Unraveling electricity consumption profiles in households through clusters: Combining smart meters and door-to-door surveys

João Pedro Gouveia *, Júlia Seixas

CENSE – Center for Environmental and Sustainability Research, Department of Science and Environmental Engineering, Faculty of Science and Technology, Universidade NOVA de Lisboa, Portugal

A R T I C L E   I N F O

Article history:
Received 15 September 2015
Received in revised form 14 December 2015
Accepted 30 January 2016
Available online 1 February 2016

Keywords:
Daily electricity consumption
Hierarchical clustering
Consumers’ segmentation
Smart meters
Household surveys
Electricity demand management

A B S T R A C T

Improvements of energy efficiency and reduction of Electricity Consumption (EC) could be pushed by increased knowledge on consumption profiles. This paper contributes to a comprehensive understanding of the EC profiles in a Southwest European city through the combination of high-resolution data from smart meters (daily electricity consumption) with door-to-door 110-question surveys for a sample of 265 households in the city of Évora, in Portugal. This analysis allowed to define ten power consumption clusters using Ward’s method hierarchical clustering, corresponding to four distinct types of annual consumption profiles: U shape (sharp and soft), W shape and Flat. U shape pattern is the most common one, covering 77% of the sampled households.

The results show that three major groups of determinants characterize the electricity consumption segmentation: physical characteristics of a dwelling, especially year of construction and floor area; HVAC equipment and fireplaces ownership and use; and occupants’ profiles (mainly number and monthly income).

The combination of the daily EC data with qualitative door-to-door survey-based data proved to be a powerful data nutshells to distinguish groups of power consumers, allowing to derive insights to support DSOs, ESCOs, and retailers to design measures and instruments targeted to effective energy reduction (e.g. peak shaving, energy efficiency).

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Greenhouse gases (GHG) emissions will hold steady or might even increase in developed countries if effective reduction of energy consumption will not be taken [1], contrary to policy goals aiming a transition toward low carbon economies. The need for energy consumption reduction is also linked to energy supply security and affordability, and climate change strategies. Therefore, increased search for energy efficiency, GHG emissions reduction and increased share of renewable energy sources, as established in the European Union (EU) goals by 2030 [2] requires more successful and directed actions.

Energy consumption in residential buildings deserves special attention since they represent a significant share of final energy consumption in OECD (Organization for Economic Co-operation and Development) countries, 27% in EU28 in 2013 [3]. In Portugal, residential buildings consume approximately one third of total electricity, with a growth of 70% from 1995 to 2012 [4]. This consumption is a complex issue that can be explained by a combination of physical, technological, demographic, climatic and behavioral characteristics of a dwelling and its occupants.

Understanding the determinants that govern energy consumption has thus been the subject of abundant international literature for more than 30 years (e.g. [5,6–8]). More recently, Jones et al. [9] presented a literature review of the existing research investigating the socio economic, dwelling and appliance related factors that affect electricity consumption in the residential sector.

In this area of study, smart meters have been gaining higher interest in demand side management initiatives and for utilities, and are seen as an important instrument for giving energy consumption feedback to households and for consumers’ profiles analysis [10]. With growing deployment of smart meters and
real-time home energy-monitoring services, adequate data allowing the study of electricity consumers’ profiles in households and its the determinants are becoming available.

Hayn et al. [11] worked on daily electricity household profiles through segmentation based on lifestyles, socio demographic factors, and electric appliances and on new technologies for heat and electricity generation. Crossed the information delivered by the smart meters with the main determinants of energy consumption in each household, allows for market segmentation identifying consumers with similar needs and behaviors [12]. This valuable knowledge promotes individually and personalized feedback evaluation to households or groups of similar consumers being important for energy savings. Furthermore, tailoring of energy efficiency measures based on specific household profiles enables the change of behavior and equipment with the ultimate goal of an effective energy consumption reduction and load curve consumption peaks minimization.

There are studies on the residential electricity consumption profiles using smart metering data. Seo and Hong [13] with a 30 households sample for Daegu in South Korea characterized power consumption patterns and presented summer load profiles. Rhodes et al. [51] using 103 homes for Austin in Texas determined representative residential electricity use profiles within each season drawing correlations to the different profiles based on survey data. Lee et al. [14] demonstrated profiles of electricity consumption for 60 low energy-housing houses in South Australia. Ramos et al. [15] identified daily load profiles of medium voltage customers applying several clustering algorithms; McLoughlin et al. [16] presented a methodology for electricity load profile characterization through clusters for Ireland using 3941 customers.

The Southwest European region have not yet be analyzed in terms of electricity consumer profiles, which has been seen as a bottleneck for the identification of opportunities for energy reduction and further energy efficiency achievements. Usually, there are statistics and knowledge regarding the national level, although, for effective opportunities of policy instruments or services toward energy efficiency and reduction there is the need for data and knowledge at a more local level.

An analysis of the data available for Évora indicates that, 82% of the residential buildings are associated with single-family houses (mainly terraced houses) and only 8% with apartments [17]. This presents a relevant difference from the EU average countries with 64% of residential buildings being single-family houses and the remaining 36% being apartments [18].

A substantial share of the buildings stock in Évora, as in other European cities, is older than 50 years. More than 20% of the residential buildings have been constructed before the 1940s when energy-building regulations were very limited. A large increase in construction in 1946–1990 is also evident, with the buildings constructed in this period representing around 36% of the current city stock [17].

This paper aims to identify, understand and explain representative yearly electricity consumption profiles of households, for the case study of Évora municipality. We applied a clustering approach to electricity consumption data, gathered from smart meters, and linked it with a dedicated survey for the same households to identify and characterize target groups of consumers.

We argue that the proposed methodology and the achieved results are useful to derive insights to support utilities, retailers and ESCO’s for marketing segmentation and innovative policies for effective energy reduction, as it is the case of tariff design, demand side management strategies, energy efficiency improvements, among others.

The paper is organized in four sections. Section 2 describes the methods and discloses the data used. Section 3 presents selected results regarding electricity profiles by consumption clusters and related explaining variables. Section 4 concludes.

2. Methods and data

This section describes the methodology used. Through the combination of the smart metering dataset provided by an electricity distribution company as in Wyatt [19] and Bartusch et al. [20]; and a door-to-door survey as in Kavousian et al. [21] and Gram-Hansen et al. [22]; we have made an in-depth analysis through segmentation of consumers based on clustering electricity consumption, aiming to identify distinct yearly electricity consumption profiles and to characterize their determinants. Fig. 1 explains the work was developed and how the different steps were addressed. Each step will be described next.

2.1. Door-to-door household surveys

The door-to-door survey consisted in 110 questions and encompassed information on location, socio economic data (e.g. average monthly income, family size), equipment’s ownership and use (e.g. number of hours of use in a day) and physical characteristics of the dwellings (e.g. bearing structure).

The fieldwork of the survey was carried out through the municipality of Évora during July and August 2014, including urban and rural areas. The identification and selection of the locations to make interviews was supported on the existing internal districts of the municipality i.e. parishes, which are the lowest spatial unit with available statistical data. Évora municipality has twelve parishes, three in the urban area comprising around 80% of the population and nine in the rural areas. Therefore, for our purpose, four districts were identified: we combined all the rural parishes in one sector and the three urban parishes were individually kept as districts.

Due to budget limitations we set a maximum of 400 interviews to be done. Because of onsite difficulties, time and transport logistics and interviewers availability constraints, we were able to collect 389 valid surveys, representing 97% of the total expected surveys, being 37% of the surveys answers collected in rural areas, and the remaining in the urban area. This way we were able to capture different households characteristics and consumer types.

2.2. Smart meters dataset

This study also relies on data from a massive smart metering system conducted for the first time in Portugal in the municipality of Évora, within the InoCity project (EDP Distribuição S.A. [23]). It contains measurements of electricity consumption registries gathered from 31,000 household every 15 min since April 2010. The installed equipment’s in Évora are concentrators from EFACEC and Jans meters with PLC communication (FSK modulation) in the GENELEC-A frequency band (35–91 kHz). Data collection of load diagrams from the meters to the distribution transformer controller is done on a daily basis starting at 00:00 and for every 6 h. The InoCity project is being carried out by the main Portuguese electricity distribution company, hence the smart meters component is integrated within a full smart city philosophy, which comprises better network management, remote and centralized control stations; electric mobility and distributed generation systems (EDP Distribuição S.A. [23]).

Residential electricity consumption has strong temporal variation, which is not captured with low-resolution consumption data such as monthly bills, thus high–resolution electricity consumption data from smart meters is vital. Therefore, making use of this data, a sample was collected; the household surveys were linked to the smart meters database through the household meter number, while preserving the confidentiality of the house owners. As
mentioned by Wijaya et al. [24], consumer segmentation alone is not enough. Therefore combining these two sources of information allows an extensive and coherent household data analysis to develop effective and efficient policies better targeting different consumers.

Of the total number of collected surveys (i.e. 389), we were able to identify and link 64% of them with the respective smart meters (275). The reasons for the gap are twofold: (1) the interviewers were not able to identify the number of the meter so we were not able to link the survey to the 31,000 smart meters database (32%) or (2) no smart meter is installed in that household (4%).

Data availability is dependent on the smart meters rollout in the municipality, since not all the meters were installed in the beginning of the project (i.e. 2010). For our objective, while avoiding too much granularity (by using 15 min data), daily electricity consumption data was retrieved for the years 2011 to 2014. By excluding 2010 data we were able to collect a more complete database. Despite the information acquired from the surveys referred only to 2014 with possibilities of changes in household socio-economic details (e.g. tenure, number of people, income); we assumed that the characteristics mostly apply for the electricity profiles of 2011–2014.

Information on the type of tariff (dual and single) and contracted power (kVA) was also obtained for improved knowledge on the sampled households. The type of tariff is related to the costs of electricity, depending on the hours of consumption (day, night and weekends), while the contracted power (e.g. 1.15 kVA, 3.45 kVA, 6.9 kVA) constrains the number of electrical appliances that could be used simultaneously.

A data trimming of the electricity dataset was made. The electricity registries from the Distribution System Operator (DSO) have the information per meter of accumulated electricity consumption. To have the daily electricity consumption we had to subtract the registry of the previous day from the registry of the day. Following Torsten et al. [52] meters with annual readings with less than 80% of available electricity readings were discarded. Thus, we screened the full data set of 275 meters to identify major data faults (i.e. missing of sporadically daily registries; missing several days/months in sequence). 10 meters were excluded in this step.

For further analysis, the daily electricity consumption data were averaged for the four years (2011–2014), preserving the intra-annual variability for each household. This approach will allow us to identify important and distinctive profiles of consumption and make typifications of consumers’ characteristics (e.g. dwelling characteristics and occupants’ profiles) for each electricity consumption profile. The 265 meters remained with few days with missing data (less than 1%), for which we imputed values based on the average values of the neighboring days.

2.3. Data analysis methods

An exploratory data analysis of the final sample of 265 households’ daily electricity consumption data sets from smart meters was made, as well as a clustering analysis. The cluster analysis was carried over the daily means (per household), i.e., averaged over 2011–2014 for each day. After the previous explained electricity data trimming, we applied a hierarchical clustering using the Ward’s Method [25] with a measured interval through the squared Euclidean distance, allowing an analysis of variance approach to evaluate the distances between clusters. This method is regarded as very efficient, however, it tends to create clusters of small size [26]. Therefore, through an iterative process, we evaluated the clustering results for a number of clusters ranging from 3 to 12. We concluded that still maintaining robustness and statistical significance of the clustering, only increasing the number of clusters allows to capture distinct yearly consumption patterns that would be interesting to unravel and compare, in order to create types of consumer for which different policy and energy reduction measures could be targeted. The 10 clusters option with similar means and standard deviations were selected for further profiles analysis. It met a well balance option to illustrate differences of the annual profiles with significant number of meters/surveys.

After allocating each survey to the correspondent cluster, a screening of the answers of the surveys was made in order to recognize the most relevant parameters (e.g. dwelling characteristics, occupants profiles, electrical appliances ownership and use) that further explain the electricity consumption patterns.
and major similarities/distinctions within clusters allowing an increased knowledge on the clusters segmentation.

From the information collected in the households survey, we retain the following variables to characterize the households: (i) location (Urban and Rural) (as in e.g. [27,28]), (ii) dwelling type (as in [29,30]), (iii) dwelling age [31,32], (iv) dwelling total floor area (e.g. [21,31]), (v) type of glazing and windows framing, (vi) bearing structure and (vii) type of external walls. The following socio economic variables, which might influence electricity consumption, were selected: (i) the number of occupants (according to [6,22]), (ii) education of the household responsible person (e.g. [22]), (iii) household income [34,35] and (iv) employment status (e.g. [36]). For factors associated with electrical appliances and heating and cooling equipment we selected the following variables: (i) ownership of heating and cooling technologies (as in [29,37]), (ii) ownership of white electrical appliances (as in [30,38]), (iii) type of tariff and (iv) contracted power.

Statistical analysis performed over very high temporal resolution data allows the characterization of the electricity consumption profiles. Significant differences and similarities within cluster groups were assessed, which can be useful to support market segmentation and tariff design for DSOs and to improved knowledge on groups of consumers for ESCO’s and for electricity retailers to feed specific energy and pricing reduction recommendations.

### 3. Results and discussion

In this section, we aim to explore the results from the clustering analysis portraying the different yearly consumption profiles, and its most relevant determinants gathered from survey data to explain household electricity consumption clustering. Fig. 2 presents the daily electricity consumption for the sampled meters (265 households) averaged for the four years with consistent available data (2011, 2012, 2013, and 2014) and the corresponding daily minimum temperature. A higher daily average consumption in the winter months of December and January and in the summer months of July is apparent, presenting stronger inverse correlation with the minimum daily temperatures ($r = -0.82$) (Fig. 2), maximum daily temperatures ($r = -0.77$) and with the daily average temperature ($r = -0.80$). These correlations show a potential direct link between electricity consumption and the use of cooling and heating systems.

#### 3.1. Electricity data clustering from smart meters

The clustering method applied split the 265 smart meters dataset into 10 clusters, showing a distinct distribution of meters (with at least 25 meters per cluster) within clusters with mean daily electricity consumptions below 12 kWh (cluster 1 to 5), totaling 217 meters (more than 82%) and the other five clusters. The remaining 47 meters are included in clusters 6–10 fitting the high levels of consumption and/or variability with daily median consumption of almost 28 kWh (i.e. cluster 8) (Fig. 3).

The box-and-whisker plot unveils the descriptive statistics of the clusters (Ci) regarding their dispersion and skewness, and the existing outliers. The distribution of electricity consumption data from C1 to C5 is similar, with C1 presenting the lowest statistics (median 3.99 kWh and standard deviation of 2.10 kWh) and C2 the highest variance (standard deviation of 4.26 kWh). The short box plots within these clusters (and also C9 and C10 at a certain extent) suggests that, generally, the consumption data have similar profiles. Differences within these clusters can be further evaluated in Table 1. Clusters C6–C8 present tall box plots depicting significant variances (standard deviations ranging from 6 to 11 kWh) within clusters already unveiling possible differences among the seasons of the year. Cluster C7 shows the highest data variability (standard deviation of 10.87 kWh) and highest consumption. With the exception of clusters C6 and C7, all the other clusters have a consistent distribution of data within the second and third-quartile.

Table 1 advances on the annual electricity consumption profiles of each one of the ten clusters allowing to identify important distinctions of inter annual consumption patterns. Three distinctive profiles can be concluded, named (a) U profile; (b) W profile, (c) Flat profile. The rational behind this identification is a visual analysis further supported on the evaluation of the differences between the levels of electricity consumption in winter and in summer. This segmentation has the objective of acknowledging the different yearly electricity consumption patterns of household consumers and to further evaluate them.

Under the U Profile we include six clusters (77% of the sampled households): C1, C2, C3, C4, C7 and C8 (in orange). This profile could be further disaggregated, in the Sharp U profile (C1, C2, C3 and C7) and the Soft U profile (C4, C6 and C8) branches. Sharp U profile is characterized by differences of electricity consumption on winter and the remaining seasons of the year (around 50% lower in summer) possible explained by the lower ownership and use of cooling electrical when compared to electricity-supported systems. Sharp U profiles present a higher difference on winter and the remaining seasons (around 70% lower in summer in some clusters) portraying the inexistence or low use of cooling equipment in the summer compared to a strong use of electricity-based technologies for space heating in the colder months of winter (December, January and February) which is corroborated by the findings in Tables 2–5.

Under the W Profile (in blue), we include clusters C6, C9 and C10 (12% of our sample), which present clear distinctions of electricity consumption between summer and winter and the inter-seasons period. These clusters with high values of daily consumption present a strong hump-shaped consumption in summer,
Table 1

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph 1" /></td>
<td><img src="image2.png" alt="Graph 2" /></td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Cluster 4</td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph 3" /></td>
<td><img src="image4.png" alt="Graph 4" /></td>
</tr>
<tr>
<td>Cluster 5</td>
<td>Cluster 6</td>
</tr>
<tr>
<td><img src="image5.png" alt="Graph 5" /></td>
<td><img src="image6.png" alt="Graph 6" /></td>
</tr>
<tr>
<td>Cluster 7</td>
<td>Cluster 8</td>
</tr>
<tr>
<td><img src="image7.png" alt="Graph 7" /></td>
<td><img src="image8.png" alt="Graph 8" /></td>
</tr>
<tr>
<td>Cluster 9</td>
<td>Cluster 10</td>
</tr>
<tr>
<td><img src="image9.png" alt="Graph 9" /></td>
<td><img src="image10.png" alt="Graph 10" /></td>
</tr>
</tbody>
</table>
Table 2
Summary of selected variables characterizing the dwellings of Clusters 1, 5, 7 and 9.

<table>
<thead>
<tr>
<th>Shape of annual electricity profile</th>
<th>Cluster</th>
<th>Characteristics of dwellings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Location (%) Type (%) Period of construction (%)</td>
</tr>
<tr>
<td>U shape (soft)</td>
<td>C1</td>
<td>61 39 30 52 18 16 20 55 9 – 90</td>
</tr>
<tr>
<td>Flat</td>
<td>C5</td>
<td>34 66 52 34 14 3 22 38 34 3 117</td>
</tr>
<tr>
<td>U shape (sharp)</td>
<td>C7</td>
<td>78 22 33 33 33 22 5 – 25 50 25 162</td>
</tr>
<tr>
<td>W shape</td>
<td>C9</td>
<td>38 63 24 38 38 – – – – – – –</td>
</tr>
</tbody>
</table>

Table 3
Summary of selected variables characterizing the household occupants of Clusters 1, 5, 7 and 9.

<table>
<thead>
<tr>
<th>Shape of annual electricity profile</th>
<th>Cluster</th>
<th>Characteristics of household occupants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average number of persons per household Age of household members (%) Gender of household members (%) Education of the head of the family (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;5 years 5–17 18–49 50–64 ≥65 years old Male Female &lt;9th grade 9th–12th grade Graduation, M.Sc or Ph.D.</td>
</tr>
<tr>
<td>U shape (soft)</td>
<td>C1</td>
<td>2 – 7 32 18 43 44 56 53 35 12</td>
</tr>
<tr>
<td>Flat</td>
<td>C5</td>
<td>2.86 – 16 48 14 21 52 48 44 44 12</td>
</tr>
<tr>
<td>U shape (sharp)</td>
<td>C7</td>
<td>2.44 – 9 32 36 23 45 55 13 24 63</td>
</tr>
<tr>
<td>W shape</td>
<td>C9</td>
<td>4 – 3 19 59 19 – 56 44 – 50 50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shape of annual electricity profile</th>
<th>Cluster</th>
<th>Characteristics of household occupants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Monthly average income of the household (%) Employment status (%) Household occupation contract (%) Relation of household members (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≤750€ 751–1500€ 1501–2500€ ≥2501€ Working Retired Student Other Owner Rented Tenant for free Family Room mates Other couple</td>
</tr>
<tr>
<td>U shape (soft)</td>
<td>C1</td>
<td>60 29 11 – 32 47 15 6 60 38 2 89 9 2</td>
</tr>
<tr>
<td>Flat</td>
<td>C5</td>
<td>26 53 21 – 51 23 20 5 86 14 – 100 – –</td>
</tr>
<tr>
<td>U shape (sharp)</td>
<td>C7</td>
<td>20 40 20 20 43 23 17 17 89 11 – 100 – –</td>
</tr>
<tr>
<td>W shape</td>
<td>C9</td>
<td>– 17 33 50 50 9 38 3 100 – – 100 – –</td>
</tr>
</tbody>
</table>
Table 4

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Heating technology (%)</th>
<th>DHW technologies (%)</th>
<th>Cooking technology (%)</th>
<th>Electric store</th>
<th>Gas stove</th>
<th>Non electric (gas, oil, wood)</th>
<th>Non electric (gas, oil, wood, heat pumps)</th>
<th>Electric (resistance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>86</td>
<td>95</td>
<td>79</td>
<td>43</td>
<td>57</td>
<td>27</td>
<td>46</td>
<td>12</td>
</tr>
<tr>
<td>C2</td>
<td>79</td>
<td>43</td>
<td>73</td>
<td>3</td>
<td>6</td>
<td>24</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>C3</td>
<td>89</td>
<td>73</td>
<td>67</td>
<td>6</td>
<td>17</td>
<td>64</td>
<td>57</td>
<td>27</td>
</tr>
<tr>
<td>C4</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>C5</td>
<td>95</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>C6</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>C7</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>C8</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>C9</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>C10</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Profile</th>
<th>White appliances (%)</th>
<th>Other electrical equipment (%)</th>
<th>Cluster</th>
<th>Appliances ownership</th>
<th>White electricity profile</th>
<th>Other electricity profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>16</td>
<td>52</td>
<td>U Shape (soft)</td>
<td>100</td>
<td>61</td>
<td>90</td>
</tr>
<tr>
<td>C2</td>
<td>16</td>
<td>52</td>
<td>U Shape (soft)</td>
<td>61</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>C3</td>
<td>16</td>
<td>52</td>
<td>U Shape (sharp)</td>
<td>100</td>
<td>61</td>
<td>90</td>
</tr>
<tr>
<td>C4</td>
<td>16</td>
<td>52</td>
<td>U Shape (sharp)</td>
<td>61</td>
<td>90</td>
<td>80</td>
</tr>
</tbody>
</table>

Clusters 1, 5, 7 and 9 are most prominent in C9 and C10. These profiles suggest that the respective households might have high ownership rates and use of HVAC systems for cooling and low use of electrical equipment for heating in the winter (C6) or both high use of electrical systems for cooling and heating (C9 and C10). Nevertheless, we acknowledge that C6 is not a very distinct W profile, looking more as a transition profile between a Soft W and Soft U.

Cluster C5 (in green) is in the lower levels of consumption of all the sampled data, and we considered it as having an annual flat consumption profile (11% of the sampled households) since the consumption has small intra annual variations. The minimum of electricity consumption is 62% of the maximum consumption value. This type of profile could be explained by the inexistence of electrical equipment for cooling or heating being the electricity consumption only linked to end uses like cooking, lighting, washing clothes and dishes and other electric equipment (e.g., televisions, computers) which usually do not have seasonal variations.

For an in-depth characterization of the electricity consumption of the households behind the four recognizable distinct profiles, we crossed the meters data of each cluster with the correspondent survey results. Considering the statistical behavior and yearly patterns presented previously, three clusters are selected as examples of each profile to present the results: (a) W profile – Cluster 1 and Cluster 7, (b) W profile – Cluster 9 and (c) Flat profile – Cluster 5. The selection was based on distinct consumption profile of clusters regarding the mean (low, medium, high), dispersion (low and high) and annual profile (similar along the year or different in winter and/or summer months).

van Raaij and Verhallen [5] recognized several factors that drive household electricity consumption behavior, such as energy-related attitudes, personality, socio-demographic factors, building characteristics, energy prices, feedback and general information about energy use. Kelly [8] identified for England the number of household occupants, floor area, household income, dwelling efficiency, and household heating patterns and living room temperature as the main drivers behind residential energy consumption. For Germany, Gruber and Scholman [7] showed that electricity consumption is strongly influenced by the number of existing equipment, household area and annual income. Bartiaux and Gram-Hansen [6] exposed for Belgium and Denmark that family size; dwelling area and number of equipment are strong determinants for electricity consumption.

Tables 2–5 disclose selected variables collected in the surveys, to be compared throughout the chosen clusters. By evaluating the survey results for the households in each cluster, it is possible to identify important similarities and differences regarding socio-economic determinants, dwelling characteristics and appliances use and ownership, which could further explain the different clusters’ aggregation and levels of consumption and profiles.

Cluster 1 is characterized by a predominance of terraced dwellings located in urban areas, in small houses (around 90 m²) built between 1946 and 1990 period. Following the period of construction, materials and techniques, the predominant bearing structure of the dwellings comprised in this cluster is masonry wall with or without plate associated with brickwork single layered in the external walls. The majority of the dwellings (83%) have single glazing and wooden window framing.

Regarding occupants characteristics, we can say that this clusters’ households are portrayed by the smallest families of all clusters (average of two persons per household), generally older than 65 years old with low levels of education (secondary level), retired and with households monthly average income below 750€. It is in this cluster that the level of owner occupied houses is the lowest, with a relative important share of rented houses (38%).

The electricity profile of this cluster (Soft U Shape), defined by a significant difference of consumption on winter months is...
backed up by the survey results with predominant ownership and use of electric heating equipment (88%). Only 46% of these cluster dwellings have cooling equipment. From which, near 80% own fan coils, that consume a lot less than HVAC systems. Still, it is in this cluster that the ownership and use of fans is the lowest.

In C1, the overall smallest ownership of white appliances, computer equipment and lamps from all the clusters combined with the dominant number of houses (78%) with low contracted power (under 3.45 kVA) also explain the lowest levels of daily electricity consumption in this cluster when compared to others. 71% of the houses in this cluster still have single tariffs not taking advantage of the lowest prices at night of dual tariffs.

Being the cluster with the higher number of dwellings from our sample (21%), C1 is, as seen, characterized by the lowest electricity consumption levels and annual consumption profile portraying the lack of fulfillment of thermal comfort levels inside households both in summer and winter, suggesting a case of fuel poverty. As described in a project from EPEE [39], by Moore [40] and Thomson and Snell [41], backing up our results, the combination of low incomes, low performance dwellings with defective insulation (windows, walls, roofs) and older households are enablers of fuel poverty. Also consistent with our findings, Wand [42], under EU fuel poverty network, pointed out that currently in Portugal, around 28% of the population is unable to keep their home adequately warm.

**Cluster 5:** Opposing with C1, C5 is characterized by a high share of households located on rural areas (66%), and with higher prevalence of more recent built houses of the semi-detached type. Furthermore, other characteristics describing the households within this cluster are: average size dwellings around 117 m², built after 1946 but with a high share built after 1991, also shown in the higher amount of concrete houses (33%). Increased share of insulation levels justified by the entrance of more restraining thermal regulations also represent important differences when compared to C1. The sampled houses of this cluster have a similar distribution of single and double-glazing but the majority of them has aluminium framing in the windows (64%).

Regarding occupation, C5 is established by higher number of occupants inside the households (2.86), contrasting with C1 concerning the age of occupants, household income, employment status and household occupation status: 64% of the occupants aged below 50 years, and 51% working full time reflected on medium levels of monthly income (i.e. 53% of houses earning between 751 and 1500€).

The construction characteristics combined with the very high ownership and use of non-electricity based equipment for cooking (103%), heating (57%) and DHW (97%), enable us to better understand the annual flat electricity consumption profile. When available, the space heating falls back predominantly on fireplaces, and the space cooling is majorly carried out with fans. Consistent with the increase in daily average consumption when compared to C1, C5 has 66% of the households with contracted power between 4.6 and 6.9 kVA, and a very high share of single tariff users. The electricity consumption profile of C5 portrays a standard comfort household.

**Cluster 7** has the highest share of urban dwellings. It presents an even occurrence of the three considered types of houses (terraced, detached and semi detached), thus not being an explanatory variable distinguishing the houses in this cluster compared to other clusters. Construction characteristics (e.g. period of construction, external wall and glazing) of dwellings are very similar to the ones illustrating C1.

The deepest differences on the amount and the seasonality of electricity consumption between C1 and C7 include the higher average household area (40% higher in C7), and the number of persons per household (2.44), suggesting more space heating needs in winter months. In this cluster, the bulk of the age of the household members is below 64 years old, 80% of the monthly income are above 750€ and directly related to the high levels of education of the head of the family (63% have at least a graduation). These socioeconomic characteristics are effective drivers of the C7 electricity consumption profile.

Regarding appliances ownership, C7 presents one of the highest levels of penetration of space heating equipment (89%), from which 73% have electric heaters or HVAC (the majority bought after the 2005 summer heat wave in Portugal). 67% of the houses in this cluster own equipment for cooling but the lion share being fan coils which once again explain the sharp difference of seasonal consumption.

Besides all the previous characteristics, the very high daily average consumption is also justified by the high penetration of white appliances and other electrical equipment. The penetration of refrigerators (133%), freezers (100%), microwaves (111%), dish washing machines (89%), electric stoves (67%) and number of lamps per household are higher than in the previous assessed clusters, showing a clear evidence of consumers with a higher levels of disposable income.

Regarding contracted power, all the dwellings in C7 households have at least 4.6 kVA (72%), with once again a dominance of single tariffs contracts (56%). We may state that C7 portray what we may name as ‘fat energy’ households with opportunities for potential reduction of electricity consumption, either through energy efficiency options and/or more rational energy behaviors.

**Cluster 9:** The electricity consumption in this cluster households follow a W profile, recognizable by the high levels of consumption in winter and summer months when compared to the inter seasons months. The dwellings are predominantly located in rural areas (63%), with a strong predominance of houses constructed after 2006 with high average floor areas (162 m²). In the research carried out by Zhao et al. [53], there is a clear distinction between the patterns of energy use in urban and rural households, due to higher energy services demand in urban households. In our work this is not recognizable. Despite the important share of houses in rural areas, they are still close (less than 30 km) to the urban city environment, therefore with similar urban patterns consumption.

Dwelling characteristics, as bearing structure, type of wall and windows (87% with double glazing and 63% with aluminium framing) arise to distinctively characterize this cluster. Similarities within other important explaining determinants of electricity consumption such as household occupants include: the average number of four persons per household; 59% aged between 18 and 49 years old; 100% of the adult members have at least the secondary level of education; 50% of the members either have full time jobs (50%) or are students (38%); 50% of the households have an income level above 2500€ per month, the highest share of all clusters.

The high income relates with the ownership of electrical equipment both impacting the quantity and quality of the appliances (e.g. [49]). A large body of literature has also concluded that energy consumption increases with income [32,44,45]. However, the opposite have also been identified by other studies (e.g. [49]). All these socio economic features can typify middle to high-class family, with two working adults and two children, and explaining the high consumption levels throughout the year but especially in winter and summer seasons.

When evaluating the survey results for the houses in C9, we can conclude that the identical levels of consumption in winter and summer are validated by the dominance of air conditioning systems for cooling and a mix of electric (36%) and non-electric (64%) equipment for heating. The lower ownership and use levels of electrical heating equipment, as oil heaters and HVAC, closes down the gap between both seasons (i.e. winter and summer) consumption. Also supporting the high daily electricity consumption is one of the higher ownership levels of white appliances, computers
and lamps of all the clusters. Desktops, laptops, refrigerators, freezers, microwaves, cloth washing machines and televisions have all ownership levels higher than 100%. Cloth drying machines have in this cluster the highest penetration rate of all the clusters. This high daily consumption cluster has the double of the average lamps per household (i.e. 20) of the lowest consumption cluster – C1.

As expected by the electricity consumption profile, 40% of the households have a contracted power higher than 6.90 kVA with 62% taking advantage of dual tariff pricing. C9 also can be considered ‘fat energy’ households with a different profile, with opportunities for effective reduction of electricity consumption.

The relationship between area, persons per household and consumption portrayed in this cluster is also referred by Larsen et al. [46], Kaza [44] and Gram-Hanssen [47] that present the number and the use of appliances correlated to the number of people living in the house; but for Kaza [44], the space cooling and heating use is likely to be same irrespective of number of people. However, it is more energy efficient to live more people together, as families with more members consume less electricity per capita [31,46].

Our analysis suggest to conclude for three major groups of determinants that influence the residential electricity consumption segmentation: (i) physical characteristics of a dwelling, especially year of construction and total floor area; (ii) electrical heating/cooling equipment and fireplaces ownership and use; and (3) occupants profiles (mainly number of occupants and monthly income).

### 3.2. Insights for policy and stakeholders

The characterization of the dwellings, in terms of construction type, socio economic factors and equipment, beneath the consumption of the clusters highlight and explain the wide range of electricity consumption profiles, within consumers of the same region. This illustrates the relevance of consumer segmentation for policies and measures design and implementation, tailored to energy reduction.

Following other studies outcomes (e.g. [50]), our results unfold that higher average household area also reveals higher energy consumption. However, when comparing the clusters on household occupants we can state that there is (but not on all clusters) a non-linear relationship between household electricity consumption and the number of occupants, as also suggested by Brounen et al. [32], Kavousian et al. [21] and Hayn et al. [11].

According to these four clusters evaluation, we can say that tariff while being similar to several clusters is not a paramount explanatory variable of the segmentation. Furthermore, we might also conclude that gender, type of household occupation contract (contrary to the findings of Ndaiye and Gabriel [48]) and relation of household members are also variables that not significantly distinct the consumption profiles. Other determinants collected in the surveys which do not make a distinction between clusters therefore not being group specific to tackle individually measures are: penetration of electric equipment for DHW, high substitution of incandescent lamps for compact fluorescent lamps and widespread ownership of refrigerators and cloth washing machines near or above 100%.

Our results on typification of electricity consumption profiles and description of the characteristics of the households beneath them, unfolds important results for several stakeholders in the electricity services supply chain. The DSO would benefit from better handling the peak demand, making use of seasonal tariffs and balancing changes in contracted power. Besides, in constrained budget availability and knowing the amount of consumers in each cluster, it is important to target measures to the most relevant group of consumers. For example, the U and W shape consumers could benefit from changes of contracted power in the seasonal variations of electricity consumption, reducing their annual expenses with electricity.

Electricity retailers and ESCO’s can also benefit from the detailed awareness of consumers’ profiles at local level, in order to make tailor-made measures targeted to groups of consumers with similar needs, equipment and socio economic profiles. In fact, there is a significant difference between groups of consumers within the same municipality, with some consumers struggling to achieve minimum comfort levels in winter and summer months, while others consuming three times the average along the year, which require a well-balanced interplay of policies and measures. For some consumers, focusing on improvement of dwelling characteristics through insulation measures, improvement of roofs, walls and window materials is decisive. For others, the ultimate goal of energy reduction coupling energy efficiency measures (i.e. equipment’s substitutions) and behavioral changes might be the focus.

### 4. Conclusions

This paper examines how the combination of smart meter data and door-to-door survey information can deliver important results and meaningful knowledge regarding households’ electricity consumption profiles. Exploratory analysis and hierarchical clustering were applied. The annual consumption profiles were extensively characterized and explained through socio economic characterization of the household members, dwellings characteristics and equipment ownership.

The analysis of the electricity consumption profiles of 10 clusters, showed four distinct types of annual consumption patterns in the municipality of Évora: U shape (soft and sharp), W shape and Flat, which might bring different insights for public policies and stakeholders decisions. U shape profile is the most common one, covering 77% of the sampled houses, and is characterized by a significant difference of electricity consumption between winter and the rest of the year unraveling the low ownership levels of air conditioning system for space cooling in Évora (confirming also the pattern at national level).

Based on three major groups of electricity consumption determinants: dwelling’s physical characteristics, especially year of construction and total floor area; electrical heating/cooling equipment and fireplaces ownership and use; and occupants profiles,
mainly number of occupants and monthly income), we typify and distinguish three main groups of consumers: fuel poverty, standard comfort, and “fat energy” households. Therefore, future policies and measures, as well as energy services companies, should take into account these differences to better serve simultaneously energy efficiency and thermal comfort levels.

The fieldwork was conducted in a southwestern European city, however the methodology can be applied to any region equipped with a smart metering network. This paper also discloses the importance of the future widespread use of smart meters, to benefit both the consumers interest and the stakeholders of the electricity services supply chain. In fact, despite acknowledging that such a consistent dataset of information with an extensive characterization of consumers is still rare and is unlikely to be collected by electricity retailers or DSOs, we consider that the outcomes of our analysis could also be used as a starting point for utilities looking at peak shaving and electricity demand shifting inside households derived from market segmentation.

Acknowledgements

The work supporting this paper was funded by the EU project INSMArt, Integrative Smart City Planning, under the Seventh Framework Programme grant agreement No.: 314164 and by the Portuguese Science and Technology Foundation (FCT) through the Scholarship SRH/BD/70177/2011. The authors would like to thank EDP Distribuição, S.A., especially Vera Nunes and Miguel Andrade for their contribution regarding smart meters data; Nuno Bilo and António Valentin for the help on the household surveys and Shiming Luo and Joana Monjardino for their assistance with the smart meters data set handling. The authors would also like to thank the interviewers and citizens that contributed for to the completion of the household surveys.

References

[27] C.R. Wand, % of households unable to keep their home adequately warm, 2013, Available at: http://fuelpoverty.eu.


