Chapter 10

Multi-Criteria Assessment of Data Centers Environmental Sustainability

Miguel Trigueiros Covas, Carlos A. Silva and Luis C. Dias

Abstract The size and capacity of Data Centers (DCs) is growing at a rapid pace to meet the increased demand of data processing and storage capacity requested by a digital information society. Since DCs are infrastructures that have large energy consumption, there is a need to change their design approach to make them more efficient and more environment friendly. This research was motivated by the planning of a new DC in Portugal. It proposes a multi-criteria framework to assess the sustainability of a DC, which includes a new metric to evaluate the DC efficiency taking into account the environmental conditions of the DC location. ELECTRE TRI was chosen for aggregating different metrics concerning the environmental sustainability of a DC into sustainability categories. The evaluation methodology allows some freedom for each DC to place more weight on the aspects in which it is stronger, an analysis facilitated by the IRIS decision support system.

10.1 Introduction

As our society shifts from paper-based to digital information management, the demand for data processing and storage has increased significantly across all activity

---

Miguel TRIGUEIROS COVAS
Portugal Telecom, MIT Portugal Program, Sustainable Energy Systems
e-mail: `miguel.a.covas@telecom.pt`

Carlos A. SILVA
IN+ Center for Innovation Technology and Policy Research (IST), Technical University of Lisbon, Portugal, MIT Portugal Program, Sustainable Energy Systems
e-mail: `carlos.santos.silva@ist.utl.pt`

Luis C. DIAS
INESC Coimbra and Faculty of Economics, University of Coimbra, Portugal, MIT Portugal Program, Sustainable Energy Systems
e-mail: `lmcdias@fe.uc.pt`

sectors. Data Centers (DCs), by their data processing and storage capacity, are essential for the development of the new paradigm of collaborative networked society. With the increasing use of the Internet, telecommunications services, and IT networks internal to organizations, the number of servers and their power consumption has risen rapidly over the last years. Due to this, the implications to the capacity of power grids to supply larger amounts of electricity and the carbon emissions associated with electricity generation are getting the attention and concern from both industry and public policy makers [EPA, 2007, European Commission, 2008]. In addition, the increasing costs of electricity are making IT companies aware of the importance of implementing an optimized infrastructure necessary to support the new IT equipment. Therefore, social, environmental and economic interests are leveraging the development of more sustainable DCs.

This work was motivated by a new DC that is going to be built in Portugal by a telecommunications company. Its purpose was to help this telecommunications company assess the sustainability of its planned new DC in a simple way that could be used as a standard by this industry. There were already concrete plans for the new DC (namely its location has already been made public) but a few design options were still under consideration and not yet definitely decided. Nevertheless, rather than focusing solely on this particular problem, we intend to propose a general framework to assess DC sustainability taking into account several environmental criteria.

The study involved the authors and three DC experts from the telecommunications company, spanning around three months in time. The authors were the analysts, who suggested a Multi-Criteria Decision Analysis (MCDA) methodology and provided guidance in its use. The set of criteria was based on a literature review on DC metrics and incorporating other sustainability concerns of the problem owner. The discussions among analysts and experts led to the suggestion of replacing the most used criterion in the industry by a variant that takes meteorological data into account. This discussion led the team to propose a benchmarking tool that would encourage DC designers to take as much advantage as possible of the opportunities they have to increase the energy efficiency of the DC.

This chapter proceeds as follows: Section 10.2 briefly introduces the problem of assessing sustainability. Section 10.3 presents a review of the main metrics to evaluate the DCs performance. Section 10.4 presents a new metric to evaluate the DC efficiency taking into account the environmental conditions of the DC location. In Section 10.5, a framework to assess the DC environmental sustainability performance is proposed. Section 10.6 describes the application of the proposed framework that motivated this work. Finally, Section 10.7 draws the main conclusions of this study.

10.2 Sustainability Assessment

Sustainable development is a compromise between environmental, social and economic goals of a community enabling the well-being for the present and future generations. In the words of the World Commission on Environment and Development [World Commission on Environment and Development (WCED), 1987] sustainable development is a "development that meets the needs of the present without compromising the ability of future generations to meet their own needs". Several definitions have been proposed since then [Pezzoli, 1997] but there seems to be a consensus that sustainable development is a multi-dimensional issue that can involve a large amount of complex information [Ghosh et al., 2006]. According to Ciegis et al. [2009], addressing sustainability implies the problem of its measurement. This can be addressed by indicators, which are quantitative or qualitative measures, that should be simple (with a transparent method of calculation), and should have wide coverage and that allows setting trends [Böhringer and Jochem, 2007, Ciegis et al., 2009, Valentin and Spangenberg, 2000]. Thus, the development of sustainable strategies without indicators or qualitative reasoning would be lacking a solid foundation, as indicators are an instrument to evaluate environmental, social and economic goals. Though there is no single measure that could encompass all aspects of the concept sustainability [Ciegis et al., 2009], a collection of indicators chosen and analyzed under multiple criteria could better describe such complex concept [Gasparatos et al., 2008].

These criteria must be credible, relevant, attainable and measurable/verifiable [Global Ecolabelling Network (GEN), 2004]. Elghali et al. [2007] established three sets of criteria to assess the sustainability of a bioenergy system: economic viability; environmental performance, including, but not limited to, low carbon dioxide emissions; and social acceptability. For an urban sustainability assessment, Munda [2005] established the city product per person for the economic dimension, the use of private car and the solid waste generated per capita criteria for the environmental dimension and the crime rate, houses owned, and the mean travel time to work among others criteria for the social. Other criteria examples used to assess sustainability, as gross domestic product, pollution emissions (CO2, SO2; NOx) and water consumption, can be found in the literature [Gasparatos et al., 2008, Munda and Saisana, 2011]. But how can this set of multi-dimensional indicators be aggregated?

Often, some indicators improve while others deteriorate. For example, when incomes grow, SO2 might go down while CO2 increases [Munda, 2005]. The aggregation of several criteria implies taking a position on the fundamental issue of compensability. Compensability refers to the existence of trade-offs, i.e. the possibility that a good score on one indicator can compensate a very bad score on another indicator, which is often considered to be unwarranted or at least undesirable for sustainability assessment [Munda and Nardo, 2005]. But at the end, the different stakeholders want to have a clear and simple message regarding the aggregated analysis of the different sustainability criteria. This can be addressed by a label.

According to Boer [2003] it is difficult to fully specify what sustainability ideally means at a level of a product, production process or producer. In the absence of a fully specified ideal model, two strategies based on sustainability labeling can be developed. One strategy, based on identifying relevant "ideals" to pursue (e.g. recycling), or a strategy based on identifying "ills" to escape from (e.g. dependence on pesticide use). For Lindblom [1990] it is easier for a heterogeneous society to agree on the "ills" (e.g. poverty) to be avoided than on the "ideals" to be achieved (e.g. distribution of income). To Boer [2003], sustainability labeling is similar to quality assurance in the marketplace, as it reveals differences between more sustainable and less sustainable practices. It is not just a message about a product or service, but a claim stating that it has particular properties, and this is the goal of a label.

There are multiple stakeholders with interest in sustainability labeling. For industry, labeling products or services is a way to improve its competitive position in the market. For consumers, a label is a distinctive symbol revealing differences between more or less sustainable practices. For policymakers, it is a tool to address the economic interest of consumers (correction of asymmetric information), or to achieve broader sustainability objectives. For non-governmental organizations, creating a sustainability label could be a way to pressure the industry or consumers to make progress towards sustainability.

We can establish two main types of labels [Wiel and McMahon, 2005]: endorsement labels and comparative labels. Endorsement labels are essentially "seals of approval" given according to specified criteria. An endorsement label could be specifically conceived for energy efficiency (e.g. US Energy Star) to provide accurate information to end users to make an informed choice and to select more energy efficient products [Saidi et al., 2011] or for environmental friendliness (e.g. the European Union Eco-Label) [Harrington and Damnics, 2004] to provide critical quality assurance information on environmental impacts of the products [Bratt et al., 2011], endorsing products that have low impact on a wide range of environment factors. Comparative labels, as the European Union Energy Label, allow consumers to compare performance among similar products using either discrete categories of performance or a continuous scale. Both endorsement and comparative labels can coexist, and can be mandatory or voluntary. Several labeling programs examples (e.g. air conditioners; fans; heat pumps) can be found around the world [Harrington and Damnics, 2004].

Labels can help organizations to better understand and improve the sustainability of their products or services. However, the authors could not find in the literature a label or tool addressing the Data Center environmental sustainability in more than one dimension.

10.3 Data Center Metrics

The objective of building a framework to assess the environmental performance of DCs led the authors to perform a literature review to compile the most used metrics in this context. These are presented in Table 10.1.

The Green Grid defines several metrics to evaluate DCs. The Power Usage Effectiveness (PUE) and the Data Center Infrastructure Efficiency (DCiE) metrics address the energy efficiency of the DC infrastructure [Green Grid, 2007]; the Carbon Usage Effectiveness (CUE) [Green Grid, 2010a] addresses the carbon emissions associated with the DC operation (in terms of CO2-equivalent emissions); the Water Usage Effectiveness (WUE) [Green Grid, 2011] addresses the water usage in DCs, including the water used for humidification and water evaporated on-site for energy production or cooling of the DC and its support systems. The Green Grid developed also the Energy Reuse Effectiveness (ERE) [Green Grid, 2010b] metric, to measure the benefit of reusing the energy produced in the DC on other external infrastructures, and the Compute Power Efficiency (CPE) metric [Green Grid, 2008a], which seeks to quantify the overall efficiency of a DC taking into account the fact that not all electrical power delivered to the IT equipment is transformed by that equipment into a useful work product.

Metric	Definition
Power Usage Effectiveness (PUE)	$\text{PUE} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$
Data Center Infrastructure Efficiency (DCiE)	$\text{DCiE} = \frac{1}{\text{PUE}} \times 100$
Carbon Usage Effectiveness (CUE)	$\text{CUE} = \frac{\text{CO2 emitted (kgCO2eq)}}{\text{Unit of Energy (kWh)}} \times \text{PUE}$
Water Usage Effectiveness (WUE)	$\text{WUE} = \frac{\text{Annual Site Water Usage}}{\text{IT Equipment Energy}}$
Energy Reuse Effectiveness (ERE)	$\text{ERE} = (1 - \frac{\text{Reuse Energy}}{\text{Total Energy}}) \times \text{PUE}$
Compute Power Efficiency (CPE)	$\text{CPE} = \frac{\text{IT Equipment Utilization}}{\text{PUE}}$
Power Overhead Multiplier (SI-POM)	$\text{SI-POM} = \frac{\text{DC Power Consumption at the Utility Meter}}{\text{Total hardware power consumption at the plug for all IT}}$
Hardware Power Overhead Multiplier (H-POM)	$\text{H-POM} = \frac{\text{AC Hardware Load at the Plug}}{\text{DC Hardware Compute Load}}$
Deployed Hardware Utilization Ratio (DH-UR)	$\text{DH-UR} = \frac{\text{No. of Servers Running Live Applications}}{\text{Total No. of Servers Actually Deployed}}$
Corporate Average Data Center Efficiency (CADE)	Facility Efficiency x IT Asset Efficiency

Where:

Facility Energy Efficiency (%) = IT load / Total Power Consumed by the DC
Facility Utilization (%) = Actual IT load (servers, storage, network equipment) used / Facility Capacity
IT Utilization (%) = Average CPU Utilization
IT Energy Efficiency (%) = CPU Loading / Total CPU Power
Facility Efficiency = Facility Energy Efficiency (%) × Facility Utilization (%)
IT Asset Efficiency = IT Utilization (%) x IT Energy Efficiency (%)

Table 10.1 Data Center Metrics.

The Uptime Institute [Stanley et al., 2007] defined other metrics. The Site Infrastructure Power Overhead Multiplier (SI-POM), similar to the PUE metric, indicates how much of the DCs site power is consumed in overhead instead of being used by the IT equipment. The IT Hardware Power Overhead Multiplier (H-POM) addresses the IT equipment efficiency, by evaluating how much of the power input in the hardware is wasted in power supply conversion losses or diverted to internal fans, rather than in useful computing components. The Deployed Hardware Utilization Ratio (DH-UR) indicates the fraction of the deployed IT equipment in the DC that is not running any application or handling data. Finally, the Corporate Average Data Center Efficiency (CADE) metric, defined together by the Uptime Institute and McKinsey [Kapan et al., 2008], addresses both the physical infrastructure and the IT systems.

Other examples of DC metrics can be found on the Uptime Institute [2012] or in the Green Grid [2012] organizations. All these metrics are ratios that can be used to assess the efficiency of a DC. However it seems that no attempt has been done to develop an integrated indicator or a label. And despite the diversity of the existing metrics, the PUE has been used worldwide by the industry as a tool for measuring and benchmarking DCs energy efficiency. In fact, the European Union, the United States of America and Japan established in February 2010 an agreement to use the PUE as the metric guide for DCs energy efficiency [Energy Star, 2011].

The use of PUE only for benchmarking purposes must be, however, carefully analyzed, since we need to understand the conditions of the DC infrastructure. First, the operation constraints such as the redundancy level of a DC can influence the PUE, as the use of more levels of standby electrical infrastructure to reduce downtime may introduce additional power losses. Furthermore, PUE does not account for the environmental conditions of the DC site, which can influence the energy efficiency of the facility due to the needs of the cooling systems, which are responsible for the consumption of a considerable amount of energy (around 37%, according to Emerson [2007] in Figure 10.1). In particular, depending on the geographic location of the DCs, the potential to use free cooling solutions (air-side economizer system) can vary and influence directly the amount of energy that can be saved by using cold outside air to directly or indirectly cool the computer room, avoiding the use of the chillers as cooling systems. We propose therefore a new metric that copes with this issue.

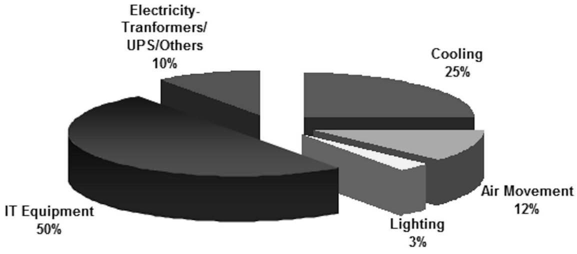


Fig. 10.1 Energy Usage in a Data Center [Emerson, 2007]

10.4 A New Metric: TRUE

As mentioned in the previous section, PUE has been adopted as the metric guide to evaluate DCs efficiency. However, the use of the PUE metric in this study raised some concerns among the analysts and participating DC specialists as they concluded that it does not necessarily foster energy efficiency practices. The following example demonstrates this.

Consider two DCs at two different locations with the same systems (IT; power distribution, generators, UPS, etc.) but with different cooling systems. Let us suppose that one of the DCs is located in a warmer climate, but they have the same PUE. This means that their cooling infrastructures (all else being equal) are using the same amounts of energy, which means that under the PUE metric, they could be considered equally efficient. However, it is easier for a DC located in a very cold climate (e.g., in the Arctic region) to profit from free cooling than for a DC located in a warmer climate such as Portugal. Thus, the DC located in the colder region is not taking advantage of the local conditions and it is not being as energy efficient as it could be.

The problem owners were not considering the possibility of building a DC in the Arctic. The problem owners do need that some of their employees live near the DC and that other personnel, including clients of the DC, can easily visit the DC for maintenance or other operations. The public image of the telecommunications company might also be at stake if it opted to build its flagship DC in another country. The problem owners main concern is that the DC is as energy efficient as it could be. Thus, benchmarking based on the PUE metric may not be totally fair.

The PUE metric does to some extent penalize DCs in countries with warmer climates. If the climate is taken into account, we could consider that the DC located in the colder climate is less efficient, since it is not able to profit from a better free cooling potential in order to use less energy for the cooling infrastructure. It is therefore necessary to develop another metric that will encourage a DC to profit as much as possible from the free cooling potential of the region it is located at.

We developed a metric called Temperature of the Region Usage Effectiveness (TRUE), to take into account the efficient use of the air free cooling potential - and corresponding impact on energy use in a DC location. As discussed, the cooling system represents more than one third of the energy consumption but the type of system and its efficiency depends on the location temperature conditions. The TRUE metric is thus a correction factor to the PUE metric that tries to incorporate the temperature conditions impact on efficiency, using a correction factor which describes the number of hours per year at which the temperature of the region where the DC is located has an average value that allows to use free cooling systems.

Free cooling can be used only in climate zones where the outside air temperature and humidity conditions are appropriate. For the ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers) [ASHRAE, 2008] and the European Commission [2010], DCs should be designed to operate at "inlet" (supply air) temperatures between 18^{o}C and 27^{o}C, to maximize energy efficiency. Following these recommendations and considering that for "inlet" temperatures up to 25^{o}C the server fans will already run at high speed and consume more energy [ASHRAE, 2008, Sartor and Greenberg, 2008, J. H. Bean, 2011], and that the usual temperature gains when using free cooling technology are 4^{o}C [Dunnavant, 2011], we established 21^{o}C as the top limit for the outside air temperature to allow the use of free cooling. We did not set any limit for humidity levels because depending of the free cooling technology (direct or indirect) this constraint can be easily managed.

The TRUE metric is defined as follows. Let us first note that PUE (Table 1) can be rewritten as:

$$PUE = \frac{\text{Cooling System Power+IT Eq.Power+Lighting Power+Other(e.g.UPS)}}{\text{IT Equipment Power}}. \tag{10.1}$$

Thus, we define TRUE as:

$$TRUE = \frac{\text{IT Eq.Power+Lighting Power+Other(e.g.UPS)}}{\text{IT Equipment Power}} + C_f \times \frac{\text{Cooling System Power}}{\text{IT Equipment Power}}, \tag{10.2}$$

where:

$$C_f = \frac{1}{8760} \sum_{t=-\infty}^{21} n_t. \tag{10.3}$$

Here, C_f is the Correction Factor, n_t is the annual number of hours per year with average temperature t under 21^{o}C and 8760 is the total number of hours per year.

The success of PUE is mainly due to two reasons: it is easy to understand and it does not require complex mathematical formulas to be applied. With the aim to maintain a straightforward and easy way to calculate the metric, we defined the simplified TRUE metric as follows:

$$TRUE' = 1 + C_f \times (PUE - 1) \tag{10.4}$$

In this case, the correction factor is also affecting the entire electric infrastructure system of the DC (e.g. lighting, that could represent 3% of the energy consumption in the DC), even though its performance is not influenced by the outdoor temperature conditions. All the calculations reported in this article are based on this simplified version of TRUE.

In Figure 10.2 we can see the impact of the climate region around the world on the TRUE metric. For instance, a DC in Lisbon with a PUE of 1.25 has better TRUE (1.193) than one DC with a PUE of 1.25 in London (TRUE=1.24). This means that a DC in London that has the same IT infrastructure than one in Lisbon and has the same PUE is not really taking full advantage of its location in terms of free cooling potential.

This metric could be a stimulus for the organizations to pursue the effective use of the natural resources to maximize the operational efficiency and reduce the impact on the environment and resources, i.e., an enabler for the development of more sustainable DCs. However, it still does not take into account other issues, such as carbon emissions associated with the electricity consumption or the use of water.

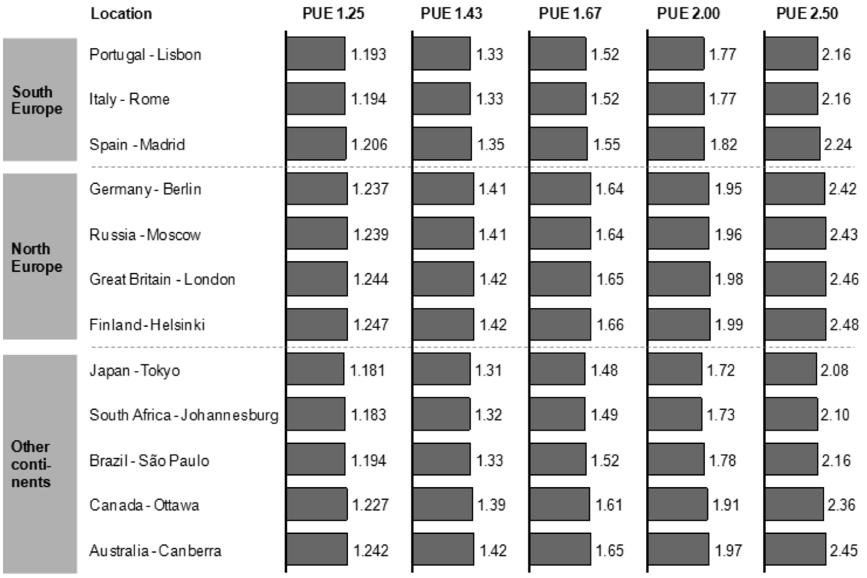


Fig. 10.2 TRUE metric for temperatures below 21^{o}C – Cities around the world

Thus, the analysis of DCs sustainability and efficiency should take into consideration other metrics.

10.5 A Framework to Assess the Data Center Environmental Performance

In the previous sections we reviewed the most common metrics to assess DC efficiency and introduced a new metric that considers the temperature of the region where the DC is located. The discussion about this new metric led us to conclude that the assessment methodology should be designed in a way that fosters DC planners to be as efficient as possible given the conditions of the DC's environment. Metrics can help organizations to better understand and improve the sustainability of their DCs, as well as, to help the decision makers in the deployment of new DC. It is important for the organizations to continue improving the effective use of resources to maximize operational efficiency and reduce the impact on resources and environment.

This section presents a tool to help DC managers to assess the environmental sustainability performance of their DC, using MCDA. This tool was developed by taking into account the DCs experts point of view. These experts (a team of three) possess a large experience in running DC facilities and were also able to represent the points of view of the telecommunications company.

10.5.1 Criteria

The establishment of the relevant criteria is an important step of the model. The construction of a criteria list was done taking into account the review and discussions presented in the previous sections. The analysts initially proposed a list with four criteria to the DCs experts: one criterion to evaluate the carbon emissions, a criterion to evaluate the facility efficiency (e.g. TRUE metric), a criterion to evaluate the energy reuse in the DC and a criterion to evaluate the DC local environmental impacts (e.g. noise, interference with protected areas, etc). These criteria were analyzed and discussed with the DC experts. This analysis and discussion provided a better understanding of the criteria and allowed to confirm the use of some of the criteria and also allowed to suggest modifications and addition of new criteria. For example, the experts suggested the inclusion of a criterion that evaluates the IT equipment in the DC, e.g. the server utilization. Regarding the local environmental impact, since DCs can consume large amounts of water, the DC experts suggested that the water usage in the DC should also be assessed. After this procedure, the DC experts' team approved the criteria hierarchy, described in Figure 10.3. This model with five criteria was the basis to assess the DC environmental sustainability.

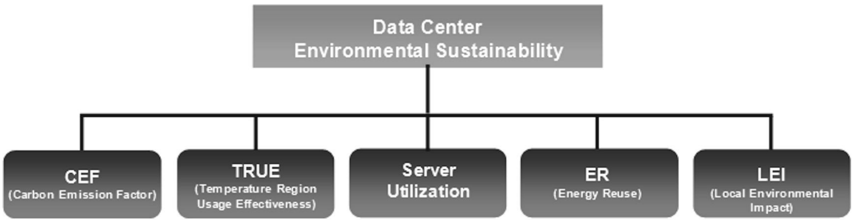


Fig. 10.3 Main criteria to assess the Data Center Environmental Sustainability Class Performance

10.5.2 Criteria Evaluation

The criteria to assess (and encourage) the DC environmental sustainability performance are from different nature and both qualitative and quantitative data are used. This diversity led to the use of different types of scales for the different criteria. The criteria assessment was done by the analysts in collaboration with the DCs experts.

1. CEF (Carbon Emission Factor): DCs can be large consumers of electricity, and depending on the electricity energy resource they can be responsible for large amounts of carbon emissions. CEF represents the carbon emissions per kWh consumed by the DC, evaluated in kgCO2eq/kWh. Its evaluation can range from 0 (e.g. use of only renewable energy resources) to infinity, where 0 is the best score.

2. TRUE (Temperature of the Region Usage Effectiveness): The TRUE metric described in Section 10.4, measures the DC support infrastructure energy efficiency taking into account the temperature in the area of its location. Its evaluation results from applying the metric, and can range from 1 to infinity (1 is the best score).
3. Server Utilization: This criterion addresses the efficiency of the IT systems. Server Utilization can include multiple systems (Central processing unit (CPU), memory, disk, network). According with the DCs experts, this depends on the intended use of the utilization data and the sophistication of the management infrastructure and applications. The experts evaluate this criterion by considering the CPU average utilization in percentage that can range from 0% to 100% (the best score).
4. ER (Energy Reuse): This criterion measures the percentage of the thermal energy generated by the DC (e.g. heat released by the IT equipment) that is being reused in other parts of the facility (e.g. dehumidification) or in other nearby external facilities (e.g. greenhouses). Its evaluation can range from 0% to 100% (the best score).
5. LEI (Local Environmental Impact): The aim of this criterion is to assess the local environmental impact of the DC on a qualitative scale. This criterion considers three different sub-criteria: Water Usage (WU), Interference with Protect Areas (IPA) and Local Impact Pollution (LIP), described as follows:

   a. *Water Usage (WU)*: the DCs can be very large consumers of water. For example, a 1MW DC operating with water-cooled chillers and cooling towers can consume up to 68.000 liters per day to dissipate the heat generated by the IT equipment [Sharma et al., 2009]. However, in general it is very difficult to establish water consumption values, so the WU evaluation was done by the perception of the DCs experts regarding the impact of the use of water. They attributed levels between 1 and 5 (the best score), as described in Table 10.2. In the evaluation, the use of water is related to the site location also, as in some areas the use of water, even if large, may not be an important constraint (e.g. if the DC is located on the shore of a large river).
   b. *Local Impact Pollution (LIP)*: to evaluate this criterion we initially proposed a list of three criteria to the DCs experts: noise (no noise impact or no impact on neighbors) and water pollution (no pollution of surface/underground natural watercourse and no water temperature increase) and ozone depletion (no degradation of the ozone layer caused by the use of cooling refrigerants or fire extinguishing gases). DCs experts suggested the inclusion of a new criterion that evaluates recycling programs, namely plans for IT equipment and packaging material recycling. Therefore, there was agreement on four sub-criteria for LIP evaluation: noise, water pollution, ozone depletion and recycling programs. Five LIP levels were defined (see Table 10.2) according to fulfillment of the described criteria.
   c. *Interference with Protect Areas (IPA)*: reserves and natural parks should be avoided to locate the DC. Therefore DCs located in protected regions, e.g.

natural parks, should be classified as having a poor environmental concern. If located outside protected areas, they should be classified as having an excellent environmental concern, as depicted in Table 10.2.

Water Usage (WU)		**Local Impact Pollution (LIP)**		**Interference with Protect Areas (IPA)**	
Description	Level	Description	Level	Description	Level
No water consumption or use	5	Local impacts are negligible	5	Outside reserves and natural parks	5
Consumption of alternative water sources, i.e. non potable water (e.g. rainwater harvesting)	4	Non-negligible but small impact in one of the aspects	4		
Use of water (from a lake, river, ocean) but without/negligible water consumption	3	Non-negligible but small impact in two of the aspects	3		
Consumption from large potential water sources (lakes, rivers, reservoirs, aquifers), i.e. ample water availability, without putting in risk the water resource over time	2	An obvious impact in one or more aspects or a small impact in three or more aspects	2		
Consumption of potable water, public water supply, water-scarce region	1	An excessive impact (given existing norms) in one or more aspects	1	Inside reserves and natural parks	1

Table 10.2 WU, LIP and IPA qualitative scale description

As mentioned the aim of LEI is to assess the local environmental impact of the DC. The authors and the DCs experts agreed that a good performance in one criterion should not be allowed to compensate a poor performance in another criterion. After discussing this issue, it was defined that the LEI performance is set by the minimum value performance of the WU, IPA and LIP criteria. For example, if a DC has a 4 in the WU, a 5 in the IPA and a 2 in the LIP criterion, the LEI performance will be 2.

10.5.3 The ELECTRE TRI Method as the Evaluation Tool

The ELECTRE TRI method [Yu, 1992] belongs to the ELECTRE family of multicriteria 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. ELECTRE TRI allows an evaluation in absolute terms, i.e., alternatives are not compared against each other but to predefined norms. The result of such analysis is a partition of the set of alternatives into several categories defined with respect to these norms (called by Roy [1996] the sorting problem formulation). Another feature is that ELECTRE models allow incomparability. Incomparability occurs when some alternatives are so different that a direct comparison is hard to justify. ELECTRE TRI does not require converting the performance criteria into a uniform scale range, as it allows the inclusion of criteria measured in different units and even measured in qualitative terms.

To establish the environmental sustainability framework to assess the DC performance several categories are pre-established to represent different environmental sustainability levels. Considering the nature of the criteria (different value scales and different value domains, qualitative data), the preference for a non-compensatory method and the possibility of using a method where the assignments of alternatives are independent from each other, are the main reasons for choosing ELECTRE TRI.

The pessimistic variant of ELECTRE TRI method was applied using the decision support software called IRIS [Dias and Mousseau, 2003b,a], which was designed to address the problem of assigning a set of alternatives to predefined ordered categories, according to their evaluations (performances) at multiple criteria. For details about the ELECTRE TRI variant used see Dias and Mousseau [2003b,a], Dias et al. [2002].

10.5.4 Model Parameters

The use of ELECTRE TRI requires to set the parameters that represent the preferences of the decision makers. In this case, this was done by the authors in collaboration with the DCs experts. The group agreed that for the study's purposes it would suffice to sort the possible alternatives (Data Centers) into five categories (levels) of environmental sustainability performance, in accordance to what it is used in energy efficiency comparative labels, according to the Table 10.3 (columns 1 and 2).

In addition to the categories definition, it is also necessary to define the category boundaries or limit profiles that represent the limit between two consecutive categories. The definition of the limit profiles was performed taking into account some support information.

For the CEF criterion profiles, it was considered the carbon emissions produced by the different power generation technologies [EDP, 2012]. The values have the following rationale: according to EDP [2012] the CEF is 0.36 kgCO2eq/kWh if electricity is generated by natural gas power plants, 0.78 kgCO2eq/kWh in case electricity is generated by fuel power plants, and 0.9 kgCO2eq/kWh in case electricity is generated by coal power plants. Illustrative scenarios were used to assess the category limits presented in Column 3 of Table 10.3: category A could represent

Environmental Sustainability Categories	Description	CEF (kgCO2e/ kWh) (Carbon Emission Factor)	TRUE (Temperature Region Usage Effectiveness)	Server Utilization (%)	ER (%) (Energy Reuse)	LEI (Local Environmental Impact)
A+	DC with an excellent performance	$\leq$ 0.18	$\leq$ 1.25	≥ 50	≥ 45	5
A	DC with a very good performance	]0.18, 0.36]	]1.25, 1.43]	[35, 50[	[30, 45[	4
B	DC with a good performance	]0.36, 0.57]	]1.43, 1.67]	[20, 35[	[15, 30[	3
C	DC with a reasonable performance	]0.57, 0.79]	]1.67, 2]	[5, 20[	[0.1, 15[	2
D	DC with a poor performance	> 0.79	> 2	< 5	< 0.1	1

Table 10.3 Environmental sustainability categories: levels and boundaries

a DC powered with a mix of 50% of natural gas power plants and 50% of nuclear and renewable plants; category B could represent a mix of natural gas (50%) and fuel power plants (50%); and category C could represent a mix of natural gas power plants (25%), fuel power plants (50%) and coal plants (25%).

The TRUE criterion considers the classification proposed by the Green Grid [2008b] for the PUE metric, see column 4 of Table 10.3. For the Server Utilization criterion, the DC experts established the limits presented in column 5 of Table 10.3. For the ER criterion, the experts defined values described in column 6 of Table 10.3. For the LEI it was straightforward to establish the values presented in column 7 of Table 10.3.

Based on the criteria description presented, the categories range values for each criterion were established. Five categories were considered, where D (category C1) is the worst, described as Poor performance, and A+ (category C5) is the best category described as Excellent performance, see Figure 10.4. Table 10.4 indicates the values considered for the reference profiles, e.g. for indicator ER we have boundaries b_1=0.1, b_2=15, b_3=30, b_4=45.

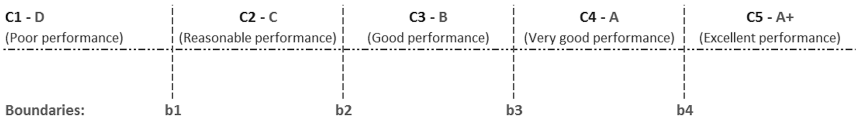


Fig. 10.4 Relation between the Environmental Sustainability Categories and the Category Profiles

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 $(v_j(b_h))$ thresholds. 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. Table 10.4 displays the IRIS thresholds that must be entered; these thresholds can be set independently for each category.

	CEF	**TRUE**	**Server Utilization**	**ER**	**LEI**
Parameter	k_1	k_2	k_3	k_4	k_5
$g(b_1)$	0.79	2	5	0.1	2
q_1	0.0395	0.100	0.25	0	0
p_1	0.11	0.165	2.5	0	0
v_1	-	-	-	-	-
$g(b_2)$	0.57	1.67	20	15	3
q_2	0.0285	0.084	1	0.75	0
p_2	0.11	0.165	7.5	7.45	0
v_2	-	-	-	-	-
$g(b_3)$	0.36	1.43	35	30	4
q_3	0.018	0.072	1.75	1.5	0
p_3	0.105	0.12	7.5	7.5	0
v_3	0.43	0.57	30	29.9	2
$g(b_4)$	0.18	1.25	50	45	5
q_4	0.009	0.063	2.5	2.25	0
p_4	0.09	0.09	7.5	7.5	0
v_4	0.39	0.42	30	30	2
Preference direction	Min	Min	Max	Max	Max

Table 10.4 Category boundaries and preference discrimination thresholds.

It was established that an alternative (a DC) should be at least C (Reasonable performance) in all criteria in order to be classified as A (Very good performance), by setting a veto threshold for all criteria to reach category C4 $(g_j(b_3) - v_j(b_3) = g_j(b_1))$. Similarly, it was also established that an alternative should be at least B (Good performance) in all criteria in order to be classified as A+ (Excellent performance), by setting a veto threshold for all criteria to reach category C4 $(g_j(b_4) - v_j(b_4) = g_j(b_2))$.

Regarding the indifference and preference thresholds, it was defined for all criteria (except the LEI) an indifference threshold of 5 % of the boundary value, and a preference threshold equal to half of the difference between adjacent categories.

Concerning the criteria weights, there was a consensus that the evaluation methodology should allow some freedom for each DC to place more weight on the aspects in which it is stronger, in the spirit of Data Envelopment Analysis (see Madlener et al. [2009]) for another example in which this type of approach was used). This avoids the controversial question of defining a precise weight for each criterion and allows DCs with different profiles to attain the best categories. Let us note however that the veto thresholds prevent a DC with a major weakness from reaching the top categories. It was defined that each criterion weight can vary from 10% to 30%. As we have five criteria, with this range of values it is guaranteed that each criterion cannot have a weight greater than the sum of the minimum weights of the number of criteria that can constitute a majority (i.e. three criteria). With this approach it is also guaranteed the possibility of all criteria having the same weight (i.e. 20%, the midpoint weight).

For the majority threshold, the value 0.66 was set in order to ensure a robust majority. This means that in order to say that an alternative is at least as good as a category limit profile, at least 66% of the criteria must be in concordance with this affirmation (after the weighting).

10.6 Application of the Model

In this section the proposed framework has been applied to a case study, the new DC that is going to be built in Portugal by Portugal Telecom (PT). The application of the model aimed at evaluating some different variants for the DC, as well as to assess the impact of the TRUE metric by considering (hypothetical) scenarios in which the same DC would be built in other locations.

PT announced on February 4^{th} 2011 the construction of a new DC in Covilhã region (a region in the center of Portugal). According to the company this new DC will be the largest ever built in Portugal, with an initial investment cost between 30 to 50 million Euro, and its main focus is to provide cloud computing services. It will have an area over 45.000 m2 and a power capacity of 40 MW. This critical facility will be a worldwide energy efficiency reference, with an expected PUE of 1.15, and it will use free cooling solutions and renewable energies (the DC will be powered by its own wind farm of 28 turbines in a total of 56 MW installed capacity [Camara Municipal da Covilhã, 2011, Portugal Telecom, 2011]).

The hypothetic scenarios refer to Data Centers located in areas with different environmental conditions, such as free cooling potential or electricity carbon emission factor, represented by the countries UK, Poland and France.

10.6.1 Criteria Evaluation

1. CEF (Carbon Emission Factor): We used the CEF published by the International Energy Agency [2010]. For Portugal, the CEF was 0.395 kgCO2eq/kWh. For the other countries considered in the comparative analysis we considered

for United Kingdom that the CEF was 0.497 kgCO2eq/kWh, for Poland it was 0.660 kgCO2eq/kWh and for France it was 0.086 kgCO2eq/kWh. Taking into account that the DC in Covilhã will have a wind farm that will be able to produce 23%[1] of the annual electricity needs, the CEF can be reduced to 0.305 kgCO2eq/kWh.

2. TRUE (Temperature of the Region Usage Effectiveness): this value was obtained taking into account a PUE of 1.15 and the temperature data of the different regions.
3. Server Utilization: according to PT, the main focus of this DC is to provide cloud computing services. Since the average CPU utilization in typical DCs is around 20% according to Meisner et al. [2009], VanGeet [2011], and considering the PT orientation, we established an average CPU utilization of 40%.
4. ER (Energy Reuse): according to PT's plans, we assumed that the DC will reuse at least 5% of the waste heat from the IT equipment in heating the offices spaces.
5. LEI (Local Environmental Impact): it was established the WU of 3, since the water consumption will be negligible and will be pumped from river Zezere (a large river in the neighborhood). The DC is located outside of reserves and natural parks and it was considered that local impact pollution of the DC will be negligible.

10.6.2 Data Center Environmental Sustainability Performance Results

This section presents the IRIS results for the DC sustainability. We considered seven different scenarios described as follows:

1. DC_PT_Covilha: the base scenario, taking into account the criteria values described in the previous section.
2. DC_PT_Covilha S1: similar to the first scenario, but without the existence of the wind farm, i.e., with an increase of the CEF.
3. DC_PT_Covilha S2: similar to the first scenario, but without the existence of the wind farm and without reuse of the energy from the IT equipment.
4. DC_PT_Covilha S3: similar to the first scenario, but without the existence of the wind farm, without reuse of the energy from the IT equipment and a decrease of the WU to 1 (i.e. LEI criterion performance equal to 1).
5. DC_PT_London: similar to the second scenario (no wind farm), but in this case the location of the DC is in London, with different temperature conditions, as well as, with different CEF.
6. DC_PT_Krakow: similar approach to the DC_PT_London scenario.

---

[1] Considering the following assumption: 56 MW of installed capacity; 2000 h/year equivalent production at maximum capacity; 40 MW average consumption power directly consumed by the DC; the excess energy from the wind park is considered to be injected into the national grid and thus is not considered.

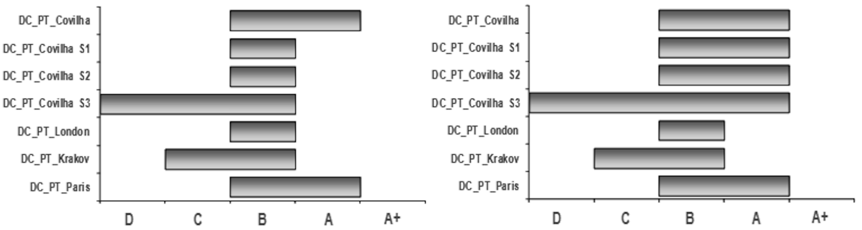
7. DC_PT_Paris: similar approach to the DC_PT_London scenario.

The criteria values used in this section for the several scenarios are summarized in Table 10.5. Although the scenarios have the same Server Utilization, this indicator contributes to define the category assignment.

Data Center	*Location*	*Free Cooling (h/yr)*	*PUE*	CEF (kgCO2e/ kWh) $g_1(.)$	TRUE $g_2(.)$	Server Utiliz. (%) $g_3(.)$	Energy Reuse (%) $g_4(.)$	LEI $g_5(.)$
DC_PT_ Covilha	*Covilhã, Portugal*	7400	1.15	**0.305**	**1.127**	**40**	**5**	**3**
DC_PT_ CovilhaS1	*Covilhã, Portugal*	7400	1.15	**0.395**	**1.127**	**40**	**5**	**3**
DC_PT_ CovilhaS2	*Covilhã, Portugal*	7400	1.15	**0.395**	**1.127**	**40**	**0**	**3**
DC_PT_ CovilhaS3	*Covilhã, Portugal*	7400	1.15	**0.395**	**1.127**	**40**	**0**	**1**
DC_PT_ London	*London, U.K*	8551	1.15	**0.497**	**1.146**	**40**	**5**	**3**
DC_PT_ Krakow	*Krakow, Poland*	8308	1.15	**0.66**	**1.142**	**40**	**5**	**3**
DC_PT_ Paris	*Paris, France*	8178	1.15	**0.086**	**1.140**	**40**	**5**	**3**

Table 10.5 Criteria Values for Portugal Telecom Data

The results from IRIS are depicted in Figure 10.5a). Since the weights are allowed to vary, there are cases in which IRIS yields more than one category: in these cases the category would depend on more precise choices for the weight values. In a "benefit of doubt'' [Cherchye et al., 2007] or benevolent perspective in the spirit of Data Envelopment Analysis [Madlener et al., 2009], each DC would be attributed the highest category allowed by the results.



a) b)

Fig. 10.5 IRIS Results for the PT DC. a) IRIS parameters set with the established Veto conditions. b) IRIS parameters set without veto conditions for category A, only the defined veto conditions for A+.

In this case, the PT DC can reach an A level or label, but not an A+. If the wind farm is not considered or if the DC does not reuse the heat from the IT equipment, the maximum performance is reduced to B. A decrease in the WU criterion may bring the classification down to D, but only if the water criterion is considered to be more important than the others.

If a similar DC was installed in London, the potential performance (B level) would be similar to the DC installed in Covilha (without the wind farm), but if it was installed in Krakow the potential performance would vary between B and C. This is due to the high level of carbon emissions of Poland (0.660 kgCO2eq/kWh). However, and due to the low carbon emissions of France, if a similar DC was installed for example in Paris, then CEF would decrease considerably and the DC could reach an A label (even without a wind farm).

In Figure 10.5b) we change the veto conditions previously established by eliminating veto condition for category A and maintaining the condition of veto for the category A+. In this scenario the range of the results can be improved, in particular the alternatives in Portugal could all reach category A. Is also interesting to highlight that London and Krakow alternatives did not change their potential.

It is also interesting to notice that no DC can ever achieve an A+ label, even if we reduce requirements of the veto conditions for this level (e.g. from all alternatives must be at least B (Good performance) to C (Reasonable performance)). In this particular case, there are no differences in the results when compared to the ones obtained in Figure 10.5b).

Comparing the two sets of results, it was felt that it would be important to take into account the veto thresholds under this type of evaluation, so that the best categories will not be achieved by DCs that are weak in one of the criteria. Another aspect that was considered noteworthy is that this tool encourages PT to build a wind farm in order to improve the CEF. Alternatively, if this were a possibility, PT could buy cleaner energy in terms of CO2 from a different grid, since this allows the Paris scenario to attain category A.

Given this set of results, the DC experts and the authors felt the tool was able to produce results that were aligned with the preferences they constructed during the discussions they had for this study.

10.7 Conclusions

We present a tool to assess the DC environmental sustainability performance using MCDA, pursuing the idea that the environmental categorization of DCs should encompass multiple dimensions. The tool is based on the outranking method ELECTRE TRI due to the nature of the data (quantitative, qualitative, different scales), the desire to avoid compensation among criteria, as well as the recognition that it would be misleading to provide results other than a separation of the alternatives among categories. With the involvement of a team of DC experts, five categories (levels) of environmental sustainability performance were defined, where D is the least preferred (worst category) and A+ is the most preferred (best category), as well as, the several criteria to evaluate the sustainability.

In total, 5 criteria were established, one of which was a new metric reflecting the concerns of the problem owners about the PUE metric. The DC should be encouraged to profit as much as possible from free cooling but taking into account what the location of the DC allows, bearing in mind that there are practical barriers to choosing an ideally cold place. DCs in warmer climates have less opportunity to improve energy efficiency by drawing upon external air than DCs in cooler climates. The TRUE does not contradict that the regions with a higher free cooling potential are potentially better areas to install a DC; it indicates that certain installations, even if they have a higher PUE, can in fact be more efficient if they take more advantage of the climate. A smaller differential between PUE and TRUE leads to a lower potential investment needed to achieve the same levels of efficiency. The TRUE provides a way to improve DCs benchmarking but mostly to determine opportunities to consider the use of alternative cooling strategies. It can help DC decision making processes related to site planning.

The tool was applied to several scenarios for the new PT data center, some of which were only hypothetical. Due to the lack of information some assumptions were taken. Considering identical DCs, more effort is required from a DC located in Portugal to reach A+ level than from one in France, due to the low carbon emission factor in this country. The results show that the tool helps to visualize the state of the DC quickly, and the flexibility in assigning weights according to the type/use of the DC provides additional value to this tool, because the decisions makers (e.g. DC managers) have the ability to control the importance of each variable in the problem resolution in a transparent way, giving them the sense of ownership of the evaluation model.

Using these results, further analysis with more accurate data can be conducted to identify possibly improvements in the DCs. Indeed, one of the intangible results of this study was the knowledge transfer from the analysts to the company about the use of MCDA. The company wishes to use this approach for helping making architecture choices, in the planning phase for a new data center or a major renovation of an existing infrastructure. It can also help DC decision making processes related to Data Center site selection, e.g. looking for sites where can be increased the use of the heat recovered from the DC (e.g. swimming pools; greenhouses, etc). In general terms, we can also conclude that this study helped the company to better define the vision or goals of what should be a sustainable DC, identifying what must be assessed to evaluate a DC's environmental sustainability performance.

Although this study did not aim at selecting the best design for the new DC, it had an indirect impact through the learning process that occurred. The classification of the alternatives, although imprecise, helped the telecommunications company to understand which options would be important to obtain a good classification if this type of labeling was adopted by the industry. Initially the company was mostly concerned about PUE, but this study contributed to the emergence of other concerns leading it to pursue good performances in other criteria. For instance, the company is now considering architecture choices in order to improve the heat recovering from the DC. These efforts will be pursued even though this would not be sufficient to achieve the best category (despite the company's effort to have a strategy focused on energy efficiency and an IT resources optimization via increasing the virtualization and server utilization levels).

With more sustainable DCs, organizations can better manage the increased computing needs, i.e. they can meet the future business needs and at the same time lower their energy costs. The future poses serious challenges for DC managers, such as energy cost, water cost, carbon taxation, and general environmental concerns. Organizations that proactively focus on these issues will manage better the DC total cost of ownership and consequently reduce their business risks.

Our contribution with this tool is not to provide an absolute measure of the DC environmental sustainability, but instead to provide a way according to the specificities of each DC to assess their potential sustainability by addressing several issues and help organizations to determine strategies to improve DCs operational efficiency and reduce the impact on resources and environment. If new indicators appear that the DCs managers would like to analyze and incorporate, with this flexible tool we can easily adopt them if necessary.

The authors hope that this framework may help the industry to have a common understanding of Data Center Sustainability measurement, and can generate dialogue to improve it. An industry consensus about a sustainability assessment tool for DCs such as the one proposed here can yield several benefits: it can foster the promotion of sustainable DCs industry internationally, it can facilitate transparency and accountability by organizations and provides to stakeholders a universally applicable and comparable framework, from which one can understand disclosed information. Finally, such a framework could be a tool to communicate with customers, to help them to buy services from more eco-friendly Data Centers.

10.8 Acknowledgments

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 LARSYS and INESC Coimbra (PEst-C/EEI/UI0308/2014).

References

ASHRAE. ASHRAE environmental guidelines for datacom equipment. Technical report, 2008.

J. Boer. Sustainability labelling schemes : The logic of their claims and their functions for stakeholders. *Business Strategy and the Environment*, 12:254–264, 2003.

C. Böhringer and P. Jochem. Measuring the immeasurable - a survey of sustainability indices. *Ecological Economics*, 63:1–8, 2007.

C. Bratt, S. Hallstedt, K.-H. Robèrt, G. Broman, and J. Oldmark. Assessment of eco-labelling criteria development from a strategic sustainability perspective. *Journal of Cleaner Production*, 19(14):1631 – 1638, 2011.

Camara Municipal da Covilhã, 2011. URL `http://www.cm-covilha.pt/simples/?f=5443`. 11/12/2011.

L. Cherchye, W. Moesen, N. Rogge, and T. Puyenbroeck. An introduction to benefit of the doubt composite indicators. *Social Indicators Research*, 82:111–145, 2007.

R. Ciegis, J. Ramanauskiene, and G. Startiene. Theoretical reasoning of the use of indicators and indices for sustainable development assessment. *Inzinerine Ekonomika-Engineering Economics Journal*, 3:33–40, 2009.

L. Dias and V. Mousseau. IRIS: A DSS for multiple criteria sorting problems. *Journal of Multi-Criteria Decision Analysis*, 12:285–298, 2003a.

L. Dias and V. Mousseau. IRIS - interactive robustness analysis and paramenters inference for multi criteria sorting problems (version 2.0). Technical report, 2003b. User Manual, INESC Coimbra - Institute of Systems Engineering and Computers.

L. Dias, V. Mousseau, J. Figueira, and J. Clímaco. An aggregation/disaggregation approach to obtain robust conclusions with ELECTRE TRI. *European Journal of Operational Research*, 138(2):332–348, April 2002.