

This is a PDF file of an unedited manuscript that has been accepted for publication in *Sustainable Computing: Informatics and Systems*. 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/10.1016/j.suscom.2012.12.001>

On locating sustainable Data Centers in Portugal: problem structuring and GIS-based analysis

Miguel Trigueiros Covas^{a,e,*}, Carlos A. Silva^{b,e}, Luis C. Dias^{c,d,e}

^a Portugal Telecom

^b IDMEC, Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal

^c Faculty of Economics, University of Coimbra, Portugal.

^dINESC Coimbra, Coimbra, Portugal.

^eMIT- Portugal Program, Sustainable Energy Systems

**Corresponding address*: Instituto Superior Técnico (IST), MIT Portugal Program, Av. Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal. Tel: +351 210407022. Fax: +351 214233598

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

Abstract

Data Centers (DCs) are high energy intensive facilities due to the large power requirements of the computing equipment and the infrastructure needed to support it, in particular cooling systems. Nowadays, the capacity and consequently the energy demand of DCs is growing rapidly to meet the digitalization of the society. In order to meet sustainability concerns, the construction of a new DC must address the economic and environmental impacts of the different design options and one of the most crucial choices to be made concerns the location of the DC. This paper presents a brief overview of the current landscape of Portuguese DCs, based on the results of a survey. It also proposes a methodology to identify the best areas to locate a new DC in Portugal, using a new structure for evaluation criteria, based on inputs from industry experts. The relevant data was collected in a Geographical Information System (GIS) and successive filters were applied to progressively reduce the number of candidate sites. Finally, an analysis to the benefits of having DCs on those sites in comparison with their current locations was made, and the results show that the chosen sites to locate these facilities were not the best solution to achieve a more sustainable DC.

Keywords: Data Centers; Energy Efficiency; Site Location; Geographic Information System;

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 sectors. See for example the exponential growth of general internet communications at a rate of approximately 10 percent per year worldwide [1], the increased use of electronic transactions in the financial, commerce and services sectors, or the storage of electronic medical records, government information, etc. Data Centers (DCs) are critical infrastructures to develop this new paradigm of collaborative networked society. In particular, cloud computing - an emerging computational framework in which the applications, the data and IT resources are provided as a service to users over the network - is evolving rapidly and will be a new driver for DCs growth in coming years [2].

DCs are high energy intensive facilities due to the high power requirements of the computing equipment and the infrastructure needed to support the computing equipment, in particular its cooling systems [3] as shown in Figure 1. DCs are therefore more similar to industrial facilities in terms of energy consumption than commercial or service buildings. According to [4], the electricity demand of a DC can be 25 to 50 times higher than in standard office spaces with the same area. The worldwide DC power consumption in 2005 was already equivalent to the capacity of about seventeen 1.000 MW power plants [5] and the impact of the increasing usage of DCs in the energy sector is expected to be significant: the U.S energy consumption for DCs operation has been doubling every five years [1] up to 2007 and was expected to increase in Europe by 85% from 2007 to 2020 [6]. However, a very recent study [7] has shown that the energy growth rate at DCs not been as high as expected due to virtualization and economic crisis in US and Europe.

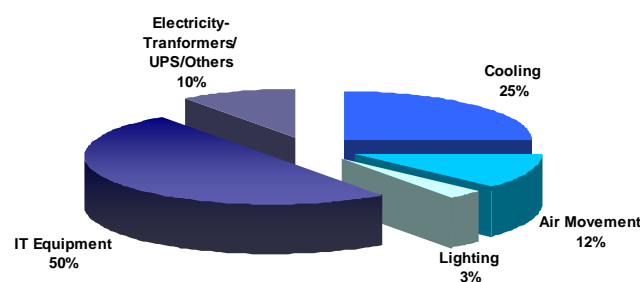


Figure 1 - Energy Usage in a Data Center [8]

To assess the energy efficiency of DC infrastructure (electrical and mechanical systems), the Green Grid, ASHRAE and the U.S Environmental Protection Agency, among others international organizations, have adopted the Power Usage Effectiveness (PUE) metric [9,10]. This metric is a ratio between the total energy consumed by a DC and the fraction of energy consumed by the IT infrastructure only. The result is always higher than one and the closer it is to one, the more efficient the DC is. Some organizations, e.g. European Commission [6], also use the inverse of the PUE ratio, called the DCiE (Data Center infrastructure Efficiency) metric. For example, a DC with the energy usage indicated in Figure 1 would present a PUE of 2 or a DCiE of 50%.

The direct energy use of IT and infrastructure equipment may also have a positive effect in the energy use [11]. Data processing and communication services provided by DCs lead to reductions in energy use in the broader economy, as for example e-commerce, digital invoices, teleworking or telemedicine, as it helps to reduce transportation energy use. Nevertheless, DCs have to be increasingly more efficient and greener, which requires reducing the energy needed to run computer infrastructures, ensuring longevity of the systems and ensuring energy consumption based on renewable energy sources [12].

Since DCs energy consumption is so high, the operation costs of the DC are mostly due to energy consumption. Thus, in a context of increasing energy costs, when a company decides to build a new DC, one of the main variables of the decision process is the energy cost and the availability

and reliability of electrical power, since these factors are highly dependent on the location of the facility. But these are not the only important variables. The security of data processing and storing and the reliability of the DC operation are also closely related to the location of the facility since it can be affected by the security of intrusion, the probability of occurrence of natural disasters, the access and reliability of telecommunications networks, the transportation accessibilities or labor availability. These are all critical aspects to the DC operation, thus, during the design phase of a DC, the choice of the location should follow a holistic approach that considers multiple criteria. However, the choice of a convenient location for a DC is an aspect often disregarded in the literature about green DCs, which usually focuses on technological hardware options or management software.

A Portuguese telecommunications operator is considering the construction of a new DC in Portugal with 12.000 m² of white space and 40MW of electrical power to support its increasing business activities. The goal is to build a state of the art facility, particularly in terms of energy efficiency, meaning that the DC is aimed to have a PUE equal to or lower than 1.25. This will be the largest DC ever built in Portugal and its design is expected to be used as a reference not only in terms of energy efficiency but also in terms of general sustainability. And this depends largely on finding the best possible location to attain all these objectives.

This paper proposes an approach to identify the best locations to build DCs that tackles economic, environmental and risk concerns, in order to aid decision makers on the planning process of a new DC. This approach, which involves structuring a hierarchy of objectives in interaction with relevant stakeholders, was applied in the design process of a new DC of a telecommunications operator in Portugal, by identifying the best set of parishes to build the facility, among a total of 4.050. Another contribution of this work is to present an assessment of the current landscape of DCs in Portugal based on a written survey and interviews, and to estimate the potential loss of opportunity in terms of energy savings due to their locations.

The paper proceeds as follows: Section 2 presents the main results of a survey done to the main DC operating in Portugal. Section 3 presents a literature review of the methodological approaches for the design of DCs and describes the sustainable design approach proposed in this paper. Section 4 describes the application of the proposed methodology to the Portuguese geographical reality, and evaluates the impact of using the proposed approach in terms of energy consumption. Section 5 draws the conclusions and presents the future research guidelines to improve the methodology.

2. Data Centers in Portugal: the current landscape

The construction of a new DC in Portugal that aims to be an international reference in terms of global sustainability of DCs requires that all the stakeholders involved in the design phase understand the country specificities regarding DC construction, operation and performance. To do

this assessment, a written survey with some follow-up interviews was done to the main DCs operating in Portugal Mainland.

The survey took place in the end of 2009 and was repeated at the end 2011. It was sent to a panel of eleven companies (national and international companies) from different sectors (telecoms, finance services, IT - ISP/ASP/hosting providers) and the respondents were IT directors or DC managers, covering a total of 17 DCs. In particular we asked to indicate the site location, the power requirements, and PUE of the facility. The survey covered also other features like the IT room area, the cooling technology, and the existence of energy efficiency plans.

The survey showed that most of the DCs are currently installed in office buildings (85%) and are all located in urban or suburban areas, mostly in Lisbon region (82%) and the remaining ones in Porto (18%).

In terms of power requirements, as described in Figure 2, more than one third (35%) are rated between 1MW to 5 MW and no DC requires more than 5 MW of power. This indicates that all the existing DCs in Portugal are relatively small.

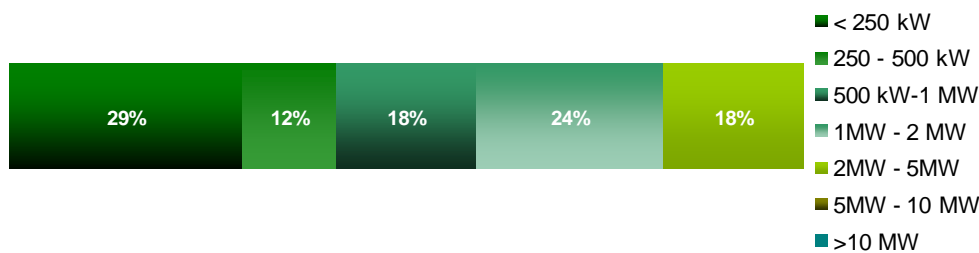


Figure 2 – Required power of the main Data Centers in Portugal (Base=17 Data Centers; percentages may not total 100% due to rounding)

In terms of energy efficiency, 41% are unfamiliar with or do not use the PUE measure. From the remaining ones, 23.5% have a PUE higher than 2. Overall, only 31% had a PUE lower than 2 and no DC in Portugal has a PUE lower than 1.25. Considering the DC Green Grid rating proposal (see legend in Figure 3) from the larger DC (power requirements >1 MW), only 14.5% of them are Bronze Data Centers, 28.5% are only Recognised Data Centers, and the remaining 57% are unfamiliar with or do not use the PUE measure. This indicates that the energy efficiency of the facility is in most cases unknown in detail, mostly due to the fact 85% of the surveyed DCs are located in office buildings, sharing the facilities with other services.

The most efficient DC surveyed is located in Lisbon region and has a PUE between 1.25 and 1.43. It is located in an office building and it is the only one that uses free cooling technology to reduce cold air generation needs. Figure 3 shows the location of the surveyed DCs and their classification using the Green Grid rating proposal.

The survey showed that the approach for DC construction in Portugal has been in general to retrofit office buildings in the main urban centers, and because of this - and despite the

implementation of some energy efficiency measures (e.g. cold aisle/hot aisle configuration) - most of the DCs surveyed cannot significantly improve further their efficiency. More than two thirds of the DCs have energy efficiency plans implemented, but more than half expected to increase their energy consumption over the next 12 months.

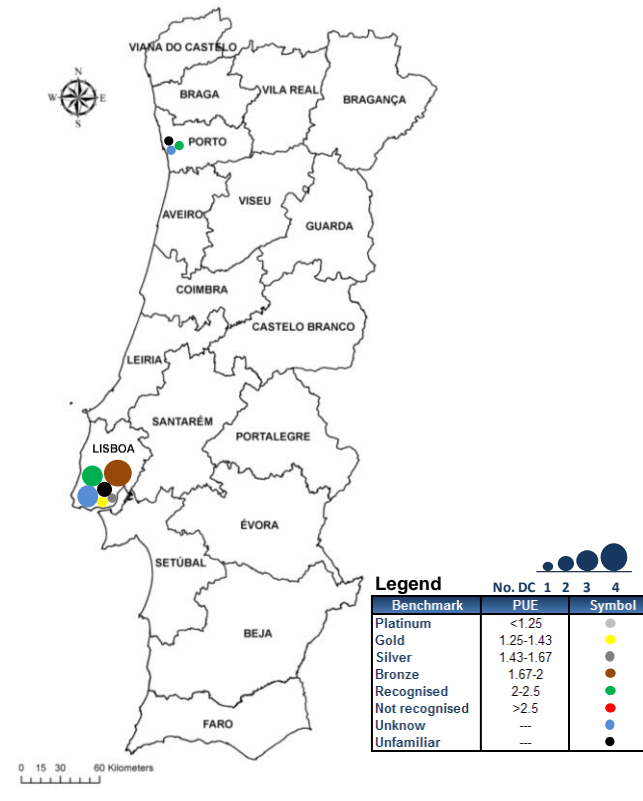


Figure 3 - Portugal Main Data Centers Locations and Efficiency

3. Design approach for a sustainable Data Center

Designing a DC involves several different activities, which include designing the layout of technical rooms and the location of rack cabinets, defining the cooling and air management systems and the power distribution level, and also choosing the type of servers, the backup power, the fire-protection and other safety systems.

Some of these activities - in particular the definition of cooling solutions or power distribution level - depend significantly on the location of the facility and a wrong site decision may affect the performance of the facility during the complete lifetime of the project, since there is no flexibility to change the facility or modify it significantly once it is built. Thus, it is essential to ensure that the site will fulfil not only the immediate facility needs, but especially future needs.

To address this problem, multiple criteria must be taken into account. Sun Microsystems [13] considers that site location should take into consideration natural hazards (e.g. seismic activity, tornados, hurricanes), man-made hazards (e.g. industrial pollution, vibration), the availability of utilities services (e.g. power supply, telecommunications) or even the emergency services and vehicle access. The Code of Conduct on Data Centers published by the European Commission

[14] established some criteria for DC location to promote the energy efficiency of the facility, which includes locations where waste heat of the DC can be reused, locations with low ambient external temperature and or humidity. The Uptime Institute [15], in addition to the above-mentioned criteria, also considers economic criteria such as utility consumption rates, financial incentives (property and sales taxes, rebates, reduced permit fees), etc. These criteria are generally referred also by companies like Cisco [16], Fortrust [17], Gartner [18], APC [19], Google [20], HP [21], IBM [22], and Microsoft [23].

The scientific literature about the DC location problem is still very scarce and the authors could only find two works that take into account some of the criteria recommended by industry and other organizations. Abbasov et al [24] presented an optimal location model to locate DCs taking into account different types of risks (natural, political and economic) while minimizing two objectives (cost and risk). Chang et al [25] 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 every DC is as small as possible. However, these papers do not address location criteria that have direct impact on the energy efficiency of the DC and consequently the impact on the economic performance of the DC.

In this paper, we extend this literature by proposing a design methodology to determine the location of a DC that filters and highlights the best locations to build a sustainable DC, taking into account many of the design criteria specified above. This methodology encompasses a structuring phase in which the relevant criteria are selected and available sources of information are identified. For this, it is important to involve DC experts and other possible stakeholders, as well as to follow structured guidelines for building a family of criteria [26-28]. A second phase consists in performing a search analysis using a Geographical Information System (GIS) that manages and displays the collected information. Although a GIS can be a good basis for implementing sophisticated multi-criteria decision aiding methods [29,30], it can also be simply used as a means to progressively narrow the set of acceptable alternatives (areas) by adding successive filters.

3.1. Search Analysis using GIS

There are many variables involved in the spatial problem of finding the best areas to locate a DC. One is the ambient temperature of the location, since this influences directly the design of the cooling system and thus the energy efficiency of the facility [31,32]. The type of electric grid is another important variable, as it impacts the reliability and the energy costs. Other important variables are the proximity to the main roads access and rail stations, which have impact on the accessibility, or the telecommunications networks, which have impact on the IT performance and reliability of the facility.

The variables of concern can be represented geographically and can be used to select and classify areas as to their aptness for the proposed objective, but can also be used to exclude those that are clearly inadequate areas. A Geographic Information System (GIS) model was developed for the purpose of helping to identify the most promising regions to build a DC taking into account

the various criteria involved in the selection process. The use a GIS application allows users to create interactive queries (searches made by users), to analyze spatial information, to edit data, to map, and to display the results of all these operations [33,34]. In this study we used the ArcGIS Desktop 9.3.1 (ArcEditor) software from the Environmental Systems Research Institute (ESRI) [35].

The search area consists of Mainland Portugal (88.971 km²), which is composed by 18 districts with 278 municipalities and 4.050 parishes. This study was performed at the parishes' level of detail.

3.2. Evaluation Criteria and Restrictions

In order to perform the site search analysis, a family of criteria had to be identified. This process was based on the study of the literature, the current best industrial practices, and discussions with DC professionals in Portugal. For each of the selected criteria, the choice of an acceptability threshold was also determined. These thresholds can of course be easily changed to more, or less, demanding levels.

The resulting criteria were structured along three main categories:

1. Environmental Criteria:

- a. Number of hours per year with average air temperature below 21°C;
- b. Availability of water resources like rivers or lakes;
- c. Conflicts with reserves and natural parks;

2. Economic Criteria:

- a. Proximity to telecommunications networks, defined by parishes intersected by a backbone node (a 4km radius was considered);
- b. Proximity to electrical power, defined by parishes intersected by a power substation with high voltage/medium voltage conversion (a 4km radius was considered);
- c. Proximity to the main road access (highways, main itineraries and additional itineraries), defined by parishes intersected by a main road access (a 10km radius was considered);
- d. Proximity to railway stations, defined by parishes intersected by a railway station (a 4km radius was considered);

3. Risk Criteria:

- a. Seismic intensity, defined by the Modified Mercalli Scale;
- b. Nuclear risk, defined by parishes within a security perimeter for a nuclear accident established by the Portuguese National Protection [36] and the International Atomic Energy Agency (IAEA) safety standards [37] (a 30km radius was considered).

In the following sections we describe in detail each of these criteria. One can note that other important criteria could be taken into account, such as the price of real estate, financial subsidies, and other aspects. Such concerns can be incorporated later to distinguish among the different

areas deemed to be the most promising (in fact, some of these omitted attributes might be negotiable on an case by case basis). In this study, the option has been to consider attributes that are less likely to be negotiated or to change as time passes.

3.2.1 Environmental Criteria

As seen in Figure 1, the second largest energy use in a DC, right after the IT load, is the cooling and ventilation system. According with the ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers) [38] and the Best Practices for the EU Code Of Conduct On Data Centers [14] recommendations, DCs should be designed to operate at "inlet" temperatures between 18 °C and 27 °C, to maximize energy efficiency. The relative humidity (RH) range is limited to less than 60% with lower and upper dew point temperatures of 5.5°C and 15°C.

If the local climate is suitable, outside air can be directly employed to provide part or all cooling needs to the DC. This cooling solution, typically called air free cooling or compressor-less cooling, is based on the natural convection phenomena (usually using an air-side economizer [31]) and requires much less energy than a conventional cooling system (compression chiller), as it uses the ambient external air to cool the servers eliminating the need to chill the air. The opportunities for the utilization of free cooling solutions increase the efficiency of the facility and enable reductions up to 50% in the energy consumption of cooling system [39].

Following these recommendations and considering that usually the temperature gain in using free cooling technology is around 3°C [40], this study considers the number of hours per year where average air temperature is equal or below 21°C is an indicator of the number of hours per year that the DC could operate without conventional cooling systems (chillers). With this approach, we can ensure "inlet" temperatures around 24°C and also provide a safety band (for maintenance or repair possible faults without putting is risk the availability of the DC) taking into account the 27°C limit in this analysis.

The relative humidity (RH) is also a very important factor, since high RH environments can cause corrosion of the hardware circuits and low RH environments can cause electrostatic discharges. However, this parameter can be easily controlled by free cooling systems. If the air-side economizer system is indirect, the outside air is never mixed with the inside air. The outside air is passed through a heat exchanger to cool down inside air and then is returned to the outside, while the cool inside air returns to the servers' hall. This solution reduces the risk of outdoor air pollutants adversely affecting the hardware [41]. In this case, the outside air RH does not affect the inside air RH, which makes it easier to maintain more stable moisture levels in the data hall [40]. If the air-side economizer system is direct, the outside air is directly used as the cooling medium, after passing through a filtering system into the servers' hall. In this case, when the ambient air humidity is higher than the admissible threshold, the outside air is mixed with inside air to decrease the RH, which avoids the need to implement a dehumidification process; when the RH of the outside is too low, evaporative processes must be applied to increase the levels of RH.

Therefore, it was considered unnecessary to explicitly consider the relative humidity of the region as a criterion.

Another environmental factor that should be considered is the distance to shoreline that should be higher than 10Km to avoid the potentially corrosive coastal air [41]. The implementation of such criterion to a country like Portugal, where half of its border is a shoreline with the Atlantic Ocean, would eliminate a large number of potential locations. Moreover, the locations that would be eliminated contain most of the existing DCs in Portugal (as depicted in Figure 3), which do not suffer from costal air problems. The HP/EDS Data Center in Wynyard Park, Middlesbrough, UK is another example of a DC that uses direct free cooling with a distance to the shoreline smaller than 10 Km [21]. The influence of the costal air depends on several factors, such as wind speed and direction, land topography, and it can be mitigated by using heat exchangers. For these reasons, the 10Km limit was not considered in this study. Nevertheless, in the results section we highlight the parishes located at distances higher than 10 Km from the shoreline.

As far as other pollutants, as they depend significantly on neighbor activities, we considered in this study that they cannot be analyzed at the parish level and should be considered at a later stage of analysis, when precise sites are to be evaluated.

In regions where it is not possible to achieve 100% of air free cooling during the year, the availability of water resources to assist cooling systems is also very important. In this case, water is used in adiabatic cooling processes that reduce air temperature, which increases the free cooling potential. In Portugal the use of water resources for cold purposes is allowed, thus places with rivers or lakes are potentially places with higher potential to use free cooling systems.

The existence of reserves or natural parks takes into account the fact that in general it is not possible to build any type of facility in these locations, which means that parishes within or crossed by reserves and natural parks must be avoided.

Finally, another environmental criterion that could be considered is the availability of green power in the region. In Portugal, there is only one transmission system operator and one distribution system operator, and all electricity generation facilities are connected to these regulated operators. This means that the CO₂ emission factor is the same for all parishes, so this criterion was not considered in this study.

3.2.2 Economic Criteria

In terms of economic criteria, one of the most important factors is the availability and proximity of telecommunications operators' backbones that guarantee high bandwidth capacity and low latency. Usually DCs need high bandwidth capacity, albeit in some cases low bandwidth capacity circuits are also required (e.g. ADSL circuit for maintenance purposes, or co-location customers). The availability of a high bandwidth infrastructure reduces significantly the investment costs in

communication infrastructure (e.g. fiber optical networks) and operating costs (e.g. telecommunications circuit's fees). For distances over 4 Km [42] from the telecommunications operator node, it becomes very difficult to use xDSL technologies (e.g. VDSL), since is not possible to ensure broadband. Therefore this is the maximum distance considered under this criterion to ensure the proximity and availability of high bandwidth capacity.

Another critical criterion is the availability of electrical power. The proximity to the electrical grid reduces in general the investment cost (the cost of a high voltage electrical line can exceed 150 K€/Km in Portugal). The proximity to a very high voltage level connection can represent further reductions, since the costs of energy transport and delivery are lower, which has a direct impact on the operation cost of the DC.

The proximity to main roads access is also important since it facilitates the access of customers, DC's staff and suppliers (needed to ensure service level agreements). The considered threshold is 10 Km, which represents a 10 minutes' drive from a main road at an average speed of 60 Km/h.

The proximity of public transportation (train and/or bus) is important to employees and suppliers. 4 Km represents a 10 minutes travel by bus at an average speed of 25 Km/h. In addition to the above criteria, the proximity of airports can be also considered an important factor for locating DCs, since it facilitates the access from customers. In this case, a limit distance of 200 Km could be considered, representing a 2 hour drive at an average speed of 100Km/h. Since Portugal is a small country (the maximum length is 561 Km and the maximum width is 220 Km), with 3 commercial airports (Porto in the north, Lisbon in the center and Faro in the South), all country would be covered, thus this criterion was not applied in this study.

3.2.3 Risk Criteria

In terms of risk criteria, we should consider not only the possibility of natural disasters (e.g. tornadoes, hurricanes, earthquakes, etc.) but also disasters related to human activities (e.g. nuclear accident).

Natural hazards like hurricanes, tornados and volcanic activity were not considered, since there are no records of its occurrence in Portugal. The main natural disaster that may happen in Portugal are earthquakes (and tsunamis on costal areas), which can destroy buildings and equipments or cause power failures, damage roadways, human injury or death. For this, we have used the expected intensity of an earthquake measured using the modified Mercalli Scale, which scales the earthquake intensity in terms of effects on the Earth's surface, humans, objects of nature, and man-made structures on a scale from I (not felt) to XII (total destruction). The Lisbon area and the south of Portugal are the areas with the higher expected intensity earthquakes (X in the Mercalli scale).

A nuclear accident can cause the contamination of water and air due to the release of

radioactivity and thereby stop the DC operation, thus the DC should be located outside of the evacuation zone of a nuclear accident (30km according to the IAEA [37]).

3.3. Support Information

The application of these criteria on a GIS requires that each of the criteria is described by an assessment map. An assessment map represents the spatial distribution of an attribute and is used to evaluate the performance of alternatives to the objective associated with this attribute.

For the Portuguese administrative official areas we used the Portuguese Administrative Cartographic (CAOP) version 2009 [43], published by the Portuguese Geographic Institute, which describes the official limits of the country (the boundaries of the districts, municipalities and parishes).

The data for the criteria came from the following sources:

1. Number of hours per year with average air temperature below 21° C (Source: Instituto Meteorologia, www.meteo.pt)
2. Hydrographical network (Source: Instituto do Ambiente, www.iambiente.pt)
3. Reserves and natural parks (Source: Instituto do Ambiente, www.iambiente.pt)
4. Telecommunications networks (Source: PT Comunicações, www.ptwholesale.telecom.pt)
5. Power substations (Source: EDP, www.edp.pt)
6. Main roads access (Source: Estradas de Portugal, www.estradasdeportugal.pt)
7. Railway stations (Source: Refer, www.refer.pt)
8. Seismic intensity (Source: Instituto do Ambiente, www.iambiente.pt)

4. DC location analysis

When we have a large number of options to locate a DC - in this case thousands of parishes - a common strategy is to eliminate all that do not meet a minimum level of performance according to certain criteria. In this work we applied the conjunctive rule [44], i.e. the parish is deemed to be interesting for the DC location if and only if all of its established conditions are true. This decision making rule establishes a minimum threshold of satisfaction for each criterion, retaining only the options that are equal to or better than these thresholds.

4.1 Methodology application using a GIS

The first step consisted in characterizing the possible domain of solutions, in this case parishes, using the CAOP 2009 (Portuguese Administrative Official Areas) layer. The following steps consisted in applying successively each of the criteria described in the previous section. By adding one condition at a time it is possible to observe how stringent its requirements are in terms of the number of parishes that are eliminated from consideration. Let us note at this moment that the acceptability thresholds will be later subject to a sensitivity analysis.

4.1.1 Environmental Criteria

The second step consisted in representing the map with parishes that exceed 7.500 hours per year with average air temperature below 21° C. In this step we obtained 1.326 parishes (Figure 4). The average number of hours with air temperatures below 21° C for all the parishes in Portugal is about 7.300 hours per year. The 7.500 hours is above the average and represents approximately 313 days of the year (85%).

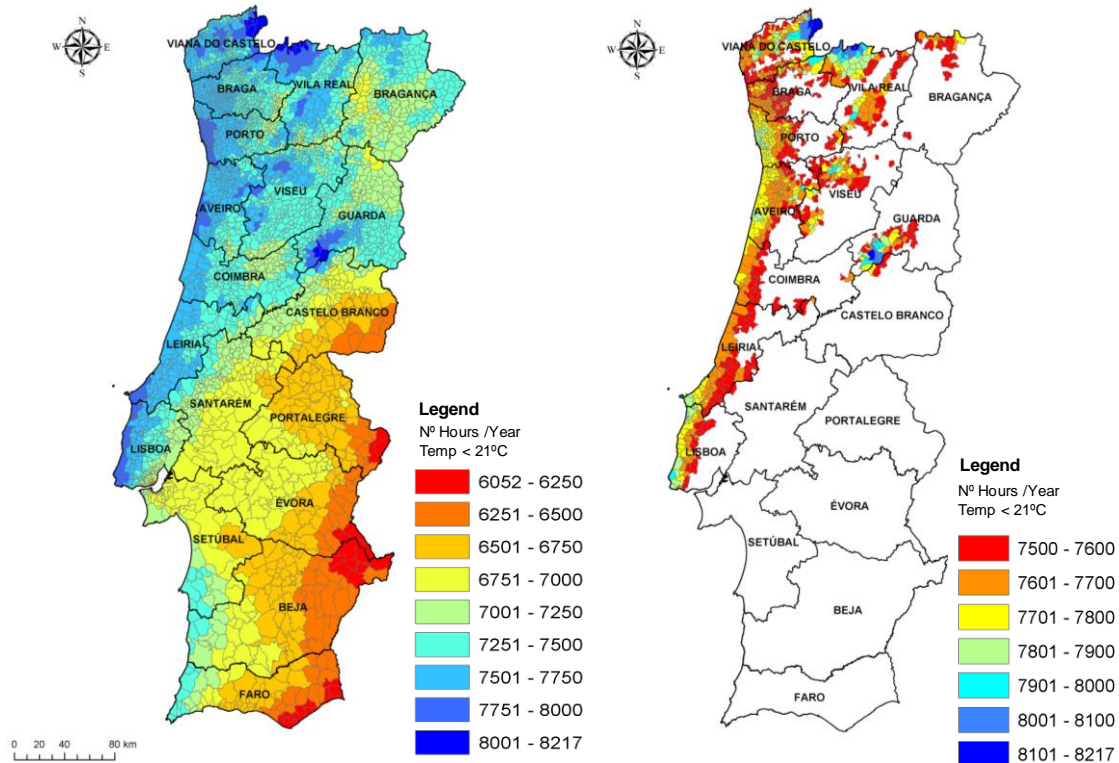


Figure 4 - The first map represents the number of hours per year with average air temperature below 21° C (Source: Instituto Meteorologia. www.meteo.pt). The second map represents the parishes that have in excess 7.500 hours per year with average temperature below 21° C.

In the third step we included the hydrographic network layer (Portuguese main rivers, hydroelectric reservoirs with flooding areas above 50 ha). In this step, the number of available parishes was reduced to 276 parishes (Figure 5).

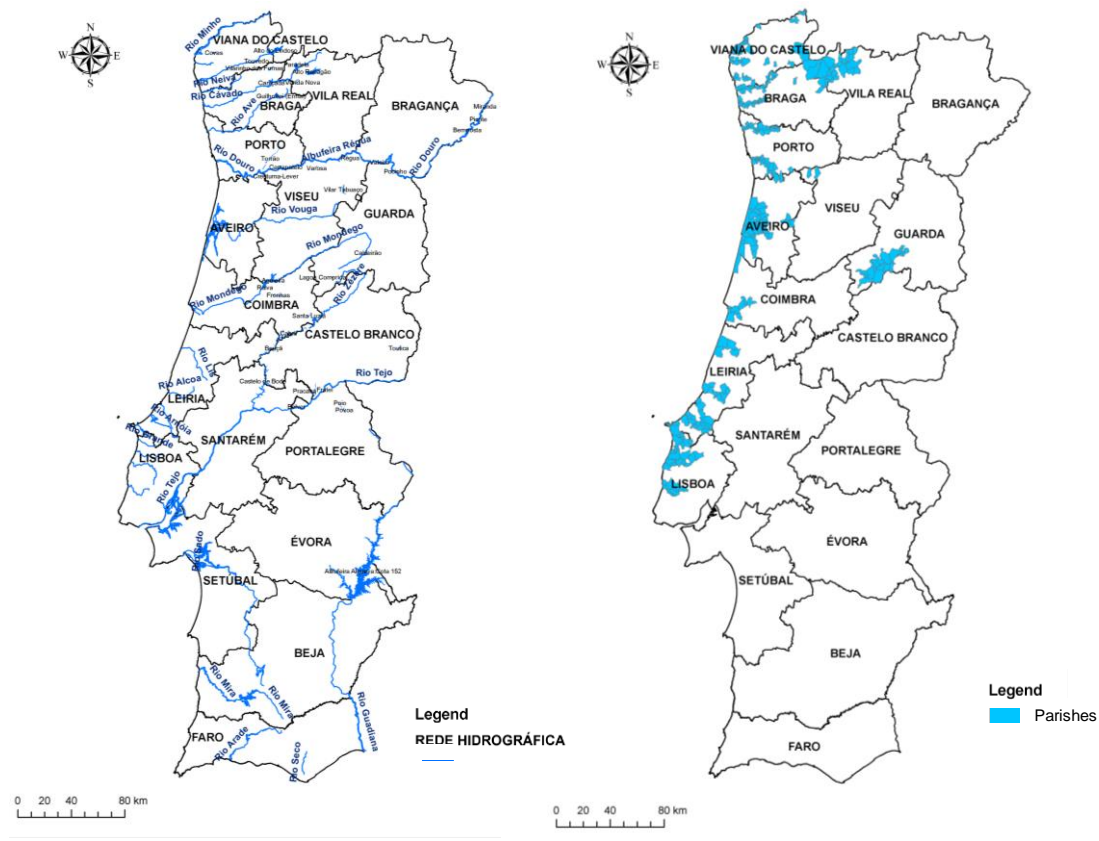


Figure 5 - The first map represents the hydrographic network layer (Source:IA). The second map represents the parishes within/intersected by a main river or reservoir.

In the fourth step we excluded areas within reserves and natural parks and the number of parishes was reduced to 231.

4.1.2 Economic Criteria

The fifth step consisted in evaluating the telecommunications networks layer (backbone nodes), i.e. parishes within or intersected by a backbone node 4km radius. In this step the number of parishes was reduced to 70 parishes. The sixth step consisted in evaluating the existence of a power substation (High Voltage or Medium Voltage) within a 4 km radius. In this step, the number of parishes remained the same.

In the seventh step, we evaluated the parishes with main roads access (highways, main itineraries and additional itineraries) and established as threshold the existence of a 10 km radius to a main road. Again, none of the 70 parishes from the previous step was excluded. In the eighth step, we considered the existence of parishes within or intersected by a railway stations 4km radius. After this step we obtained only 62 parishes.

4.1.3 Risk Criteria

In the ninth and final step, we included the seismic intensity layer allowing this information to be taken into account, but without excluding any region. Regarding the nuclear risk all parishes are

outside of an evacuation zone (distances over 30 km), since there are no nuclear power plants operating in Portugal and the closest in Spain is Alamaraz (100Km away from the closest Portuguese border).

4.2. Results

Figure 6 shows the evolution of the numbers of parishes resulting from the several steps previously described. It can be seen that the most stringent criteria were the number of hours below 21°C, followed by the existence of water resources and finally the existence of telecommunication facilities. At the end of this process there are 62 parishes (36 parishes located less than 10 Km from the shoreline depicted in light blue, and 26 parishes located at a distance higher than 10 Km from the shoreline depicted in dark blue, see Figure 7) that are deemed to be interesting to develop an efficient DC in Portugal.

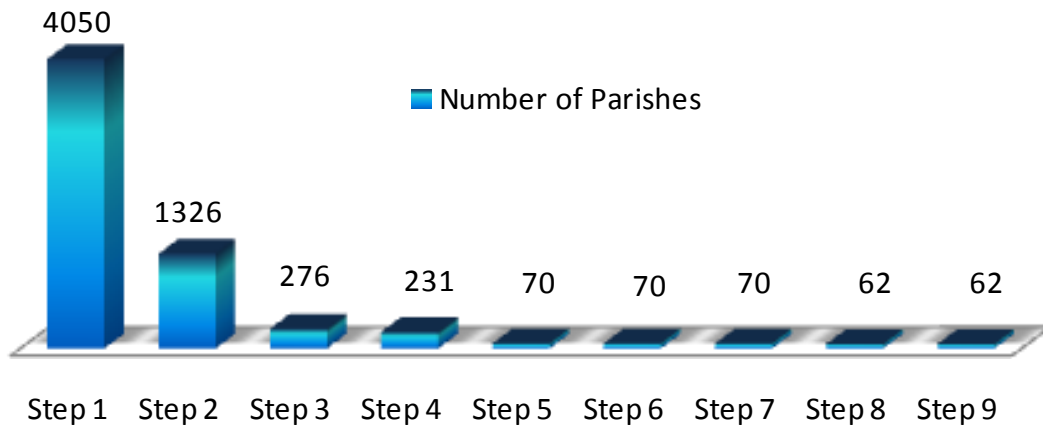


Figure 6 - Parishes Evolution from the several steps.

The best regions to locate a DC are those north of Lisbon and near the coastline. These regions are distributed over 7 districts (see Figure 7). In terms of seismic intensity these regions are between level VI (strong) and level IX (violent). The parishes that best meet the criteria of seismic intensity and temperature are located in Porto district.

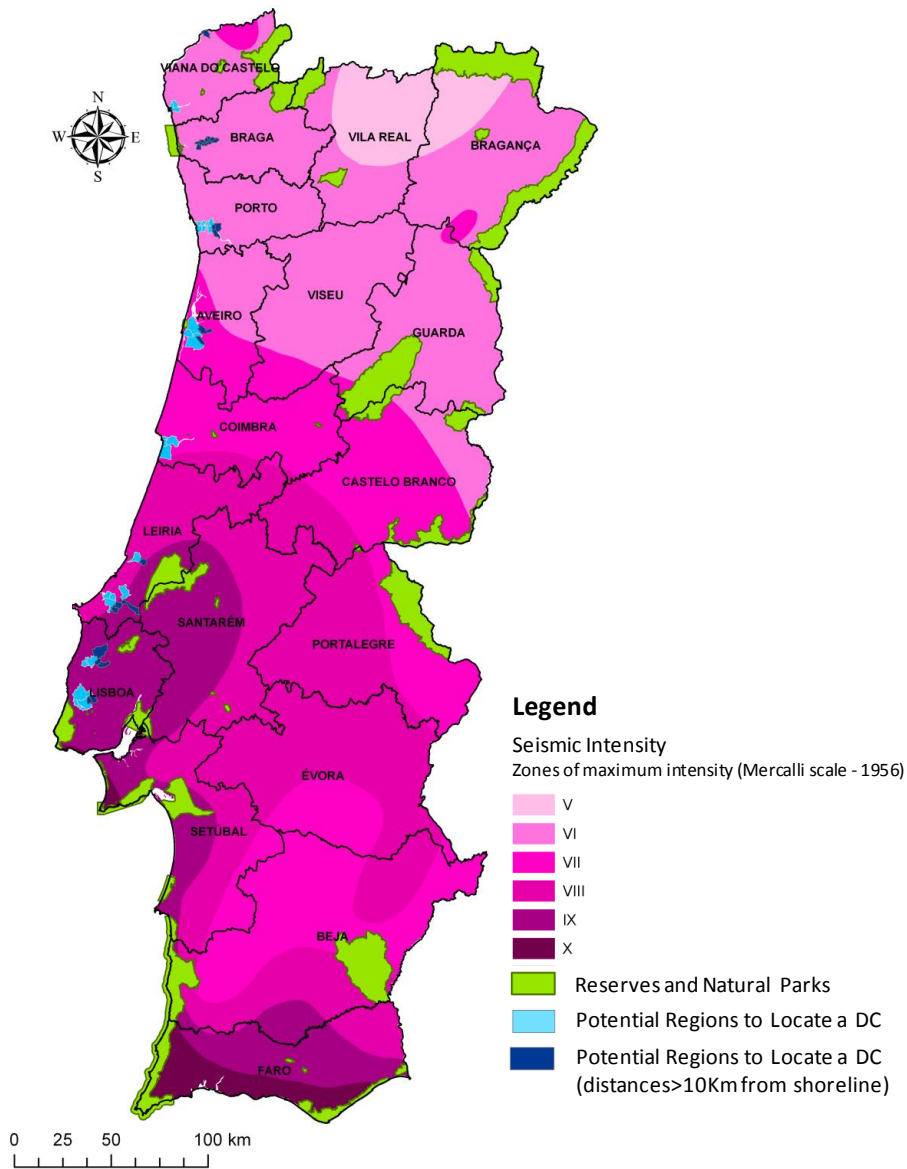


Figure 7 - Regions in Portugal that are potentially suitable to develop an efficient Data Center.

4.3. Sensitivity Analysis

A sensitivity analysis was performed in order to determine how the model results could change considering some variability in the considered thresholds. The aim is to evaluate if any interesting alternative has been eliminated by the imposed restrictions. The most questionable thresholds are those related with the Economic category. Among these, the proximity to a power substation and the proximity to main roads had not contributed to exclude any alternative. The following scenarios were developed for the remaining two thresholds in the Economic category:

Scenario 1: Increase of 20% of the of the backbone node radius (from 4 km to 4.8 km).

Scenario 2: Increase of 20% of the of the railway stations radius (from 4 km to 4.8 km).

Scenario 3: Considering the previous cases together.

Ten new parishes arose from this sensitivity analysis: 6 parishes from the scenario 1, 2 new parishes from the scenario 2 and another 2 parishes from scenario 3. Geographically, these new

regions are adjacent to regions already obtained in the modeling approach, which indicates that the latest criteria did not exclude any good environmental alternative due to the use of a very stringent economic criterion. For example, regarding the criterion number of hours per year with average temperatures below 21 °C, the maximum gain that can be obtained in these 10 parishes comparing with the average value of the 62 parishes (7.657 hours per year) is less than 1%. In this way, we can conclude that no good alternative has been excluded and therefore these ten parishes were not considered as potential alternatives.

4.4. The impact of the location site in the DC energy efficiency

According to Figure 1, the cooling system can represent a significant part of the consumed energy in the DC, right after the consumption of the IT equipment itself. As previously said, taking advantage of the opportunities to utilize free cooling solutions increases the efficiency of the facility, since it allows up to 50% reductions in the energy consumption of the cooling system. For the example in Figure 1, this could represent the difference between having a DC with a PUE of 2 or having a DC with a PUE of 1.63.

Considering the assessment of the main DCs in Portugal presented in Section 2, most of the analyzed DC are nearby the Lisbon region (85%). Taking into account the results presented in the model, a DC located in the north of Portugal has higher potential to be more efficient than one located in the Lisbon region. From this information, we can draw two conclusions. The first is that DCs built in northern Portugal are likely to be more sustainable than the ones built in other regions. The second conclusion is that DCs in northern Portugal can be more efficient than they currently are, namely if they profit from free cooling solutions – Figure 3 shows that the DCs in Porto perform worse than the DCs in Lisbon.

It is also possible to estimate the opportunity loss of not having chosen the best locations in the past. Consider a DC with 3.000 m² with a power density of 2.1 KW/m² (the average value described in [45]) that uses indirect air free cooling aided by a chiller system. Consider further that the coefficient of performance (COP) for the cooling system using chillers is 3.5 and for the free cooling solution is 10 [40]. Finally, assume that in Portugal, the emission factor is 0.47 Kg CO₂ /KWh [46] and that the average energy cost is 70€/MWh.

If this exact DC is located in the north of Portugal, it would consume less 995 MWh per year than one with the exact same characteristics located in Lisbon. For a DC with a lifetime of 15 years, this represents less 14.9 GWh energy demand, less 7018 ton/CO₂ emissions and operation costs savings of 1.045 M€, which demonstrates clearly that the choice of location of a DC has a significant environmental and economic impacts in the total lifetime of the project. If we extend this theoretical exercise to the data collected in the survey described in Section 2, we can conclude that if all the 17 DCs analyzed in the study operated in Porto using free cooling systems, the annual energy savings could have represented 45.8 GWh, 21567 ton/CO₂ avoided emissions and operational costs savings of 3.212 M€.

5. Conclusions and future research lines

The decision regarding the location to build a DC is very important and there are several key issues that must be addressed, since there is no flexibility to change or modify the site once it is selected. In this paper, we presented a new methodology to help finding suitable locations to build sustainable DCs, based on the GIS location search methodology and apply it to determine the best locations at the parish level in Portugal mainland.

From a total of 4.050 parishes, it was possible to narrow down the set of potentially interesting parishes to build a sustainable DC to 62 parishes only, which represents 1.5% of the total initial set of alternatives. Geographically, these potential regions are at north of Lisbon and near the coastline (where there are a higher number of hours per year with average temperatures below 21 °C). The developed model streamlines the selection process by reducing the search space and consequently minimizes costs and time spent in the search process.

This set of parishes does not include the parishes where the majority of the main DCs operating in Portugal (85%) are currently installed. From a survey done to a set of the largest DCs currently operating in Portugal mainland (from 11 different companies in the telecoms, finance and IT providers sectors), most of DCs in Portugal are located in the Lisbon urban area, in office buildings that have been retrofitted for that purpose and perform poorly in terms of energy efficiency. We concluded then that future DCs in Portugal should be built in northern Portugal and that the existing DCs in this area are not taking advantage of their location to be more efficient.

Using these results, a new set of criteria (e.g. proximity of emergency services, availability of skilled labor, land cost, taxes, renewable energies, etc) can be evaluated in order to rank the 62 parishes by order of interest for the location of the DC, using a Multi-Criteria Decision Making method [47,48]. After the best parish from the 62 is identified, a more detailed evaluation can be made at a smaller spatial scale in order to determine the exact best site for the DC location.

Acknowledgements

The authors would like to thank the reviewers for their detailed comments and suggestions. 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.

References

[1] Report to Congress on Server and Data Center Energy Efficiency, US Environmental Protection Agency Public Law 109-431, 2007.

- [2] R. Buyya, et al., Cloud Computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, *Future Generation Computer Systems*, 25 (2009) 599-616.
- [3] B. Shi, A. Srivastava, Unified datacenter power management considering on-chip and air temperature constraints, *Sustainable Computing: Informatics and Systems*, 1(2011) 91–98.
- [4] High Performance Data Centers – A Design Guidelines Source Book, Pacific Gas and Electric Company, 2006.
- [5] J. Koomey, World electricity used in data centers, *Environmental Research Letters*, IOP Publishing, 2008.
- [6] Code of Conduct on Data Centres Energy Efficiency Version 1.0, European Commission, 2008.
- [7] J. Koomey, Growth in Data center electricity use 2005 to 2010, Analytics Press, 2011.
- [8] Five Strategies for Cutting Data Center Energy Costs Through Enhanced Cooling Efficiency, Emerson, 2007.
- [9] Harmonizing Global Metrics for Data Center Energy Efficiency, Energy Star, 2012, http://www.energystar.gov/index.cfm?c=prod_development.server_efficiency.
- [10] Recommendations for Measuring and Reporting Overall Data Center Efficiency, Version 2 – Measuring PUE for Data Centers, Energy Star, 2012.
- [11] SMART 2020: Enabling the low carbon economy in the information age, The Climate Group, 2008.
- [12] S.K.S. Gupta, T. Mukherjee, G. Varsamopoulos, A. Banerjee, Research directions in energy-sustainable cyber-physical systems, *Sustainable Computing: Informatics and Systems*, 1 (2011) 57-74.
- [13] R. Snevely, Enterprise Data Center Design and Methodology, Sun Microsystems, 2001.
- [14] 2010 Best Practices for the EU Code of Conduct on Data Centers Version 2.0, European Commission, 2010.
- [15] P. Turner IV, J.H. Seader, V. Renaud, Operational Sustainability and Its Impact on Data Center, Uptime Performance, Investment Value, Energy Efficiency and Resiliency, Uptime Institute, 2008.
- [16] D. Alger, Build the Best Data Center Facility for Your Business, Cisco Press, 2005.
- [17] Geographic Factors for Data Center Site Selection, Fortrust, 2008.
- [18] Data Center Facility Location Selection Criteria, Gartner, 2009.
- [19] Site Selection for Mission Critical Facilities, APC, 2007.
- [20] Google, <http://www.google.com/datacenter/lenoir/faq.html>, 5/01/2011.
- [21] D. Middleton, Wynyard Park Data Centre Programme Overview, EDS/HP internal report, 2009.
- [22] Think Green Data Center Planning and Design, IBM internal report, 2009.
- [23] J. Dwyer, Microsoft Data Centres: The Next Generation, Microsoft internal report, 2009.
- [24] A. M. Abbasov, I. M. Aliev, L. H. Kerimova, Optimal Location of Internet Data Centers Taking into Account the Risks, *Automatic Control and Computer Sciences*, 43 (2009), 309–316.
- [25] S.Chang, S. Patel, J. Withers, An Optimization Model to Determine Data Center Locations for the Army Enterprise, IEEE, 2007.

- [26] Bouyssou, Building criteria: A prerequisite for MCDA. Readings in multiple criteria decision aid, Springer, 1990, pp 58-80.
- [27] R.L. Keeney, Value-Focused Thinking, Harvard University Press, Cambridge, 1992.
- [28] B. Roy, Multicriteria Methodology for Decision Aiding, Kluwer Academic Publishers, 1996.
- [29] Malczewski, J. GIS-based multicriteria decision analysis: a survey of the literature, International Journal of Geographical Information Science, 20 (2006) 703–726.
- [30] S. Chakhar, V. Mousseau, Spatial Multicriteria Decision Making, Encyclopedia of GIS. Springer, 2008, pp.753-758.
- [31] A. Shehabia, et al., Data center design and location: Consequences for electricity use and greenhouse-gas emissions, Building and Environment, 46 (2011) 990-998.
- [32] K. Kant, Data center evolution. A tutorial on state of the art, issues, and challenges, Computers Networks, 53 (2009), 2939-2965.
- [33] A.T. Murray, Advances in location modeling: GIS linkages and contributions, Journal of Geographical Systems, 12 (2010), 335-354.
- [34] J. Malczewski, “GIS-based land-use suitability analysis: a critical overview”, Progress in Planning, 62 (2004) 3-65.
- [35] ESRI, Environmental Systems Research Institute, <http://www.esri.com>.
- [36] ANPC, Autoridade Nacional Protecção Civil
www.proteccaocivil.pt/PrevencaoProteccao/RiscosTecnologicos/EmergenciasRadiologicas/Pages/CentraIsNuclearesnaEuropa.aspx, 5/01/2011.
- [37] Arrangements for Preparedness for a Nuclear or Radiological Emergency. IAEA Safety Standards (International Atomic Energy Agency), 2007.
- [38] 2008 ASHRAE Environmental Guidelines For Datacom Equipment, ASHRAE, 2008.
- [39] W.M. Kok, Three steps to perfect green, Uptime Institute, Symposium 2009, 2009.
- [40] K. Dunnavant, Indirect Air-Side Economizer Cycle Data Center Heat Rejection, ASHRAE Journal, (2011) 44-54.
- [41] Gaseous and Particulate Contamination Guidelines for Data Centers (White Paper), ASHRAE, 2009.
- [42] ANACOM, Autoridade Nacional de Comunicações,
<http://www.anacom.pt/render.jsp?contentId=891687>, 5/01/2011.
- [43] “CAOP 2009” - Carta Administrativa Oficial de Portugal (CAOP) de 2009, IGP- Instituto Geográfico Português, <http://www.igeo.pt>.
- [44] C.L. Hwang, K. Yoon, Multiple Attribute Decision Making, Springer-Verlag, Berlin, 1981.
- [45] H.S. Sun, S.E. Lee, Case study of data centers energy performance, Energy and Buildings, 38 (2006) 522-533.
- [46] Despacho nº 17313/2008 de 26 Junho, Republica Portuguesa, INCM, Diario da Republica 2ª Série, nº 122, 2008.
- [47] V. Belton, T. Stewart, Multiple criteria decision analysis: an integrated approach, Kluwer Academic Publishers, 2002.
- [48] D. Bouyssou, et al., Evaluation and decision models with multiple criteria: Stepping stones for the analyst, Springer, 2006.