15-30
High	0	8	10-20	30-50
Very High	0	>=9	>20	>50

Table 2. Risk Categories Boundaries

Regarding the values indicated in Table 2, it should be noted that:

- The Mercalli Scale has a range between 1 and 12 but the seismic intensity in Portugal ranges between 5 and 10. Considering this information, the DMT agreed to define the following boundaries  $b_1=8$ ,  $b_2=7$ ,  $b_3=5$ .
- The number of crimes against persons per 1.000 habitants recorded in the municipalities where the parishes are located reaches a maximum of 17. The DMT set  $b_1=20$ ;  $b_2=10$ ;  $b_3=5$  (Table 2, column 4), considering that Portugal has no regions with very high risk level according to this subcriterion.
- Analogously, the DMT set  $b_1=50$  for the subcriterion Crime Against Property, although the maximum number observed in the studied parishes was 33 (and the maximum number recorded in Portugal was 46). Therefore, no parish will be considered of very high risk level according to this subcriterion.

### 3.2.1.2 IRIS Parameters Set

The next step consisted in defining the indifference ( $q_j(b_h)$ ) and preference ( $p_j(b_h)$ ) threshold values for each category, as well as, the veto ( $v_j(b_h)$ ) conditions. The thresholds  $q_j(b_h)$  and  $p_j(b_h)$  intervene when checking if a criterion agrees with an outranking. A criterion agrees fully with the outranking if the alternative is not worse than the profile by a difference larger than  $q_j(b_h)$ ; it agrees partially if this difference lies between  $q_j(b_h)$  and  $p_j(b_h)$ ; it does not agree if the difference is  $p_j(b_h)$  or higher. In the latter case, it will oppose a partial veto if the difference lies between  $p_j(b_h)$  and  $v_j(b_h)$ , or it will oppose a complete veto if the difference is  $v_j(b_h)$  or higher. A veto means that the outranking is not accepted even if all other criteria support it. Fig.9 displays the IRIS screen in which the thresholds are entered; these thresholds can be set independently for each category.

	Flooding	Earthquakes	Crime Person	Crime Property
gl(b1)	0	8	20	50
q1	0	0	2	5
p1	0	0	5	10
v1				
gl(b2)	0	7	10	30
q2	0	0	1	3
p2	0	0	5	10
v2		1		
gl(b3)	1	5	5	15
q3	0	0	1	2
p3	0	0	2.5	7.5
v3	0.1	2		
MAX/min?(1/-1)	1	1	1	1

Fig.9. Risk Assessment - IRIS Categories Boundaries Layout

The DMT agreed to define two veto thresholds for two criteria. They established that an alternative with a level of 0 in Flooding cannot outrank  $b_3$  (i.e., cannot reach category C4) by setting  $v_{\text{Flooding}}(b_3)=0.1$ . Similarly, by setting  $v_{\text{Earthquakes}}(b_3)=2$  they prevent an alternative from being classified as low risk (C4) if it has a seismic risk of 7 ( $=5+2$ ) or higher. For crime against property and crime against persons, the DMT defined an indifference threshold of 10 % of the boundary value, and defined a preference threshold equal to half of the difference between adjacent categories. Regarding the constraints, the DMT was asked what were the most important ones. The DMT agreed that the flooding ( $k_1$ ) and earthquake ( $k_2$ ) criteria are equally important ( $k_1=k_2$ ). It was also established that these criteria are more important than the crime against persons ( $k_3$ ) and property criteria ( $k_4$ ). The crime against persons ( $k_3$ ) and property ( $k_4$ ) were defined as having equal importance. The IRIS software does not accept strict inequalities (e.g.  $k_2>k_3$ ), therefore these were included considering a minimum difference of 0.05 (when the DMT establishes that  $k_2>k_3$ , the constraint inserted in IRIS is  $k_2-k_3\geq 0.05$ , see Fig. 10).

	lambda	k1	k2	k3	k4	<,>=	RHS
Norm		1	1	1	1	=	1
ad1	0	1	-1			=	0
ad2			1	-1		>=	0.05
ad3			1		-1	>=	0.05
ad4				1	-1	=	0

Fig.10. IRIS - Table of constraints (relation between risk criteria).  $k_1$ =Flooding;  $k_2$ =Earthquakes;  $k_3$ =Crimes Against Persons;  $k_4$ =Crimes Against Property.

### 3.2.1.3 IRIS Results

This section presents the IRIS results for the Risk Assessment criteria, considering that the weights are free to vary (Fig. 11a), Without Constraints), and also considering that the constraints on the weights placed by the DMT should be respected (Fig. 11a), With Constraints). A third scenario was also included where a central weights vector inferred from the imposed constraints (for details see (Dias and Mousseau, 2003; Mousseau and Dias, 2004)) is considered (Fig. 11a), Central Weights). When there is some freedom in the weight values, IRIS



uses an optimization approach to determine the set of categories in which an alternative can be sorted. For instance, the first alternative (Cheleiros) could be placed in categories C1 or C2, depending on the weights chosen (without violating the imposed constraints), whereas the last alternative (Ganfei) must be always placed into C3. The alternative Monserrate (fourth from the bottom) can be placed into categories C3 or C4 if the weights are free to vary, but only into C3 if the weight constraints are to be respected. We can see that the application of the constraints does not limit too much the range of results – only 13 parishes out of the 62 reduce the number of possible categories from 2 to 1. The application of the central weights led to a specific assignment for each category. With this approach no parish reaches C4 (*Low Risk*), 16 parishes are assigned to C1, 4 parishes assigned to C2 and the majority (42 parishes) are placed to C3.

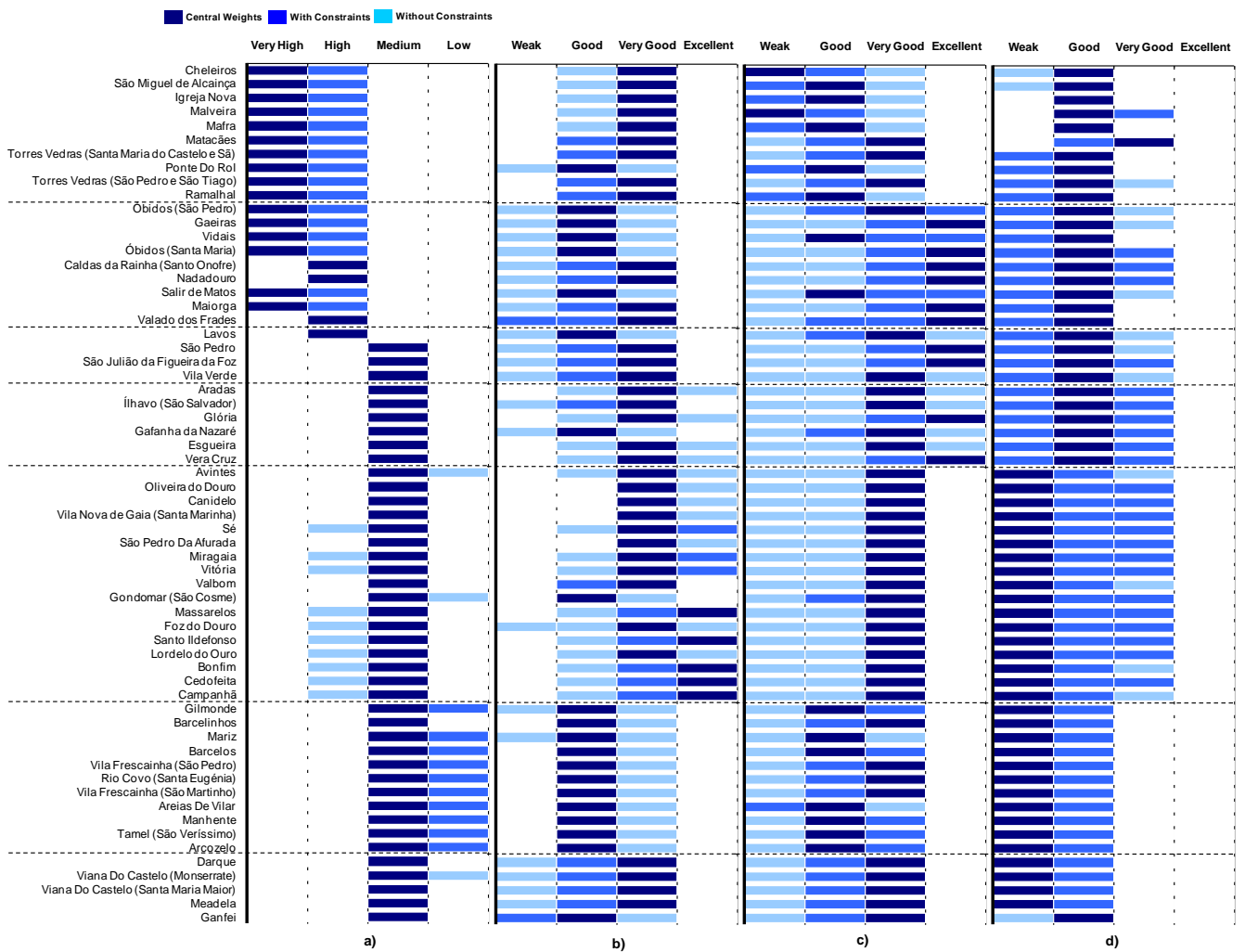


Fig.11. IRIS Results Without Constraints, With Constraints and Central Weights: a) Risk Assessment; b) Social Assessment; c) Economic Assessment; d) Environmental Assessment

### 3.2.2 Social Assessment

#### 3.2.2.1 Range Value Limits

The Social Assessment criteria describe the features that are important for people working on a DC and services provided. Taking into account the criteria descriptions in section 2.2.2, four categories were established

(Fig.12). For each social criterion, the range of values corresponding to each category were defined taking into account the DMT's inputs (Table 3), leading to the  $g_j(b_h)$  values depicted in Table 4.

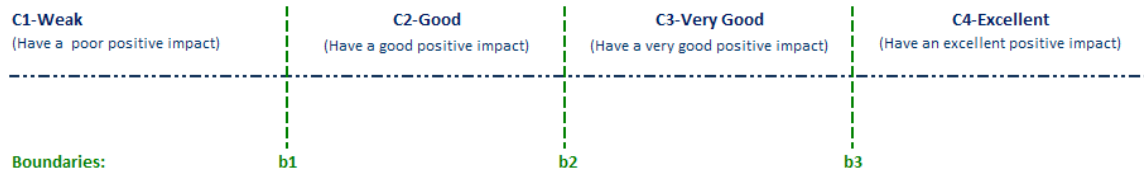


Fig.12. Relation between the Social Categories and the category profiles

Social Categories	Life Quality	Rail Station	Main Road Access	Life Cost	No.Crimes Against Persons	No.Crimes Against Property	Police Station	Fireman Station	Skilled Labor	Support Vendor
Weak	<56,6	4	>10	>250 k€	>20	>50	>5	>5	<100	1
Good	56,6-67,93	3	5-10	175-250 k€	10-20	30-50	3-5	3-5	100-400	2
Very Good	67,93-80,38	2	2-5	100-175 k€	5-10	15-30	1-2	1-2	400-600	3
Excellent	>= 80,38	1	<=1	<=100 k€	< 5	<15	0	0	>=600	4

Table 3. Social Categories Boundaries

Regarding the values indicated in Table 3, it should be noted that:

- The life quality criterion bound  $b_3$  corresponds to the 30<sup>th</sup> best municipality in Portugal,  $b_2$  corresponds to the 100<sup>th</sup> best municipality, and  $b_1$  corresponds to the 200<sup>th</sup> best municipality according to (Manso and Simões, 2009).
- For the rail station criterion, all the parishes in analysis are less than 4 Km from a rail station, and therefore none will be considered Weak in this subcriterion.
- For the main road access criterion, all the parishes in analysis are less than 10 Km from a main road node access, and therefore none will be considered Weak in this subcriterion.
- For the life cost criterion, in Portugal the average value of the urban properties transacted at the municipality level is € 126.692. Values for the parishes in analysis range from €70.000 to € 220.000.
- Concerning the emergency services (police and fireman), all parishes in analysis are less than 5 Km away from both, and therefore none will be considered Weak in these subcriteria.
- Concerning the Availability of Skilled Labor in the Region, the number of students in high school technology courses ranges from 0 to 606 in the municipalities considered.
- The values indicated for the Vender Support criterion are those defined in Table 1.

### 3.2.2.2 IRIS Parameters Set

The indifference and preference threshold values for each category, as well as the veto thresholds are depicted in Table 4. The veto threshold for Life Quality reflects the DMT's opinion that one region should be at least Good in Life Quality in order to be classified as Excellent ( $g_1(b_3)-v_1(b_3) = 80.38-23.78 = 56.6 = g_1(b_1)$ ). Similar opinions defined the veto thresholds for Skilled Labor and Vendor Support. Indifference and preference thresholds were also set in a way similar to that described in section 3.2.1.2

	Life Quality	Rail Station	Main Road Access	Life Cost	No.Crimes Against Persons	No.Crimes Against Property	Police Station	Fireman Station	Skilled Labor	Support Vendor
IRIS Variables	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10
g (b1)	56,6	3	10	250.000	20	50	5	5	100	2
q1	1,13	0	0	5.000	2	5	0	0	2	0
p1	5,66	0	0	25.000	5	10	0	0	50	0,5
v1										
g (b2)	67,93	2	5	175.000	10	30	2	2	400	3
q2	1,35	0	0	3.500	1	3	0	0	8	0
p2	5,66	0	2,5	17.500	5	10	1,5	1,5	150	0,5
v2										
g (b3)	80,38	1	1	100.000	5	15	0	0	600	4
q2	1,6	0	0	2.000	1	2	0	0	12	0
p2	6,22	0	2	10.000	2,5	7,5	1	1	100	0,5
v2	23,78								500	2
Max/Min? (1/-1)	1	-1	-1	1	-1	-1	-1	-1	1	1

Table 4. Social Assessment - IRIS Categories Boundaries Layout

The relation between the social criteria and the IRIS software variables representing criteria weights (km, n=1...10) is given in the second line of Table 4. The decision panel established the following constraints:

- The Skilled Labor (k9) and Vendor Support (k10) criteria have the same importance:  $k9 = k10$ .
- The Police Station (k7) and Fire Station (k8) criteria have the same importance:  $k7 = k8$ .
- The crime against persons (k5) and property (k6) have the same importance:  $k5 = k6$ .
- Taking into account that all parishes are at least in a reasonable distance from a emergency service, and to run a DC it is essential to ensure Vendor support (k10), as well as Skilled Labor (k9) it was assumed that these criteria are more important than the Police (k7) and Fire Station (k8) Criteria:  $k9 > k7$  (i.e.,  $k9 - k7 \geq 0.05$ ) and  $k9 > k8$ . Similarly, Skilled Labor (k9) is a more important criterion than the Life Quality (k1):  $k9 > k1$ .
- Nevertheless, Life Quality (k1) was considered to be more important than the Train Station, because it is preferable to be in the top 30 municipalities in terms of life quality than being just a 1 or 2 Km from a train station:  $k1 > k2$ .
- The Life Quality (k1) criterion is more important than the Main Road Access (k3), because is preferable to be in the top 30 municipalities in terms of life quality than being just a 1 or 5 Km from a main road access node:  $k1 > k3$ .
- The Life Quality (k1) criterion is more important than the Life Cost (k4) Criterion:  $k1 > k4$ .
- The Life Quality (k1) criterion is more important than the Crime Against Person (k5) and Property (k6) criteria:  $k1 > k5, k6$ .
- The Train Station (k2) is more important than the Main Road Access (k3) criterion. The proximity of public transportation is an important factor for the Data Center staff:  $k2 > k3$ .

### 3.2.2.3 IRIS Results

In Fig.11b) we present the IRIS results for the Social Assessment, with and without the application of the constraints defined by the decision panel, and also considering the central weights inferred by IRIS. In this case taking into account the constraints on criteria weights leads to a noticeable reduction in the range of categories to which the alternatives can be assigned to. This occurs mainly for parishes with very contrasting performances (some are very good and others weak), which can easily become globally bad or good if no

restrictions are placed on the weights. In this case only 6 parishes out of the 62 maintain the potential classification. Considering the central weights to the imposed constraints, no parish is assigned to C1 (Weak), 21 parishes are placed to C2, 36 parishes are placed to C3 and 5 are assigned to C4.

### 3.2.3 Economic Assessment

#### 3.2.3.1 Range Value Limits

The Economic Assessment is related with the costs and benefits to its owner from having a DC located in a certain region. The economic category levels are identical to the Social Assessment, see Fig.12. The range of values corresponding to each category is given in Table 5, leading to the  $g_j(b_h)$  values depicted in Table 6, were defined taking into account the DMT's inputs.

Economic Categories	Investment Costs			Attractiveness to Customers		
	Electric Grid Connection Cost	Local Incentives (NSRF IQREN)	Network Communications Connection Costs	Proximity/Distance to Main Road Access (Km)	Proximity/Distance to Rail Station	Proximity/Distance to Airports (Km)
Weak	4	0	4	distance >10	4	distance >60
Good	3	1	3	5 < distance <=10	3	30 < distance <=60
Very Good	2	1	2	1 < distance <=5	2	15 < distance <=30
Excellent	1	1	1	distance <=1	1	distance <=15

a)

Economic Categories	Operational Costs							
	Availability of Qualified Support Vendor	Taxes (IMI) %	Potential for Free Cooling by Water (Productivity =P)	Potential for Free Cooling by Air (No hours/year = Nh)	Renewable Solar Resource (No hours/year = Nh)	Renewable Wind Resource (No hours/year = Nh)	Network Communications Cost (%)	Human Resources (Salary = S )
Weak	1	0.65-0.7	50=< P <100	Nh <7.500	2.300=< Nh <2.500	Nh <1.800	4	S >=1.150
Good	2	0.55-0.6	100=< P <250	7.500=< Nh <7.700	2.500=< Nh <2.700	1.800=< Nh <2.000	3	1.050<S <=1.150
Very Good	3	0.45-0.5	250=< P <400	7.700=< Nh <7.900	2.700=< Nh <2.900	2.000=< Nh <2.200	2	950<S <=1.050
Excellent	4	0.4	P >=400	Nh >=7.900	Nh >=2.900	Nh >=2.200	1	S <=950

b)

Table 5 a); b). Economic Categories Boundaries

Regarding the values indicated in Table 5, it should be noted that:

- The connection to the electric grid and to communication networks backbones at longer distances imply higher investment costs. Longer distances to the communications network imply also higher operational costs (telecommunications circuits cost).
- In terms of local incentives, an alternative is Weak only if it is not eligible for financial incentives under the NSRF.
- A higher availability of qualified support vendor in the region will reduce the associated costs with service levels agreement.
- Concerning the potential for free cooling by water, in Portugal the aquifers productivity range from 50 to 500 m<sup>3</sup>/dayKm<sup>2</sup>.
- Concerning the potential for free cooling by air, the assessment of the criterion was made using the number of hours per year with average air temperature below 21°C. In Portugal the best parish is Manteigas with 8.217 hours/year. The parishes under analysis range from 7.500 to 7.814 hours/year. For

the decision panel 7.500 hours/year of air free cooling is a good value, therefore it was established as a Weak category any value below 7.500 hours/year (albeit there are no parishes in this category). The requirement of 7.900 hours/year for the Excellent category implies that none of the parishes under analysis is able to reach it.

- Concerning renewable energy resources, in Portugal the sunshine hours potential ranges from 1.800 to 3.200 hours/year (Ambiente, 1975) (2.400 to 2.700 hours/year for the parishes under analysis, which means none of them is Excellent), whereas the wind potential ranges from 1.000 to 3.200 hours/year (Esteves, 2004) (1.200 to 2.200 hours/year for the parishes under analysis).
- Concerning human resources costs, the chosen proxy variable (average base salary of a technical engineer) varies between €886,78 (Guarda district) and €1.434,48 (Lisbon district).

### 3.2.3.2 IRIS Parameters Set

The indifference and preference threshold values for each category, as well as the veto thresholds are depicted in Table 6. The chosen veto thresholds associated with the best category reflect the DMT's opinion that a parish should be at least good in Local Incentives (i.e. must have access to incentives), Water Free Cooling and Human Resources criteria in order to be classified as Excellent in the Economic Assessment. Indifference and preference thresholds were also set in a way similar to that described in section 3.2.1.2.

	Electric Grid Connection Cost	Local Incentives (NSRF IQREN)	Network Communications Connection Costs	Availability of Qualified Support Vendor	Taxes (IMI) %	Potential for Free Cooling by Water	Potential for Free Cooling by Air	Renewable Solar Resource	Renewable Wind Resource	Network Communications Cost (%)	Human Resources	Proximity/ Distance to Main Road Access	Proximity/ Distance to Rail Station	Proximity/ Distance to Airports
IRIS Variables	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10	k11	k12	k13	k14
g (b1)	3	1	3	2	0,6	100	7.500	2.500	1.800	3	1.150	10	3	60
q1	0	0	0	0	0	0	0	0	0	0	23	0	0	0
p1	0	0	0	0,5	0	25	0	100	0	0	50	0	0	0
v1														
g (b2)	2	1	2	3	0,5	250	7.700	2.700	2.000	2	1.050	5	2	30
q2	0	0	0	0	0	0	12	0	0	0	21	0	0	0
p2	0	0	0	0,5	0	75	100	100	0	0	50	2,5	0	0
v2														
g (b3)	1	1	1	4	0,4	400	7.900	2.900	2.200	1	950	1	1	15
q2	0	0	0	0	0	0	12	0	0	0	19	0	0	0
p2	0	0	0	0,5	0	75	100	100	0	0	50	2	0	0
v2		0,1	0	0	0	300					200			
Max/Min?(1/-1)	-1	1	-1	1	-1	1	1	1	1	-1	-1	-1	-1	-1

Table 6. Economic Assessment - IRIS Categories Boundaries Layout

Regarding criteria within each main criterion the decision panel established the following constraints:

- Investment Costs: The Local Incentive (k2) is the most important criterion, followed by the Electric Grid (k1). Therefore:  $k2 > k1 > k3$ .
- Operational Costs: The Human Resources (k11) criterion is the most important criterion, therefore:  $k11 > k_n$  ( $n= 4;5;...10$ ). For the DMT, the possibility of saving energy costs using free cooling is an important factor. The Data Center can consume large amounts of water in cooling in adiabatic cooling processes, which can increase the operational costs. This criterion assumes a higher importance compared to the air free cooling because the predetermined parishes already have a reasonable potential for the latter. Therefore:  $k6 > k7$ .

The Air Free Cooling (k7) criterion assumes the following importance:  $k7 > k_n$ , ( $n= 4;5,8;9;10$ ). Furthermore, for the decision panel it is more important to surpass the category limits for the Wind (k9)

resource than for Solar (k8) resource, since the wind technology is more mature and competitive than the solar. Therefore:  $k_9 > k_8$ .

- Attractiveness to Customers: According to the DMT, customers may need to move to the Data Center any time of the day (e.g. for troubleshooting) and therefore cannot be dependent on public transportation schedules. It was decided that the proximity to Main Road Access (k12) is a more important factor than Rail Station (k13) or Airport Proximity (k14). It is also more important to be more close to rail stations than to airports. Therefore:  $k_{12} > k_{13} > k_{14}$ .

The decision panel established also the following constraints:

- The Local Incentive (k2) is the most important criterion criteria followed by the Human Resources criterion (k11):  $k_2 > k_{11} > k_n$ ,  $n=1;3;... 10;12;13;14$ .
- The Electric Grid (k1) and the Network Communications (k3), are more important criteria than any attractiveness to customers criteria:  $k_1 > k_n$  and  $k_3 > k_n$ ,  $n=12;13;14$ .
- The Water Free Cooling (k6), Air Free Cooling (k7) and the Wind Resource (k9) criteria are more important than any attractiveness to customers criteria:  $k_6 > k_n$ ;  $k_7 > k_n$ ;  $k_9 > k_n$ ,  $n=12;13;14$ .

### 3.2.3.3 IRIS Results

Fig.11c) presents the IRIS results for the Economic Assessment, with and without the application of the constraints defined by the decision panel, and also considering central weights inferred from the constraints. Without constraints the results are not very useful, as most of the alternatives could belong to any of the categories depending on the chosen weights. When considering the constraints, the results introduce significant differences between the alternatives; in fact all 62 parishes change the potential classification. With the application of the central weights, the majority of the parishes are placed to C2 (14 parishes) and C3 (36 parishes), and only 2 parishes are assigned to C1 and 10 parishes reaches C4.

## 3.2.4 Environmental Assessment

### 3.2.4.1 Range Value Limits

The four categories defined for the environmental assessment are characterized by Fig. 12 and Table 7. The corresponding  $g_j(b_h)$  values are depicted in Table 8.

Environmental Categories	Potential for Free Cooling by Water (Productivity =P)	Potential for Free Cooling by Air (No hours/year = Nh)	Renewable Wind Resource (No hours/year = Nh)	Renewable Solar Resource (No hours/year = Nh)
Weak	$50 \leq P < 100$	$Nh < 7.500$	$Nh < 1.800$	$2.300 \leq Nh < 2.500$
Good	$100 \leq P < 250$	$7.500 \leq Nh < 7.700$	$1.800 \leq Nh < 2.000$	$2.500 \leq Nh < 2.700$
Very Good	$250 \leq P < 400$	$7.700 \leq Nh < 7.900$	$2.000 \leq Nh < 2.200$	$2.700 \leq Nh < 2.900$
Excellent	$P \geq 400$	$Nh \geq 7.900$	$Nh \geq 2.200$	$Nh \geq 2.900$

Table 7. Environmental Categories Boundaries

Regarding the values indicated in Table 7, it should be noted that for the Potential for Free Cooling by Water, the Potential for Free Cooling by Air and the Renewable Energy Sources – Wind/Solar criteria, the DMT assumed the same principle described in section 3.2.3.1 but adapted to the environmental categories.

Concerning the Interference with Protected Areas criterion, all the predetermined parishes are outside of reserves and natural parks. Since it would not be a differentiating factor this criterion was not considered.

### 3.2.4.2 IRIS Parameters Set

Similarly to the other components under analysis, the indifference, preference, and veto threshold values for each Environmental category were defined (Table 8). The DMT established that one region to be classified as Excellent should be at least Good in Water Free Cooling, Air Free Cooling and Renewable Energy Resources (wind and solar) criteria, leading to the veto thresholds placed on the best category.

	Potential for Free Cooling by Water	Potential for Free Cooling by Air	Renewable Wind Resource	Renewable Solar Resource
IRIS Variables	k1	k2	k3	k4
g (b1)	100	7.500	1.800	2.500
q1	0	0	0	0
p1	25	0	0	100
v1				
g (b2)	250	7.700	2.000	2.700
q2	0	12	0	0
p2	75	100	0	100
v2				
g (b3)	400	7.900	2.200	2.900
q2	0	12	0	0
p2	75	100	0	100
v2	300	400	400	400
Max/Min?(1/-1)	1	1	1	1

Table 8. Environmental Assessment - IRIS Categories Boundaries Layout

The decision panel established the following constraints on the criteria weights:

- To reduce CO2 emissions it is more important to surpass the category boundaries for Free Cooling by Air (k2) than for Free Cooling by Water (k1):  $k2 > k1$ .
- The wind technology is considered more mature than solar and can produce more energy, hence:  $k3 > k4$ .
- Taking into account the range values for Air Free Cooling (k2) and Wind Resource (k3), the decision panel was asked what is more important for a region to change from a Very Good to an Excellent category: is it to have more 200 (=7900-7700) hours of air free cooling or to have more 200 (=2900-2700) hours of wind? According to the DMT, the latter would be more important so:  $k3 > k2$ .

### 3.2.4.3 IRIS Results

Fig.11d) presents the IRIS results for the Environmental Assessment, with and without the application of the constraints defined by the decision panel, as well the results with central weights inferred from the constraints. It is also noticeable that in this case, the use of constraints did not change significantly the classification. It is noteworthy that there are no parishes classified as excellent. This is mainly due to the demanding boundaries (b<sub>3</sub>) placed on the best category. Nevertheless, the application of the central weights, shows that the best category that can be reached is C3 category (one parish), and the majority of the parishes are assigned to C1 (32 parishes) and C2 (29 parishes) categories.

## 4. Results

In the previous section, for each of the four dimensions in analysis, we presented the results in three



different cases concerning criteria weights: without constraints, with constraints, and with central weights inferred from the imposed constraints. The results discussed hereafter are obtained considering the cases in which the constraints of the DMT are taken into account. The DMT can analyze these results focusing on the sorting suggested by the inferred central weights or considering instead the robust sorting intervals. In the later case, the results obtained for each parish can be summarized by a vector of 2×4 elements, ranging from C1 to C4: worst and best category in the risk assessment, worst and best category in the social assessment, worst and best category in the economic assessment, and worst and best category in the environmental assessment.

The DMT can apply filters to find groups of regions that meet certain conditions. Table 9 presents a set of eighteen parishes clustered in five groups according to the application of different filters. For instance, if the DMT was interested in knowing the parishes that could achieve Excellent in the Economic dimension and Very Good (or better) in the others, this would result in the regions identified in group I. Group II represents the regions that can achieve Excellent in the Social dimension and Very Good in the others. If the DMT was interested in knowing the parishes in the Environmental dimension that at least ensure a Good or higher performance and could achieve a Good performance in other dimensions, this would result in parishes in group III and IV (in group III a Very Good Social performance is ensured, and in group IV a Good Economic performance is at least ensured).. If the DMT was interested in knowing the parishes that could achieve Excellent in the Risk dimension and with the best potential performance in the others, this would result in the regions identified in group V.

Coincidentally, each of these 5 groups represents a municipality or part of a municipality: This is due to the limitations on some of the data (e.g. Crime Against Persons; Earthquakes; Life Quality; others) that are available only at municipality level, and therefore all the parishes at the same municipality have the same performance for these attributes.

Fig. 13 provides a graphical summary of the fourfold evaluation of the groups mentioned above: the four axes represent the four dimensions and are scaled from 1 (category C1) near the center to 4 (category C4) in the extreme points. For instance, parishes in group I can be C3 or C4 (depending on the weights eventually chosen) in the Economic dimension, are always C3 in the Risk and Social dimensions, and can be C1, C2, or C3 in the Environmental dimension.

Group	Parishes	Municipality	District
Group I	GLÓRIA	AVEIRO	AVEIRO
	VERA CRUZ	AVEIRO	AVEIRO
Group II	SÉ	PORTO	PORTO
	MIRAGAIA	PORTO	PORTO
	VITÓRIA	PORTO	PORTO
	MASSARELOS	PORTO	PORTO
	SANTO ILDEFONSO	PORTO	PORTO
	CEDOFEITA	PORTO	PORTO
Group III	MALVEIRA	MAFRA	LISBOA
Group IV	MATACÃES	TORRES VEDRAS	LISBOA
Group V	GILMONDE	BARCELOS	BRAGA
	BARCELOS	BARCELOS	BRAGA
	VILA FRESCAINHA (SÃO PEDRO)	BARCELOS	BRAGA
	RIO COVO (SANTA EUGÉNIA)	BARCELOS	BRAGA
	VILA FRESCAINHA (SÃO MARTINHO)	BARCELOS	BRAGA
	MANHENTE	BARCELOS	BRAGA
	TAMEL (SÃO VERÍSSIMO)	BARCELOS	BRAGA
	ARCOZELO	BARCELOS	BRAGA

Table 9. Parishes for the development of a Sustainable DC



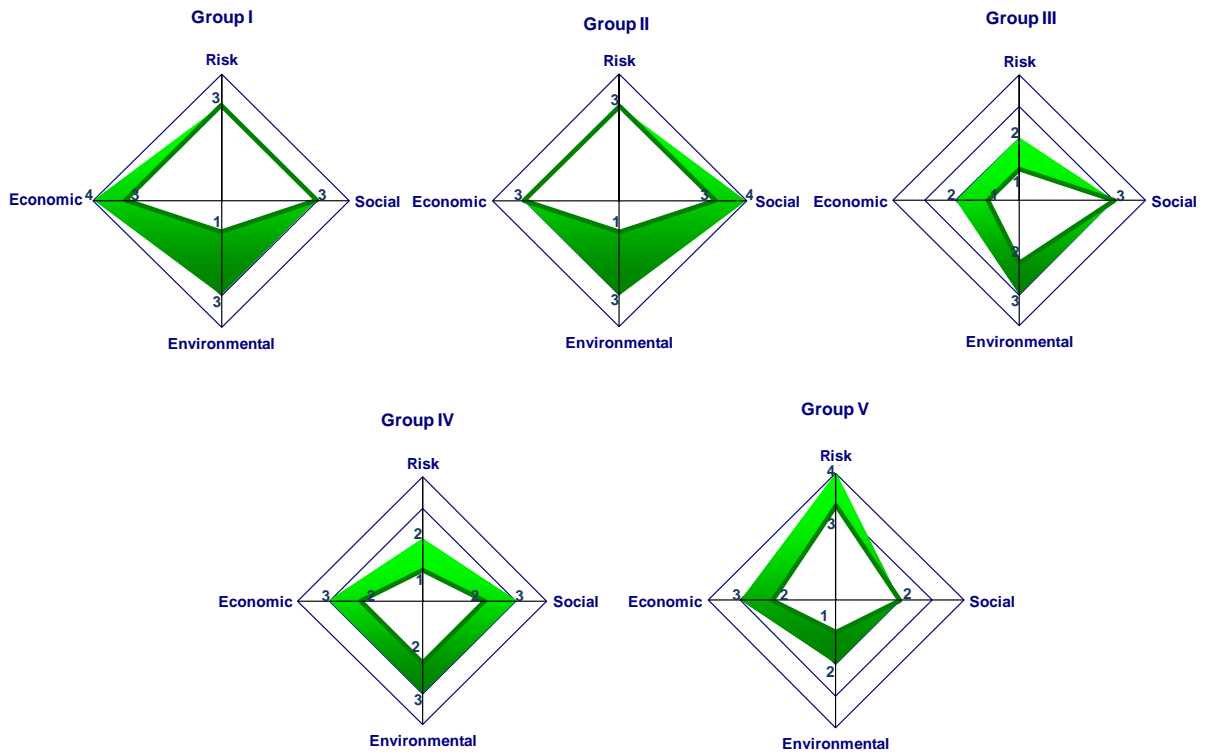


Fig.13. Group Analysis in the four components: Risk; Social; Economic; Environmental

It is possible to identify groups of parishes with the potential to reach the category C3 in all the components. Fig. 14 presents the location of these parishes in the Portuguese territory. It is also possible to identify, two groups of parishes with the potential to exceed these groups of parishes: these are the parishes in group I (two parishes in the municipality of Aveiro) and group II (six parishes in the municipality of Porto), highlighted in Fig. 14 with a darker shade. The main difference between these two groups is that the parishes in group I can achieve C4 in the economic component, while the parishes in group II can achieve C4 in the social component. In both cases however, they can also be categorized as bad as C1 in the environmental component, depending on the criteria weights chosen for this dimension.

The intervention of the analysts in this study stopped at this stage, as the DMT found the resulting methodology to be requisite for their purposes. The tool developed allows querying the IRIS results to check which parishes can reach (or are guaranteed to reach) a specified set of categories in the four dimensions, for the acceptable weights defined. Results can be made less variable by introducing tighter constraints on the acceptable weights, but the DMT did not wish to do so at this stage.

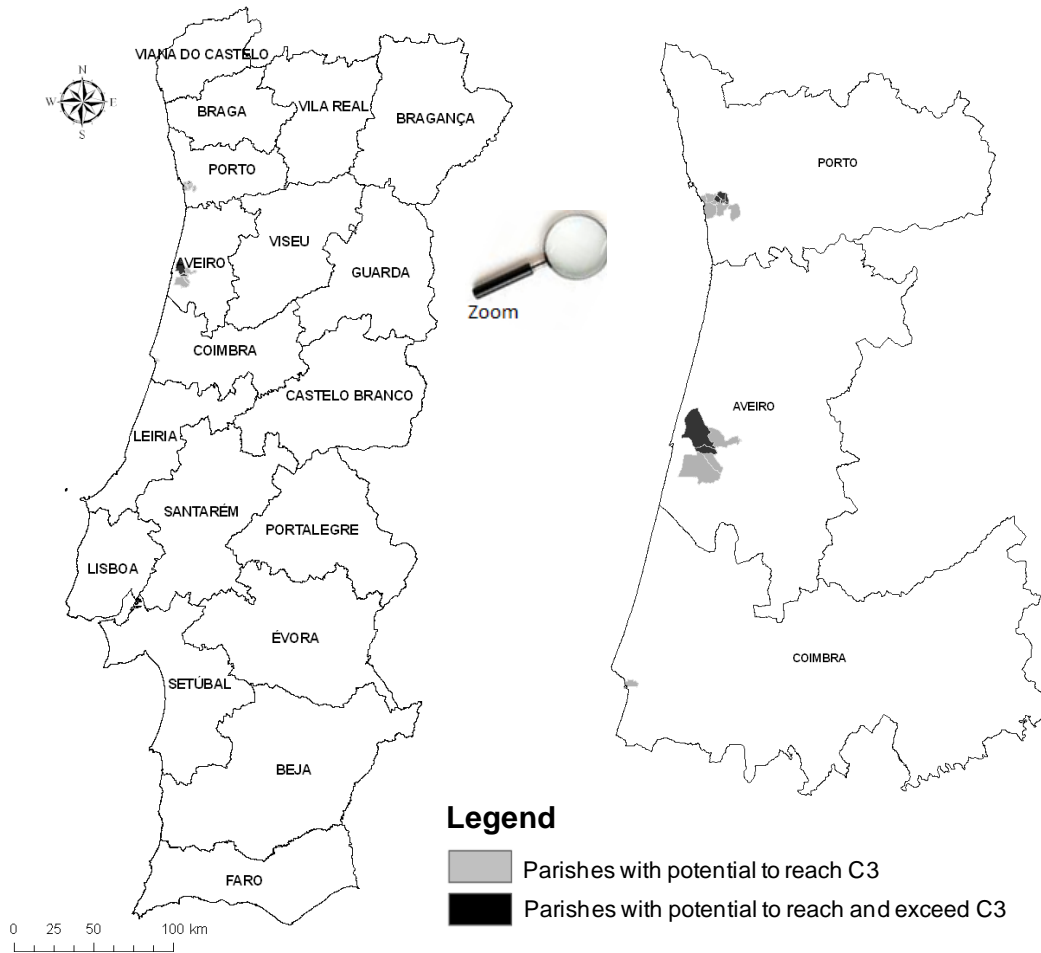


Fig.14. Regions in Portugal for the Development a Sustainable Data Center

## 5. Conclusions

This paper presents a methodology to analyze and support the decision process for identifying regions for the development of a Sustainable DC using MCDA tools. The analysis was based on the outranking method ELECTRE TRI through the IRIS software package due to the nature of the data, the desire to avoid unwarranted compensation among criteria, as well as the recognition that it would be misleading to provide other results than a separation of the alternative among categories. The crucial judgments concerning the relative importance of the criteria were facilitated by asking for simple comparisons only.

For the development of a Sustainable DC, four components were established and analyzed: Risk; Social; Economic and Environmental. Each component comprehends several criteria, in total 49 criteria were established, 14 of which were not used since they cannot be assessed at the parishes' level (the evaluation would require defining a particular piece of land). From the 35 criteria assessed, 6 belong to the Risk, 10 to the Social, 14 to the Economic and 5 to the Environmental component. For each component four categories were defined, where C1 is the least preferred (worst category) and C4 is the most preferred (best category).

The methodology is applied to a case study in Portugal, based on a previous study that has already identified a set of 62 parishes to install a Sustainable DC. The results show that there are 18 parishes out of the 62 that have the potential performance to achieve category C3 (Very Good). Among these it is possible to identify

regions with a potential to reach the best category C4 (Excellent) only in the Risk, Economic and Social dimension. Regarding the Environmental component the maximum category that can be reached is C3 (Very Good). Geographically, these potential regions are located north of Lisbon and near the coastline and even if the final solution is not located in these parishes, these solutions can nevertheless constitute a “base scenario” against which any other alternative should be compared.

The work conducted has some limitations. Many of the relevant data were available only at municipality level, and not parish level. Some of the data refers to values one or two years ago. And while the judgments concerning the criteria weights were of the ordinal type (requiring only comparisons), precise values were fixed for the remaining parameters of the model.

The authors used an original way to address the site selection for the location of a DC by evaluating separately the different sustainability dimensions of the problem. This approach contributed to a telecommunications company to have a better vision of the impact of strategies that can develop on the identification of possible regions to locate a data center. If the company chooses to have a strategy focused mainly on minimizing the risk, the regions to be analyzed to locate the DC could be different compared to a strategy focused, for example, to maximize the economic component. The results found in this study were judged to be interesting by the DMT, as it provided a vision of the Portuguese parishes where it could be very interesting to locate a DC. It also provided a methodology to address this type of problem in others regions of the world if necessary. Considering the results, the parishes should be studied in detail in order to identify a specific terrain, possibly in direct negotiations with the mayors of the respective municipalities. In this case, there are other criteria that should be considered that were not assessed in this study due to the lack of information at the parishes' level. These include the land cost or neighborhood activities and should be analyzed and evaluated with accuracy on the selection of a specific site. Using these results, further analysis can be conducted to identify (possibly using the help of the mayors of the respective municipalities) and evaluate precise locations within these few parishes to deploy the Data Center. Such a detailed evaluation could use all the criteria described in section 2.1.

The developed model was found to facilitate the site selection process, since it promotes the involvement of the DMT in the whole process, and provides to the DMT the ability of controlling the importance of each variable in the problem resolution in a transparent way, giving them the sense of ownership of the evaluation model.

## **Acknowledgements**

This work has partially been supported by MIT Portugal Program, Sustainable Energy Systems and the Programa de Financiamento Plurianual de Unidades de I&D from the Portuguese Science and Technology Foundation (FCT) to the research activities of the associated laboratory LAETA and INESC Coimbra.

## **References**

- Abbasov, A.M., Aliev, I.M., Kerimova, L.H., 2009. Optimal Location of Internet Data Centers Taking into Account the Risks. *Automatic Control and Computer Sciences* 43, 309–316.
- Al-Obeidat, F., Belacel, N. 201., Alternative Approach for Learning and Improving the MCDA Method PROAFTN. *International Journal of Intelligent Systems* 26, 444-463

- Alger, D., 2005. *Build the Best Data Center Facility for Your Business*. Cisco Press.
- Ambiente, I., 1975. CARTA I.1 - Insolação: Valores Médios Anuais (horas), Período 1931-1960. Instituto do Ambiente - Ministério das Cidades, Ordenamento do Território e Ambiente.
- ANACOM, 2010. Oferta do Lacete Local. Autoridade Nacional de Comunicações. Available at: <http://www.anacom.pt/render.jsp?contentId=891687> (accessed 2011).
- ANPC, 2010. Portuguese National Protection. Available at: [www.proteccaocivil.pt/PrevencaoProteccao/RiscosTecnologicos/EmergenciasRadiologicas/Pages/CentraisNuclearesnaEuropa.aspx](http://www.proteccaocivil.pt/PrevencaoProteccao/RiscosTecnologicos/EmergenciasRadiologicas/Pages/CentraisNuclearesnaEuropa.aspx) (accessed 2011).
- ASHRAE, 2008. *2008 ASHRAE Environmental Guidelines For Data Com Equipment*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Barda, O.H., Dupuis, J., Lencioni, P., 1990. Multicriteria location of thermal power plants. *European Journal of Operational Research* 45, 332-346.
- Belton, V., Stewart, T. J., 2002. *Multiple criteria decision analysis: an integrated approach*. Kluwer Academic Publishers.
- Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A., Vincke, P., 2006. *Evaluation and decision models with multiple criteria: Stepping stones for the analyst*. Springer.
- Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., Brandic, I., 2009. Cloud Computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems* 25, 599-616.
- Chang, S.J.F., Patel, S.H., Withers, J.M., 2007. An Optimization Model to Determine Data Center Locations for the Army Enterprise. IEEE Military Communications Conference MILCOM 2007, Orlando, FL, USA, 29-31 Oct. 2007.
- Covas, M., Silva, C., Dias, L.C., 2010. Search Analysis to Locate a Sustainable Data Center in Portugal using Geographic Information Systems. Lisbon: Euro XXIV – 24th European Conference on Operational Research.
- Dias, L.C., Mousseau, V., Figueira, J., Climaco, J., 2002. An Aggregation/Disaggregation Approach to Obtain a Robust Conclusions with ELECTRE TRI. *European Journal of Operational Research* 138, 332-348.
- Dias, L.C., Mousseau, V., 2003. IRIS: a DSS for multiple criteria sorting problems. *Journal of Multi-Criteria Decision Analysis* 12, 285-298.
- Doumpos, M., Zopounidis, C., 2002. *Multicriteria Decision Aid Classification Methods*, Kluwer Academic Publishers, Dordrecht.
- Dutta, S., Mia, I., 2010. Global Information Technology Report 2009–2010. ICT for Sustainability. INSEAD, World Economic Forum.
- Dwyer, J., 2009. Microsoft Data Centres: The Next Generation. Microsoft internal report.
- Eiselt, H.A., Marianov, 2011. *Foundations of Location Analysis*. Springer.
- Emerson, 2007. Five Strategies for Cutting Data Center Energy Costs Through Enhanced Cooling Efficiency. Emerson.
- EPA, 2007. Report to Congress on Server and Data Center Energy Efficiency. US Environmental Protection Agency Public Law, 109-431.
- Esteves, T., 2004. Base de Dados do Potencial Energético do Vento em Portugal-Metodologia e Desenvolvimento. Faculdade Ciências da Universidade de Lisboa.
- E.U.C., 2008. Code of Conduct on Data Centres Energy Efficiency Version 1.0. European Union Commission (E.U.C.).
- E.U.C., 2010. 2010 Best Practices for the EU Code of Conduct on Data Centers Version 2.0. European Union Commission (E.U.C.).
- Fernandez, E., Navarro, J., Bernal, S. 2009. Multicriteria sorting using a valued indifference relation under a preference disaggregation paradigm, *European Journal of Operational Research* 198, 602-609.
- Figueira, J., Greco, S., Ehrgott, M., 2005. *Multiple Criteria Decision Analysis: State of the art surveys*. Springer.
- Figueira, J., De Smet, Y., Brans, J.-P., 2004. MCDA methods for sorting and clustering problems: Promethee TRI and promethee CLUSTER. Tech. Rep. TR/SMG/2004-002, SMG, Université Libre de Bruxelles.

- Google, 2010. Available at: <http://www.google.com/datacenter/lenoir/faq.html> (accessed 2011).
- Hall, R.P., 2006. Understanding and applying the concept of sustainable development to transportation planning and decision-making in the US. MIT -Massachusetts Institute of Technology.
- HP, 2008. Rethinking the Data Center. Hewlett Packard (HP).
- IAEA, 2007. Arrangements for Preparedness for a Nuclear or Radiological Emergency. IAEA Safety Standards (International Atomic Energy Agency).
- IBM, 2009. Think Green Data Center Planning and Design. IBM internal report.
- INE, 2009a. *Statistical Yearbook of Lisbon Region 2008*. Portugal Statistics Institute.
- INE, 2009b. *Statistical Yearbook of Centro Region 2008*. Portugal Statistics Institute.
- INE, 2009c. *Statistical Yearbook of Norte Region 2008*. Portugal Statistics Institute.
- Kapan, J.M., Forrest, W., Kindler, N., 2008. Revolutionizing Data Center Energy Efficiency. McKinsey&Company.
- Keeney, R.L., 1992. *Value-Focused Thinking*. Harvard University Press, Cambridge.
- Manso, J.R., Simões, N.M., 2009. *Os Municípios e a Qualidade de Vida em Portugal: Proposta Metodológica Com Vista á Sua Mensuração e Ordenação*. Universidade da Beira Interior (UBI).
- Merkhofer, M.W., Keeney, R.L., 1987. A multiattribute utility analysis of alternative sites for the disposal of nuclear waste. *Risk Analysis* 7, 173-194.
- Middleton, D., 2009. Wynyard Park Data Centre Programme Overview. EDS/HP internal report.
- Mousseau, V., Dias, L., 2004. Valued outranking relations in Electre providing manageable disaggregation procedures. *European Journal of Operational Research* 156, 467-482.
- Neves, L. P., Martins, A. G., Antunes, C. H., Dias, L. C., 2008. A multi-criteria decision approach to sorting actions for promoting energy efficiency. *Energy Policy* 36, 2351–2363.
- Pavic, I., Babic, Z., 1991. The use of the PROMETHEE method in the location choice of a production system. *International Journal of Production Economics* 23, 165-174.
- Portugal Telecom, 2010. Relatório Sustentabilidade 2009. Portugal Telecom SGPS.
- Rocha, C., Dias, L.C. 2008. An Algorithm for Ordinal Sorting Based on ELECTRE with Categories Defined by Examples, *Journal of Global Optimization* 42, 255-277.
- Roy, B., 1991. The outranking approach and the foundations of ELECTRE methods. *Theory and Decision* 31, 49-73.
- Roy, B., 1996. *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers.
- Siddiqui, M.Z., Everett, J.W, Vieux, B.E., 1996. Landfill siting using geographic information systems: A demonstration. *Journal of Environmental Engineering* 122, 515-523.
- Snevely, R., 2001. Enterprise Data Center Design and Methodology (First Edition). Sun Microsystems.
- Social, M.T.S., 2010. Coleção Estatísticas – Quadro de Pessoal 2008. Ministério do trabalho e da Solidariedade Social - Gabinete de Estratégia e Planeamento.
- TurnerIV, P., Seader, J. H., Renaud, V., 2008. Operational Sustainability and Its Impact on Data Center, Uptime Performance, Investment Value, Energy Efficiency and Resiliency. Uptime Institute.
- Vulcanologia, C., 2010. Centro Vulcanologia e Avaliação de Riscos Geológicos - Universidade dos Açores. Available at: <http://www.cvarg.azores.gov.pt/Cvarg/CentroVulcanologia/atividadesismovulcanica/escalamercalli.htm> (accessed 2011).
- Yu, W., 1992. ELECTRE TRI: Aspects méthodologiques et manuel d'utilisation. Document du LAMSADE. Paris:Université de Paris–Dauphine, Working Paper n° 74.
- WCED, 1987. *World Commission on Environment and Development - Our Common Future*. Oxford University Press, Oxford.
- Zopounidis, C., Doumpos M., 2002. Multicriteria classification and sorting methods: a literature review. *European Journal of Operational Research* 138(2), 229–46.