Understanding electricity consumption patterns in households through data fusion of smart meters and door to door surveys

João Pedro Gouveia
NOVA School of Science and Technology NOVA University of Lisbon

Julia Seixas
Universidade NOVA de Lisboa

Understanding electricity consumption patterns in households through data fusion of smart meters and door-to-door surveys

João Pedro Gouveia
CENSE, Faculdade Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
FCT/UNL
Campus da Caparica
2829-516 Caparica, Portugal
jplg@fct.unl.pt

Júlia Seixas
CENSE, Faculdade Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
FCT/UNL
Campus da Caparica
2829-516 Caparica, Portugal
mjs@fct.unl.pt

Shimming Luo
CENSE, Faculdade Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
FCT/UNL
Campus da Caparica
2829-516 Caparica, Portugal
irisluo92@gmail.com

Nuno Bilo
Municipality of Évora
7005-366 Évora, Portugal
nuno.choraobilo@cm-evora.pt

António Valentim
Municipality of Évora
7005-366 Évora, Portugal
avalentim@cm-evora.pt

Abstract

This paper contributes to a comprehensive understanding on how electricity consumption patterns are explained in a Southwest European city. Household electricity consumption drivers and profiles of different type of consumers are analysed. They are built upon data gathered from a sample of 230 households with daily electricity consumption from smart meters of InovCity project, the first of its kind in Portugal consisting of a massive smart metering system (31 000 smart meters) with door-to-door 110–question surveys for the same houses in Évora. The survey encompassed information on socio economic data, equipment's ownership and use and physical characteristics of the dwellings. Ten clusters were derived using Ward's method hierarchical clustering to identify similar types of consumers based on their means and standard deviations, and three of them are deeply analysed and compared. Based on the surveys, a socioeconomic characterization of each cluster was made in order to capture the main similarities and differences within each cluster and compared to the others. We conclude that three major groups of determinants influence residential electricity consumption segmentation: physical characteristics of a dwelling 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). Urbanisation levels, bearing structure, type of tariff and contracted power are not variables that distinct the clusters grouping. This consumer profiling allows deriving insights to support utilities for marketing segmentation and policies for effective energy reduction (e.g. tariff design, demand side management strategies, peak shaving). This work is being developed under the EU project InSMART, that

involves four European cities targeting innovative methods to integrative city planning, including buildings, transport, and utilities networks.

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 (Lomas, 2010), contrary to policy goals aiming a transition towards 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, greenhouse gases emissions reduction and increased share of renewable energy sources, as established in the new European Union goals by 2030 (EC, 2014) requires more decisive action.

Energy consumption in buildings deserves special attention since they represent a significant share of energy consumption in OECD countries as 20-30% in European Union (EU) (Eurostat, 2011). In Portugal, residential buildings consume approximately one third of total electricity, with a growth of 70% from 1995 to 2012 (DGEG, 2014). This consumption is a complex issue that can be explained by a combination of physical, technological, demographic, climatic and behavioural 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. Van Raaij and Verhallen (1983) in their research in the 1980s, recognize several factors that drive household electricity consumption behaviour, such as energy-related attitudes, personality, socio-demographic factors, building characteristics, energy prices, feedback and general information about energy use. Kelly (2011) 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 Scholmann (2006) showed that electricity consumption is strongly influenced by the number of existing equipment, household area and annual income. Bartiaux and Gram-Hanssen (2005) exposed for Belgium and Denmark that family size; household area and number of equipment are strong determinants for electricity consumption. More recently, Jones et al. (2015) 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. "*As smart meters are replacing traditional electricity meters in large parts of Europe, there is now a unique opportunity to realize comprehensive consumer feedback systems that consist of much more than mere remote metering applications*" (Weiss *et al.*, 2013). With growing deployment of smart meters and real-time home energy-monitoring services, data that allows studying the determinants of energy consumption inside households and electricity consumers' profiles are becoming available.

Crossing 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 behaviours (McDonald and Dunbar, 2012). 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 behaviour and equipment with the ultimate goal of an effective energy consumption reduction and load curve consumption peaks minimization.

In other countries, there are already studies on the residential electricity consumption profiles using smart metering data. Seo and Hong (2014) with a 30 households sample for Daegu in South Korea characterized power consumption patterns and presented summer load profiles. Rhodes et al. (2014) 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. (2014) demonstrated profiles of electricity consumption for 60 low energy-housing houses in South Australia. McLoughlin et al. (2015) presented a methodology for electricity load profile characterization through clusters for Ireland using 3941 customers.

This paper presents results of a study aimed to improve the understanding of how electricity consumption patterns are explained in a Southwest European city. We combine smart meters electricity data for the case study of Évora municipality with a dedicated survey for the same houses in order to identify target groups of consumers through a clustering approach. This will be useful to derive insights to support utilities for marketing segmentation and policies for effective energy reduction (e.g. tariff design, demand side management strategies) (Armel *et al.*, 2013).

The work presented here is being developed as part of the EU project InSMART, that involves four European urban areas (Évora, Cesena, Trikala and Nottingham) targeting innovative methods to integrative city planning, including buildings, public lighting, transport, waste, water and wastewater networks (Gouveia *et al*., 2014).

The paper is organized in 4 sections. Section 2 summarizes 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, presents the limitations of the study and further work.

Methods and Data

Through the combination of a smart metering dataset provided by an electricity distribution company as in Wyatt (2013) and Bartusch et al. (2012) and a door to door survey as in Kavousian et al. (2013) and Gram-Hanssen et al. (2004) we have made an exploratory data analysis through segmentation of consumers based on clustering electricity consumption profiles identifying similar electricity consumption determinants per electricity consumption group.

Door-to-Door Surveys

The door-to-door survey consisted in 110 questions and encompassed information on location, socio economic data, equipment's ownership and use and physical characteristics of the dwellings. The fieldwork of the survey in the streets of Évora was carried out between July and August 2014. Due to 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 (400 were initially defined).

Évora municipality has twelve parishes, three in the urban area comprising around 80% of the population and nine in the rural areas. The surveys were made extensively along the entire municipality in order to collect information of a representative set of households, being 37% of the surveys answers were collected in rural areas, and the remaining in the urban area.

Smart Meters Dataset

As mentioned, our 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 InovCity project (EDP Distribuição S.A., 2014). It contains measurements of electricity consumption gathered from 31 000 household every 15 min since April 2010. This 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.

Since residential electricity consumption has strong temporal variation, which is not captured with low-resolution consumption data such as monthly bills, these high-resolution electricity consumption data is vital. Therefore, making use of this data, the surveys were linked to the smart meters database through the household meter number, although preserving the confidentiality of the house owners. Combining these two sets of information allows an extensive and coherent household data analysis.

Of the total number of surveys collected (i.e. 389) we were able to identify and link 64% of them with the smart meter database (275). The reasons 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. Thus, to have a more complete database, electricity data consumption was retrieved from 2011 to 2013. Data for the full year of 2014 was not yet available at the time of the current data analysis therefore it was not used. Despite the information acquired from the surveys refers to 2014; we assume that the characteristics mostly apply for the electricity profiles of 2011-2013.

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. 1.15kVA, 3.45kVA, 6.9kVA) constrains the number of electrical appliances used simultaneously.

According to Torsten *et al*. (2013) in a work done for Germany, data selection for analysis of households' consumption profiles requires a minimum of 80% of available electricity readings. Thus, meters with annual readings below this threshold were discarded, and the 275 meters were therefore reduced to 250.

For further analysis, the daily electricity consumption data were averaged for the three years, preserving the intra-annual variability for each household. This approach will allow us to assess the relationship between electricity consumption and structural explanatory variables, such as dwelling characteristics and occupants' profiles. This sample size was still reduced to 230 households, since, where at least 5% of the data over the 3-year period average was missing from a particular meter, that meter was excluded from the study (i.e. 20 meters).

An exploratory data analysis was made for the final sample of 230 households focused on electricity data clustering from smart meters data. The cluster analysis is based on daily means (per household), i.e., averaged over 2011-2013 for a given day. After the previous explained data trimming, we applied a hierarchical clustering through Ward's Method (Ward, 1963) 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 (Statsoft, 2015). Therefore, through an iterative process, we concluded for 10 clusters with similar means and standard deviations to allow a better evaluation regarding the linkage to the surveys parameters.

A screening of the surveys allocated to each cluster was made in order to identify the parameters (e.g. dwelling characteristics, occupants profiles, electrical appliances ownership and use) that explain the electricity consumption patterns and similarities allowing an increased knowledge on the clusters segmentation.

From all the factors collected in the households survey, we relate location (Urban and Rural) (e.g. Halicioglu, 2007; Raty and Carlsson-Kanyama, 2010), dwelling type (as in Bedir et al., 2013; McLoughlin et al., 2012), dwelling age (as in Wiesmann et al., 2011; Brounen et al., 2012), dwelling total floor area (e.g. Baker and Rylatt, 2008; Kavousian et al., 2013), type of glazing and windows framing, bearing structure and type of external walls as examples to describe characteristics of the dwellings. The socio economic factors, which might influence electricity consumption, that were selected are the number of occupants (as in Bartiaux and Gram-Hanssen, 2005; Brounen et al., 2012) education of the household responsible person (e.g. Gram-Hanssen, 2004), household income (Lam, 1998; Santamouris et al., 2007) and employment status (e.g. Cramer et al., 1985; Yohanis et al., 2008). For factors associated with electrical appliances and heating and cooling equipment we selected the following parameters: ownership of heating and cooling technologies (as in Bedir et al. 2013; Tso and Yau, 2007), white ownership of electrical appliances (as in Leahy and Lyon, 2010; McLoughlin et al., 2012) type of tariff and contracted power.

Statistical analysis performed over very high temporal resolution data allows the characterization of the electricity consumption profiles. This permits the identification of significant differences and similarities within cluster groups that could be useful for market segmentation and tariff design for utilities and to improved knowledge on groups of consumers to feed specific energy reduction recommendations.

Results and Discussion

In this section, we present results from the clustering analysis, which are linked with the most relevant determinants to explain household electricity consumption clustering. Figure 1 presents total daily average electricity consumption for all the sampled meters for the three years with consistent available data (2011, 2012, 2013). It reveals a higher daily average consumption in the winter months of December and January and in the summer month of July. A first assertion can be made regarding its relation with temperature, although we will not take further, in this paper, this variable to explain consumption data. The electricity consumption data set (3 years) presents a strong inverse correlation with the daily temperatures (-0.72), showing a direct link between electricity consumption and cooling and heating systems use. Evaluating each year individually, 2011 presents the highest inverse correlation (-0.80) and 2013 the lowest (-0.63). This might be explained by the financial constraints in Portuguese households restraining the electricity use. Summary statistics for the daily electricity consumption of the households in our sample are described in Table 1.

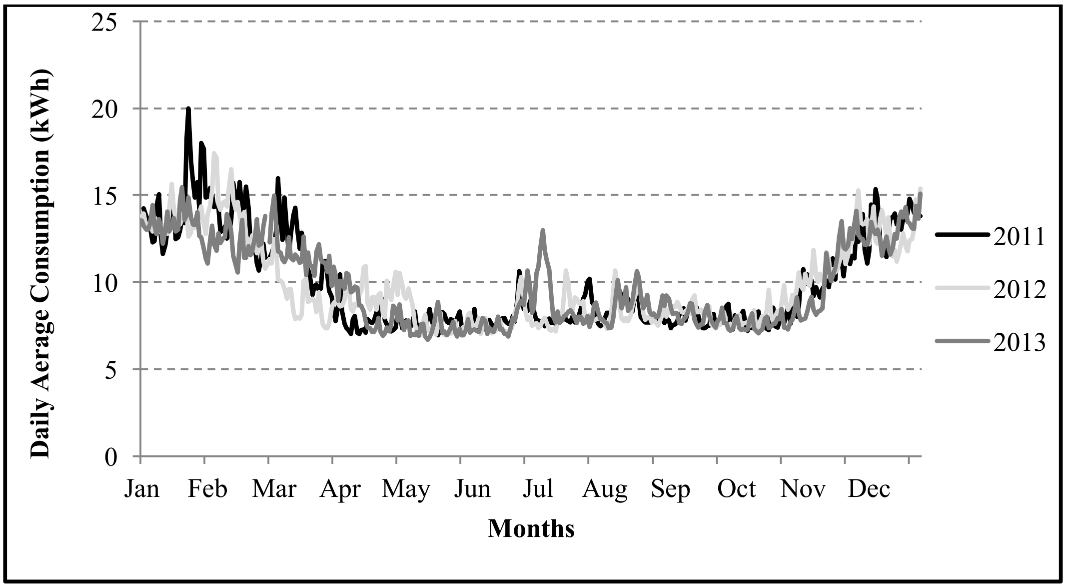

Figure 1 – Daily average consumption for the final sampled households (2011 to 2013)

Table 1 - Summary statistics for the final sampled households daily electricity consumption

Daily Electricity Consumption						
Year	**Nº of days**	**Minimum (kWh)**	**Median (kWh)**	**Mean (kWh)**	**Maximum (kWh)**	**Standard Dev. (kWh)**
2011	364	6.96	8.50	9.86	20.00	2.71
2012	365	6.95	8.83	9.80	17.40	2.43
2013	364	6.71	8.71	9,75	15.45	2.37

Electricity Data Clustering from Smart Meters

The clustering method applied splitting the 230 smart meters dataset in 10 clusters shows a similar distribution of meters (with at least 30 meters per cluster) within clusters with mean daily electricity consumptions below 15kWh (cluster 1 to 6), totalling 200 meters (more than 86%). The remaining 30 meters are included in clusters 7 to 10 fitting the high levels of consumption with daily mean consumption of almost 30kWh (i.e. cluster 10) (Figure 2).

The box-and-whisker plot in Figure 3 unveils the descriptive statistics of the clusters regarding their dispersion and skewness, and the existing outliers. The distribution of consumption data from C1 to C6 is similar, with C1 presenting the lowest statistics and C2 the higher variance. Future work will be carried, by applying statistical tests of hypothesis to assess if there is a significant difference among these six clusters, regarding both its means and variances. Cluster C7 shows the highest data variability while C8 the highest consumption. It is interesting to note that all the clusters have high maximum values that should deserve our attention within a further work, to identify the reasons for their occurrence across all the clusters or if they are outliers.

Under similar climate conditions (all clusters are located in the same city), the consumers have different profiles of electricity consumption, meaning a diversity of drivers behind those segments of consumers, which is the main purpose of this paper.

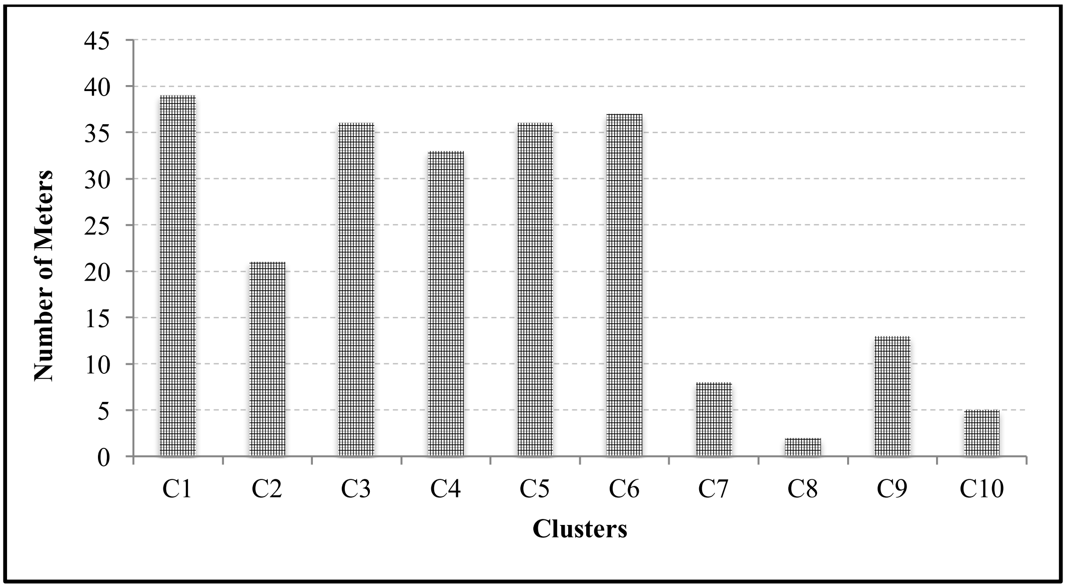

Figure 2 – Number of smart meters allocated per cluster

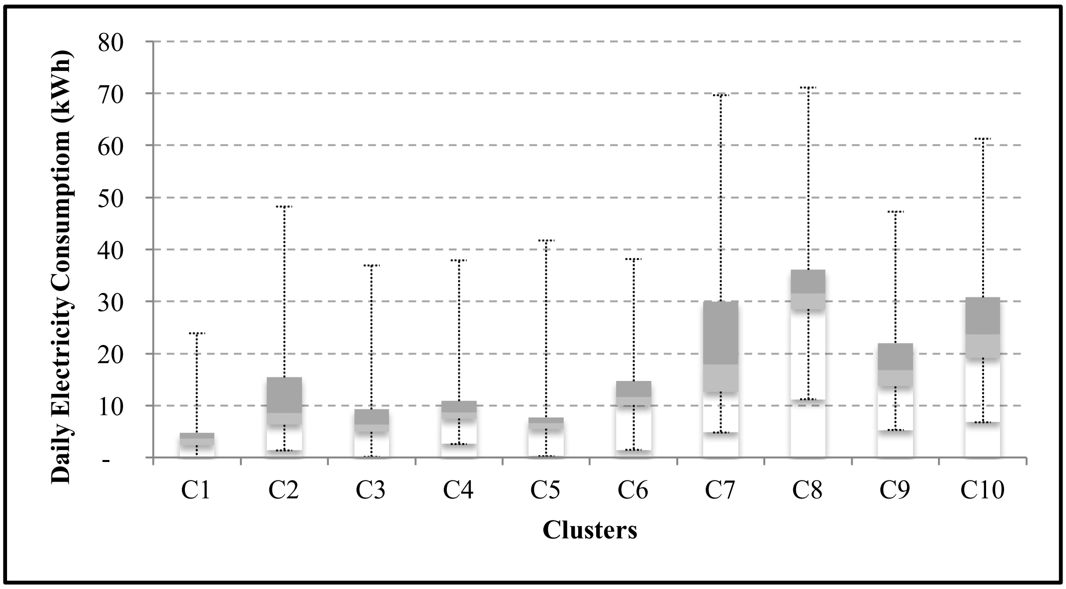

Figure 3 – Box and whisker plot with clusters distribution

By evaluating the survey results for the dwellings in each cluster, it is possible to identify important similarities and differences regarding socio economic determinants, dwellings characteristics and appliances use and ownership, that could explain the different clusters' aggregation and consumption profiles. Considering the statistical behaviour presented previously, three clusters are selected as examples for an in-depth analysis: Cluster 1, Cluster 6 and Cluster 7 (Figure 4). The selection was based on clusters with very distinct profiles 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).

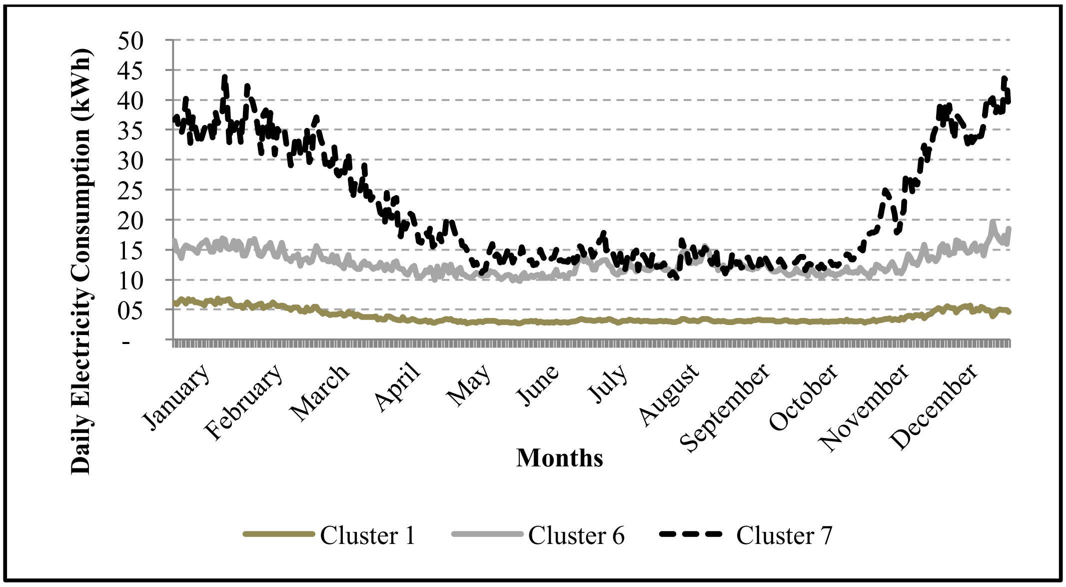

Figure 4 – Daily electricity consumption profiles of Clusters 1,6 and 7 (2011-2013 average)

Cluster 1 presents the lowest average electricity consumption of all clusters with a median of 3.6 kWh and a standard deviation of 2.3kWh. Nevertheless, the meters within this cluster show a distinct consumption along the year, with a clear difference in consumption in the winter months of December, January and February (higher) compared to the other months (around 120% higher). This suggests a strong use of electricity-based technologies for space heating which is corroborated by the findings in Table 2. Cluster 6 is in the middle of the defined smart meters clustering with an average electrical consumption of 11.7 kWh and of 3.9kWh of standard deviation. In opposition to Cluster 1, yearly consumption profile in this cluster does not present significant differences between winter months and the rest of the year (+40%). The last selected example is Cluster 7, which combines a high average consumption (17.9kWh) with a high dispersion (standard deviation of 11.3kWh). Cluster 7 as Cluster 1 presents important differences of consumption in winter month (around more 230%). Table 2 discloses a set of examples of the frequency of occurrence of variables collected in the surveys, to be compared among the chosen clusters in order to derive important factors that describe each clusters aggregation and electricity consumption profile.

Although it is out of the scope of this paper to make a detailed assessment of how the electricity is being consumed inside the households for the different uses (i.e. water heating, lighting, cooling and heating), Figure 5 discloses this information for Portugal (no similar information is available for Évora) aiming to pave the way for a better understanding of the electricity consumption profiles of the clusters presented in this paper. Evaluating the size of intra annual variations of electricity consumption clearly justified by changes of use in electrical equipment for heating (as seen in Figure 4) we would expect a high representativeness of this type use that the one portrayed in Figure 5 for Portugal. This difference could be explained by the variety of heating technologies (fireplaces, electric heaters, HVAC, gas room heaters) being used across different country regions, which is not perceived on national average statistics compared to city level data.

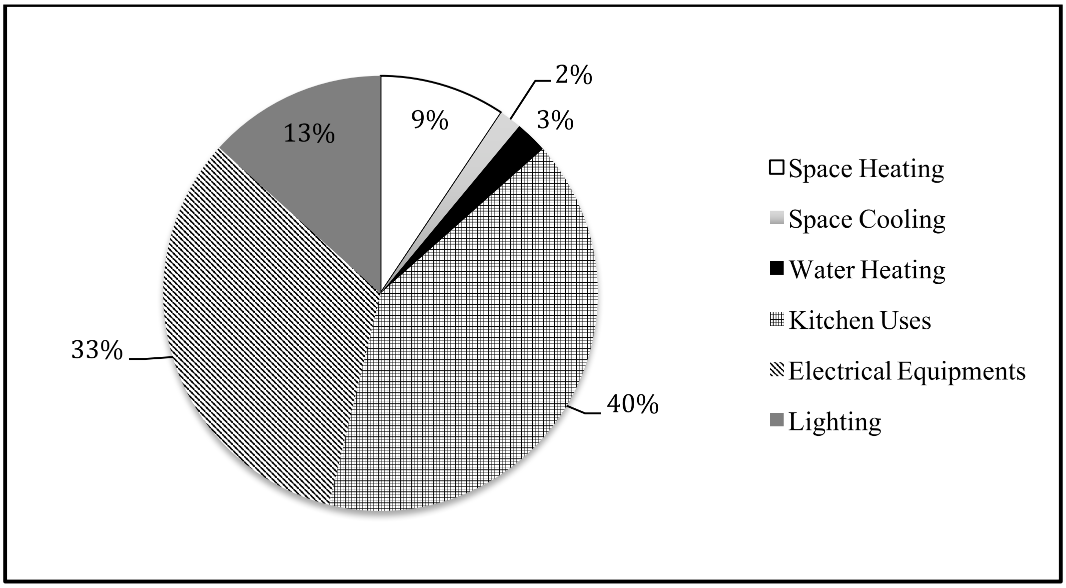

Figure 5 - Electricity consumption in Portugal by type of use in 2010 (INE and DGEG, 2011)

Table 2 - Summary of selected variables explaining the grouping of the meters per Clusters 1, 6 and 7

Explanatory Variables	Cluster 1	Cluster 6	Cluster 7
Characteristics of Dwellings			
Location			
Urban	68%	68%	87%
Rural	32%	32%	13%
Type			
Detached	24%	30%	12%
Semi Detached	30%	35%	38%
Terraced	46%	35%	50%
Age			
Before 1945	22%	14%	50%
Between 1946 and 1990	70%	51%	38%
After 1991	8%	35%	12%
Size			
Under 100m^2	66%	26%	17%
Between 100m^2 and 150m^2	31%	48%	50%
Higher than 151m^2	3%	26%	33%
Type of Glazing			
Single	91%	43%	62%
Double	9%	57%	38%
Type of Window Framing			
Aluminium	39%	81%	50%
Wood	58%	16%	38%
PVC	3%	3%	12%
Bearing Structure			
Concrete	23%	37%	38%
Masonry Wall with and without plate	74%	63%	62%
Masonry wall with loose stone	3%	0%	0%
External Wall Type			
Brickwork double layer	7%	49%	38%
Brickwork single layer	82%	43%	37%
Stone	11%	8%	0%
Rammed Earth	0%	0%	25%
Occupants Characteristics			

Number of Occupants			
1 or 2	73%	24%	37%
3 or more	27%	76%	63%
Age			
Under 18	5%	16%	0%
Between 18 and 50	39%	46%	27%
Older than 50	56%	38%	73%
Education of the Head of the Family			
Before 9th Grade	49%	41%	17%
Between the 9th and 12th grade	35%	41%	0%
Graduation, MSc or PhD	16%	19%	83%
Monthly Average Income			
Below 750€	52%	21%	0%
Between 751€ and 1500€	34%	33%	50%
Above 1501€	14%	46%	50%
Employment Status			
Working Full Time	34%	48%	41%
Retired	41%	20%	41%
Student	21%	24%	9%
Other	4%	8%	9%
Appliances Ownership*			
Heating Technologies			
HVAC (heating, ventilating, and air conditioning)	16%	54%	36%
Fireplace	3%	38%	9%
Electric Heater	78%	19%	55%
Gas Room Heater	0%	3%	0%
Heat Pump	3%	0%	0%
Cooling Technologies			
HVAC	20%	81%	80%
Fan Coil	80%	19%	20%
White Appliances			
Freezers	61%	103%	113%
Dish washers	33%	80%	100%
Microwaves	97%	97%	113%
Contractual Power Characteristics			
Contracted Power			
Lower than 6,9kVA	84%	19%	12%
6,9kVA and higher	16%	81%	88%
Type of Tariff			
Single	73%	62%	62%
Dual	27%	38%	38%

*When percentages are higher than 100 % it means that some houses own more than one equipment.

Cluster 1: As depicted in the previous table, *Cluster 1* is characterized by a predominance of terraced dwellings located in urban areas, small houses (under 100m^2) built between 1946 and 1990 period. The majority of the houses have single glazing and wooden window framing. The prevailing bearing structure is masonry walls with or without plate associated with brickwork single layered in the external walls.

When evaluating occupants characteristics, we can say that this clusters' houses are portrayed by small families (2.2 persons per household), generally older than 50 years old with low levels of education (before 9th grade), retired and with households monthly average incomes below 750€.

The electricity profile of this cluster, with a strong difference of consumption on winter months is backed up by the survey results with predominant ownership and use of electric heaters and HVAC systems for space heating (96%). Only 56% of this clusters houses have cooling equipment. From which, 80% of the houses have fan coils that consume a lot less than HVAC systems.

The overall smallest ownership of white appliances from all the clusters combined with the dominant number of houses (81%) with low contracted power (under 3.45kVA) allows also explaining the lowest levels of daily electricity consumption in this cluster when compared to others. 73% of the houses in this cluster still have single tariffs not taking advantage of the lowest prices at night of dual tariffs.

Cluster 6: *Cluster 6* presents a similar distribution of urban and rural households as *Cluster 1,* therefore not being a factor of difference between the clustering. This cluster presents an even distribution of occurrence between the three considered types thus no being a distinct explanatory variable distinguishing the houses in this cluster.

The characteristics of the dwellings describing these cluster dwellings are: average size houses with floor areas between 100 and 150m^2, built after 1946 but with a high share built after 1991, also shown in the higher amount of concrete houses. Following other studies results (e.g. Leiwen and O'Neill, 2003) higher average household area reveals also higher energy consumption. The sampled houses of this cluster have similarly distribution of single and double-glazing but the majority of them have aluminium framing in the windows (81%).

This cluster is established by higher number of occupants inside the households (3.2) with 77% of the households having more than 3 persons, also contrasting with *Cluster 1* regarding the age of occupants, household income and employment status. 62% of the occupants aged below 50 years and 48% working full time reflected on higher levels of monthly income (i.e. 46% of houses with incomes above 1501€).

Income relates with the ownership of electrical equipment both impacting the quantity and quality of the appliances (e.g. Reiss and White, 2005). A large body of literature has also concluded that energy consumption increases with income (Kaza, 2010; Cayla et al., 2011; Brounen et al., 2012). However, the opposite have also been identified by other studies (e.g. Foster et al., 2000).

This cluster presents a totally different profile of households occupants compared to *Cluster 1*, having significant impacts on the differences of electricity consumption patterns along the year. 73% of the houses have electric heaters or HVAC for space heating (the majority bought after the 2005 summer heat wave in Portugal) but it is also in this cluster that fireplaces (both with and without heat recovery) have the higher penetration (38%). 78% of the houses in this cluster own equipment for cooling.

The penetration of both freezers (103%) and dish washing machines (80%) are higher in this cluster reflected in the higher daily electricity consumption. Larsen *et al.* (2010), Kaza (2010) and Gram-Hanssen (2011) present the number and the use of appliances correlated to the number of people living in the house; but for Kaza (2010) 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 (Larsen *et al.*, 2010; Wiesmann *et al.*, 2011).

Regarding the contracted power, the majority of these cluster households have 6.9kVA (72%), with a dominance of single tariffs contracts (61%) as can also be seen in all the other clusters.

Cluster 7: The yearly electricity profile of the houses in *Cluster 7* are, as in *Cluster 1*, also portrayed by a distinction within electricity consumption in winter months and the rest of the year, but with higher daily mean electricity consumption and higher dispersion.

The households on this cluster are predominantly located in urban areas (87%), with a strong predominance of old houses (50% built before 1945) with high floor areas. Similarities within important explaining determinants of electricity consumption such as household occupants also occur in this cluster. Near 65% of the households have more than 2 persons per household; 73% older than 50 years; 83% with high education levels (i.e. Graduation, MSc or PhD) and even distribution between retired persons and working as full time professionals. As can be stated by comparing Clusters 6 and 7 on household occupants and also suggested by Brounen et al. (2012) and Kavousian et al. (2013) there is a non-linear relationship between household electricity consumption and the number of occupants. With larger households having higher aggregate electricity consumption but lower per capita consumption.

No clear similar dwellings characteristics as bearing structure, type of wall and windows arise to distinctively characterize this cluster, probably also explaining the high deviation of electricity consumption.

As in *Cluster 1,* electric appliances for space heating and cooling are dominant in this cluster. This cluster has even a higher penetration of white appliances such as microwaves (113%), dish washing machines (100%) and freezers (113%) than *Cluster 6,* also reflected in higher the levels of electricity consumption. As expected by the high average electricity consumption profile of these cluster households, 88% of the houses have a contracted power of at least 6.90kVA.

According to this three clusters evaluation, we can say that tariff and contracted power while being similar to several clusters are not paramount explanatory variables of the segmentation. Furthermore we might also conclude that the urbanisation levels and bearing structure are also not variables which distinct the clustering. Other determinants collected in the surveys which also not make a distinction between clusters are: high penetration of compact fluorescent lamps instead of incandescent ones, ownership of laptops per household; ownership above 100% of refrigerators in some households of the higher consumption clusters; high penetration of cloth washing machines (near 100%); preference of use of the electric heaters instead of HVAC systems.

Conclusions

This paper presents an exploratory data analysis of the annual electricity consumption profiles from daily consumption data of a sample of 230 households with smart meters data to better understand the electricity consumption in the residential sector of a Southwest European city. This data was combined with an extensive door-to-door survey allowing a clustering analysis supported on their annual consumption profile and socio economic characteristics.

From the analysis we conclude that three major groups of determinants influence residential electricity consumption segmentation: physical characteristics of a dwelling 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).

Despite the relevant outcomes of this work, there are some limitations, e.g. incomplete responses of data for some of the surveys, justified by difficulties regarding technical questions such as insulation type and thickness and difficulties to assess electrical appliances daily use, which have impacts on the clustering evaluation. Further work will encompass a complete assessment of the sampled households electricity consumption determinants available in the survey to identify the relative importance of each one within this smart meters data set including all clusters in the analysis performing a statistical analysis evaluating the significance in the differences across the clusters. Further statistical analysis will also be carried including daily electricity consumption for the year 2014.

The empirical work that we have conducted advances the knowledge on household consumption patterns. Besides of the identification of the factors characterizing selected electricity profiles, this paper discloses the importance of the future widespread use of smart meters, which provide: 1) sufficient information to support the design and implementation of energy reduction policies targeting specific groups of consumers based on their socio economic characteristics and energy use profile. This knowledge could also be used as a starting point for utilities looking at peak shaving and electricity demand shifting inside households derived from market segmentation.

References

Armel, K., Gupta, A., Shrimali, G., Albert, A. 2013. Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy 52 (2013) 213-234.*

Cramer JC, Miller N, Craig P, Hackett BM. 1985. Social and engineering determinants and their equity implications in residential electricity use. Energy 1985;10 (12):1283–91.

Baker KJ, Rylatt RM. 2008. Improving the prediction of UK domestic energy-demand using annual consumption-data. *Applied Energy 2008;85(6):475–82.*

Bartiaux F, Gram-Hanssen K. 2005. Socio-political factors influencing household electricity consumption: a comparison between Denmark and Belgium. In: Proceedings of the ECEEE 2005 Summer Study, European Council for an Energy Efficient Economy; 2005. 1313–1325.

Bartusch C, Odlare M, Wallin F, Wester L. 2012. Exploring variance in residential electricity consumption: household features and building properties. *Applied Energy 2012;92:637–43.*

Bedir M, Hasselaar E, Itard L. 2013. Determinants of electricity consumption in Dutch dwellings. *Energy and Buildings 2013;58:194–207.*

Brounen, D., Kok, N., & Quigley, J. M**.** 2012. Residential Energy Use and Conservation : Economics and Demographics (p. 31). *European Economic Review 56 (2012) 931–945*

Cayla, J. M., Maizi, N., & Marchand, C. 2011. The role of income in energy consumption behaviour: Evidence from French households data. *Energy Policy, 39(12), 7874–7883. doi:10.1016/j.enpol.2011.09.036*

DGEG, 2014. Energy balances. Directorate for Energy and Geology. Available at: [www.dgeg.pt]

EC, 2014. *Climate Action - 2030 framework for climate and energy policies.* European Commission. Available at: http://ec.europa.eu/clima/policies/2030/index_en.htm

EDP Distribuição, 2015. *Évora Inovcity – Smart Energy Living.* EDP Distribuição S.A. Available at: [www.inovcity.com]

Eurostat, 2011. *Energy Statistics.* Eurostat. European Commission. Available at: [http://epp.eurostat.ec.europa.eu/]

Gouveia, J. P., Seixas, J., Bilo, N., Valentim, A., Nunes, V., Giannakidis, G., Robinson, D., Irons, D., Gargiulo, M. 2014. *Integrative Smart City Planning - Buildings and Mobility in Évora.* Presented at the 4th IAEE European Energy Conference - Sustainable Energy Policy and Strategies for Europe, At LUISS University, Rome, Italy. 28-31 October 2014. Available at: [http://iaee2014europe.it/]

Gram-Hanssen K, Kofod C, Petersen KN. *Different everyday lives: different patterns of electricity use*. In: Proceedings of the ACEEE 2004 Summer Study, American Council for an Energy Efficient Economy; 2004. 7:74–85.

Gram-Hanssen, K., 2011. *Household's energy use – Which is more important: efficient technologies or user practices?* World Renewable Energy Congress 2011. 8-13 May 2011, Linkoping, Sweden.

Gruber, E., Scholmann, B., 2006. *The current and future electricity demand of appliances in German households.* Conference Proceedings of International Energy Efficiency in Domestic Appliances & Lighting Conference (EDAL'06), London, 21-23 June.

Halicioglu, F. (2007). Residential electricity demand dynamics in Turkey. *Energy Economics, 29(2), 199–210. doi:10.1016/j.eneco.2006.11.007*

INE and DGEG, 2011. *[Survey on Energy Consumption for the Residential Sector 2010] Inquérito ao Consumo de Energia no Sector Doméstico 2010.* National Institute of Statistics and Directorate for Energy and Geology.

Jones, R., Fuertes, A., Lomas, K. 2015. The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable and Sustainable Energy Reviews 43 (2015) 901–917.*

Kavousian, A., Rajagopal, R., Fischer, M. 2013. Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behaviour. *Energy 55 (2013) 184-194.*

Kaza, N. 2010. Understanding the spectrum of residential energy consumption: A quantile regression approach. *Energy Policy*, *38*(11), 6574–6585. doi:10.1016/j.enpol.2010.06.028

Kelly, S. 2011. Do homes that are more energy efficient consume less energy?: A structural equation model of the English residential sector. *Energy 36 (2011) 5610-5620. doi:10.1016/j.energy.2011.07.009*

Lam, JC. 1998. Climatic and economic influences on residential electricity consumption. *Energy Conservation Management 1998;39(7):623–9.*

Larsen, T., Knudsen, H., Kanstrup, A., Christiansen, E., Gram-Hanssen, K., Mosgaard, M., Brohus, H., Heiselberg, P., Rose, J. 2010. Occupant influence on the energy consumption of Danish domestic Buildings – State of the art. Aalborg Universitet, Institut for Byggeri og Anlaeg.

Leahy, E, Lyons, S. 2010. Energy use and appliance ownership in Ireland. *Energy Policy 2010;38(8):4265–79.*

Lee, S., Whaley, D., Saman, W. 2014. Electricity Demand Profile of Australian Low Energy Houses. *Energy Procedia 62 (2014) 91-100.*

Lomas, K. 2010. Carbon reduction in existing buildings: a transdisciplinary approach, *Building Research & Information, 38:1, 1-11.*

McDonald, M, Dunbar, I. *Market segmentation: how to do it and how to profit from it.* 4th ed. Chichester: John Wiley & Sons Ltd.; 2012.

McLoughlin F, Duffy A, Conlon M. 2012. Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. *Energy and Buildings 2012;48:240–8.*

McLoughlin, F., Duffy, A., Conlon, M., 2015. A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied Energy 141 (2015) 190–199*

Raty, R., Carlsson-Kanyama, A., (2010). Energy consumption by gender in some European countries. *Energy Policy 38 (2010) 646-649.*

Reiss, P., White, M., 2005. Household electricity demand, revisited. *Review of Economic Studies 2005: 72(3):853-883.*

Santamouris M, Kapsis K, Korres D, Livada I, Pavlou C, Assimakopoulos MN. On the relation between the energy and social characteristics of the residential sector. *Energy and Buildings 2007;39(8):893–905.*

Seo, Y., Hong, W., 2014. Constructing electricity load profile and formulating load pattern for urban apartment in Korea. *Energy and Buildings 78 (2014) 222–230*

Statsoft, 2015. *How to group Objects into Similar Categories, Cluster Analysis*. Available at: [www.statsoft.com/textbook/cluster-analysis]

Tso GKF, Yau KKW. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. *Energy 2007;32 (9):1761–8.*

van Raaij, F., Verhallen, T.. 1983. A behavioral model of residential energy use. *Journal of Economic Psychology 3 (1983) 39–63.*

Ward, J.,1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Association, Volume 58, Issue 301, 236-244.*

Weiss, M., Mattern, F., Beckel, C., 2013. *Smart energy consumption Feedback – Connecting smartphones to smart meters.* In ERCIM news, Number 92, January 2013.

Wiesmann, D.; Lima Azevedo, I., Ferrão, P., Fernández, J., 2011. Residential electricity consumption in Portugal: Findings from top-down and bottom up models. *Energy Policy 39 (2011) 2772-2779.*

Wyatt P., 2013. A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England. *Energy Policy 2013;60:540–9.*

Acknowledgements

The work supporting this paper was partly funded by the EU project INSMART, Integrative Smart City Planning, under grant agreement no.: 314164 and by Portuguese Science and Technology Foundation (FCT) through the Scholarship SFRH/BD/70177/2011. We would like also to thank EDP Distribuição, S.A., especially Vera Nunes and Miguel Andrade for their contribution regarding smart meters data.