Household Electricity Consumption Routines and Tailored Feedback

Joana Abreu, MIT Portugal Program
Francisco Câmara Pereira, Coimbra University

ABSTRACT

Evidence suggests that rational alternatives do not necessarily drive the choices made by real individuals in the privacy of their homes, Therefore, an understanding of the characteristics of the utility's customer base, is useful to help design utility programs that target the customer's ordinary daily practices.

This paper summarizes the main results of a method to highlight familial habits by datamining electricity consumption datasets. Target families were participants in a small-scale experiment that resembles a utility led advanced meter infrastructure implementation.
The process highlights load shapes, which occur more frequently for days with similar characteristics. Approximately 80% of the household electrical use can be explained within those load trends. These include: load shape for unoccupied baseline, hot working days, temperate days, cold working and cold weekend days. We conclude with a discussion of possible applications of the electricity consumption trends for feedback and segmentation.

Introduction

With the widespread adoption of smart metering in the residential sector in the U.S. and in Europe, regulatory authorities and policy makers require utilities and retailers, to provide frequent feedback to utility customers. In parallel, responsive energy demand is regarded as a potential enabler of residential energy efficiency. In this scenario, the demand can be controlled directly by the utility or retailer[1] or it can be controlled indirectly, by the homeowner, and in this case, automated algorithms and feedback can be used to replace or support human decision-making (Rochlin, 2009). Smart metering is a key enabler of responsive demand programs, as it modernizes the electric grid and enables frequent feedback mechanisms (Vasconcelos, 2009).

Usually led by electricity retailers, utilities or energy service companies (ESCO's), demand response (DR) programs involve changing the shape of the demand curve in periods of constraint. DR may imply constant adjustments to frequency fluctuations (load shaping), moving loads away from peaks to fill in valleys (load shifting) or curtailment of specific appliances through direct or indirect control by the utility (load shedding; Rochlin, 2009). In parallel, responsive load programs can be systematised according to the program objective to improve system reliability or reduce electricity provision costs.

Programs aiming at improved system reliability focus primarily on active load response, which can occur by direct control, curtailment, interruption of service or scheduling of loads, and for emergency situations, while those that aim at reducing the system operation costs are price responsive. In this case, customers receive incentives through a combination of pricing structures such as time of use, dynamic pricing or demand bidding (Conchado et al., 2010). The widespread

1 Untraditional players such as Telecoms, are beginning to see a role in home energy efficiency, see for example the case of Comcast with the company Ecofactor (where demand response is done via thermostat control), in the United States or the case of Portugal Telecom and the company Intelligent Sensing Anywhere (which provides customers with feedback about electricity use) in Portugal.

installation of smart meters, *per se*, doesn't leverage the potential held by the residential consumers as active parties in the electricity market, and public reaction to utility controlled demand response (see the example of Baskerville, California, by PG&E) provide evidence that it is necessary to develop consumer centric programs that rely on the characteristics of the population.

Traditional load segmentation methods handle low voltage data in aggregate, and are not sufficiently precise to discriminate intrinsic patterns, and regional differences of electricity consumption. Segmentation typically relies in contracted power and yearly consumption volume, (Matos *et al*., 2005) or in a combination of those with demographic characteristics. Recent research shows that even for households with similar demographics, the electrical consumption signature can be very diverse. An example is when work routines disrupt familial consumption patterns (as in when parents work shifts).

The methodology proposed in this paper is agnostic of previous knowledge about the households. Discrimination of electricity consumption patterns is therefore attained through machine learning methods, in particular, Fourier analysis, principal components analysis and clustering. The methodology, however, was applied over a small number of households, and focuses in highlighting patterns of consumption, which are characteristic of seasonal or weather variations. Future work will aim at testing the generalization of the method over increasingly larger pilot samples.

Objectives

The main objective of this research is to shed light into household routines, which can be extracted anonymously and independently of behaviours reported by individuals. We postulate that it may happen that to describe regional differences, it may only be necessary to focus attention on particular patterns or *personas*, which may be the result of the clustering of millions of individual profiles

Characteristics of the trial

The research presented in this paper is the outcome of datamining to residential electricity consumption that resulted from a small-scale trial assembled within the MIT-Portugal education and research cooperation Program. German electricity retailer YelloStrom, B.V. provided the electronic metering apparel, which was coupled with the circuit board of the houses, rather than replacing existing meters of the utility, for reasons of compliance with local regulations. The trial involved 15 participants and took place in the region of Oeiras, Portugal. Participant homes could monitor real time and historical consumption via computer interfaces.

Architecture

Coupled to the circuit board of the house, the electronic meter collected readings every second. These were transmitted through the power line to a communication's module, plugged to an Internet modem. Within the bounds of the home area network, the occupants could monitor the real time impact of turning on electrical appliances in a computer, as well as aggregated daily and historical consumption. The 15-minute aggregate readings were transmitted to a central server through the Internet. Apart from the selection process, which was based on randomization of a self-selected sample, and the number of households involved, the trial is similar to the technological solutions of existing utility based automated metering infrastructure.

Email invitations were sent to a population of 880 individuals, who worked for the municipality. A subgroup of 300 was directly invited to participate, by phone. Adherence rate was 7%, of which, 35% ended up abandoning the trial before the meters were installed.

Methodology

a) Data preparation – data received is filtered and prepared for analysis (includes interpolation of small errors). Preliminary exploratory statistical analysis includes correlation, and factor analysis.

b) Routine analysis. The method used is spectral analysis (Fourier analysis), and the objective was to isolate the most important routine cycles in the household. Given that the 24 hour cycle was the most important, the dataset originally sampled every15-minutes, was aggregated in "hourly density".

c) Feature extraction and pattern recognition. Principal components analysis is a method of dimensionality reduction, which is used to highlight the variables that construct the most important routines of a day for each individual household. Selected eigenvectors reflect up to 80% of total variability in electricity consumption. The coefficients (named in the literature "eigencoefficients) were clustered by a combination of hierarchical (unsupervised) and K-means methods, with the objective of combining similar days in groups. This method highlights daily patterns of electricity consumption, heavily correlated with weather and other loading characteristics (day of the week, season, month). Clustering conditions was: minimum 15 and maximum 150 days for datasets comprehending more than 3 months of data, and minimum 8 and maximum 30 for smaller datasets (table 1). The attributes (or loading conditions) that classify each day are: average daily temperature, weekends and official holidays: these were bundled together and constitute another descriptive characteristic of the data; working days; month of the year. The method is better described in Abreu *et al.* (2012). An example of the clusters generated by this method is presented in the next section. Each cluster comes with a representation of the temperature density distribution for a specific cluster and pie chart representation of day of the week and monthly proportion. Normal spread of weekends (including holidays) *vs* weekdays is 29:71

Table 1. Dataset dimension per household

Hsh.code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Days	434	441	377	441	112	432	446	185	359	308	428	523	108	67	408

d) Validation. This task includes the confrontation reports collected from the participants. The validity of the datamining results were cross checked against the responses to a survey and in depth interviews which provided the general overview of appliance ownership, and the more specific data about familial habits and seasonal routines. The method of interviewing was specifically designed to prevent contamination from results to the way the interview was conducted. Whenever possible, the interviewer approached all the adults living in a household with the same questions about "normal daily routines, and approximate time schedules", and seasonal variations: "when did you turn the heating on", "when did you go on vacation" or "Can you describe what, a typical Winter day, is in your household".

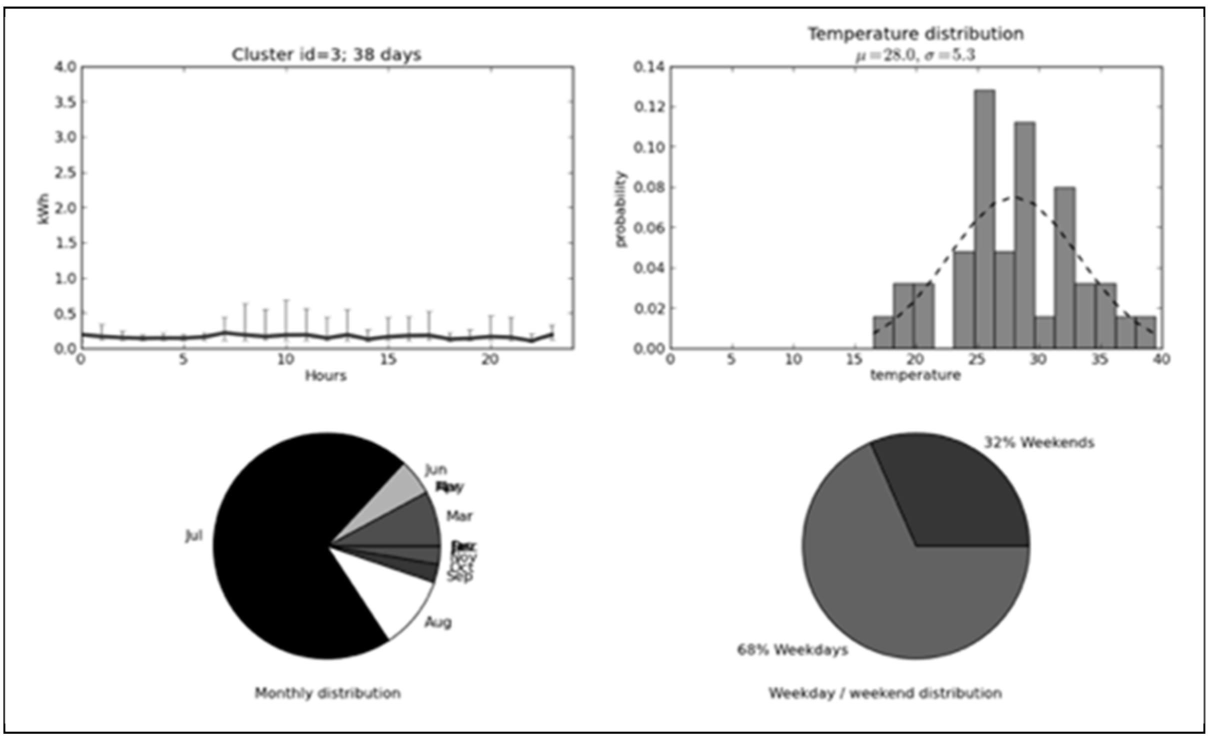
Results

Large Datasets

We consider large datasets to be a dataset with more than 3 months of collected data. The analysis method revealed that everything else remaining more or less stable (e.g.: the household keeping more or less same electrical appliances, and the number of occupants for a time period) is it possible to cluster "typical" baselines that depend on the loading-day conditions (e.g.: season, external temperature or day of the week).

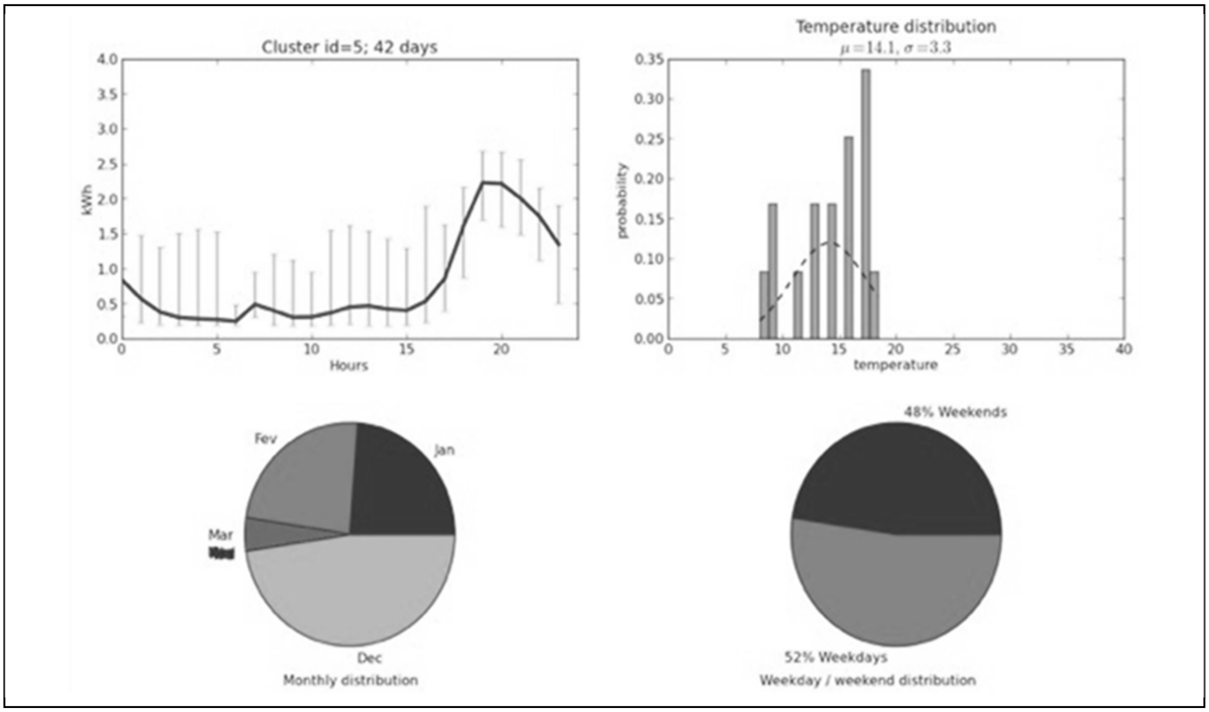
Baseline. One of the most consistent clusters is the baseline of the house, or the electricity consumption that occurs when the house is mostly unoccupied. The example is of household 2, a dataset containing 441 days of 15 minute sampled data (figure1). This is a cluster of 38 days and shows mostly the effect of standbys and off-modes, and cyclic appliances, which are characteristic of when the house is mostly vacant (during holidays, weekends off). The cluster is more frequent in July, but can also happen in August and in the spring. There's an even distribution between weekends and working days.

Figure 1: household 2 – baseline



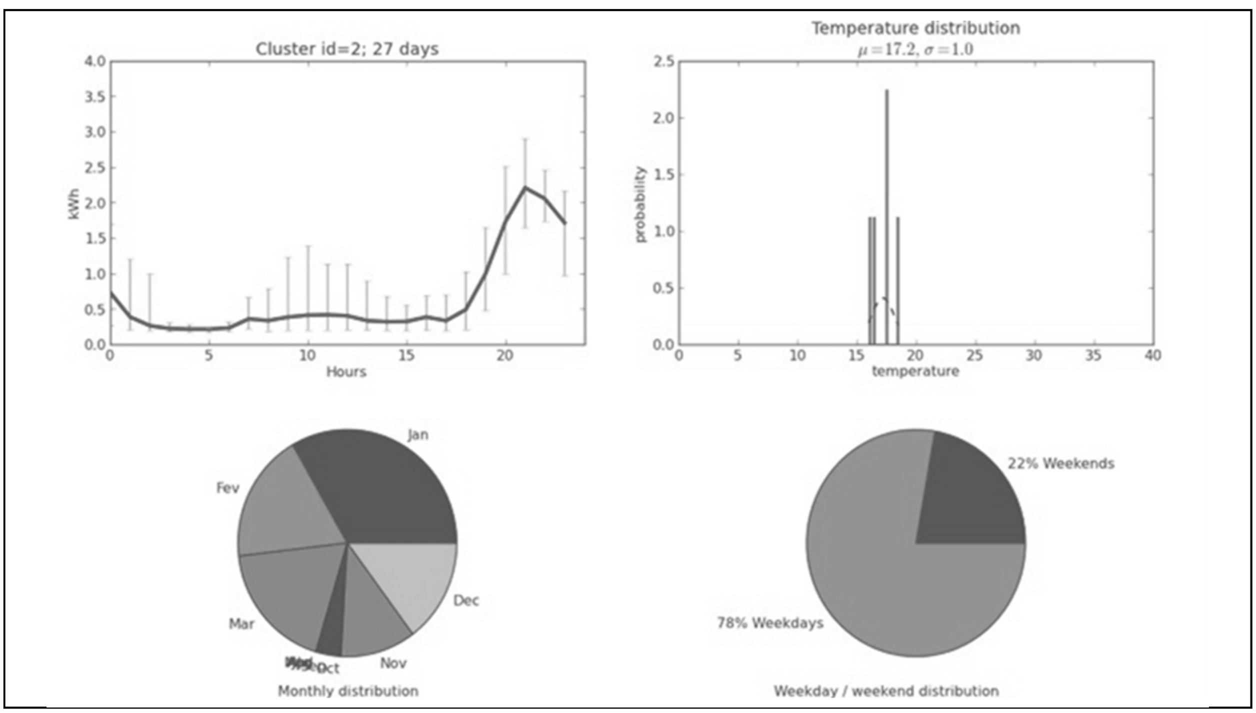
Weekends in winter. This is a 42 day cluster for household #6, which contains 436 days. It is possible to see a large range of error during the day, which is meaningful of families staying home during the day. Evening peak is high and wide. The cluster incudes a large percentage of weekends and the monthly distribution includes mostly Winter (figure 2).

Figure 2: Load profile for f#6: Winter Weekends



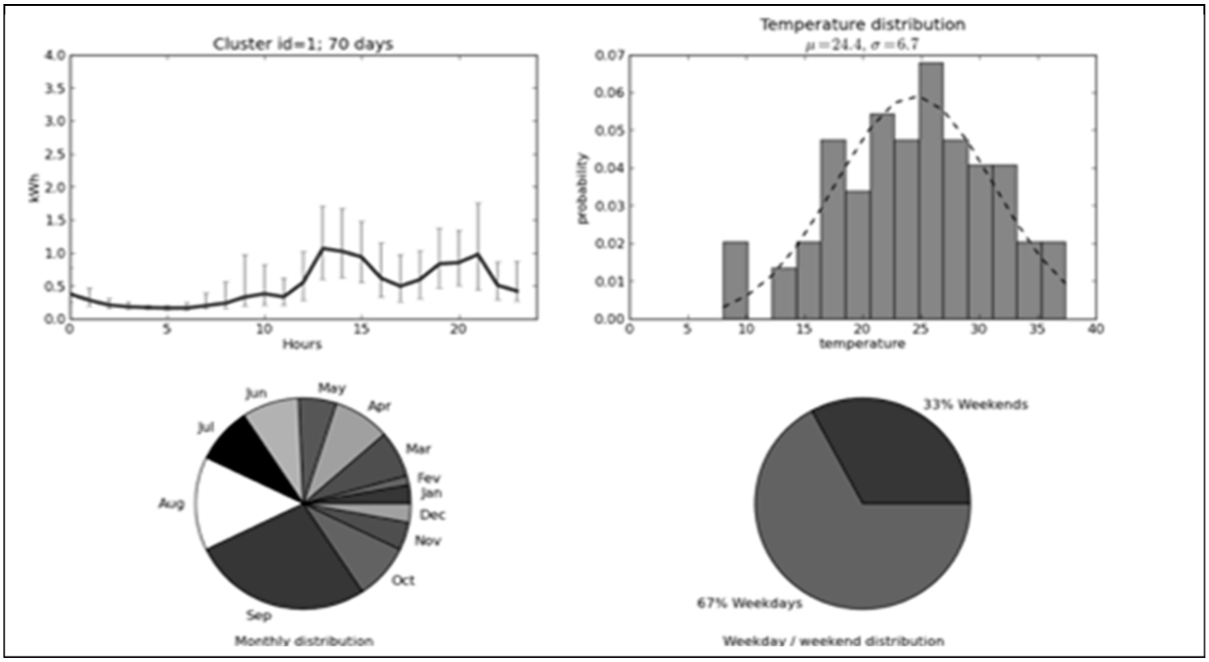
Working days in winter. This 27 day cluster for household #6, still includes winter months. Yet there's a larger input of working days in this cluster (figure 3).

Figure 3: Load profile for #f6: Winter Weekends



Temperate days. Common across the year and overemphasizing warmer temperatures, this profile includes 70 days, and shows an evening and afternoon peaks. Lunch peak is the effect of home lunch and the washer cycle, and HVAC. Load shifting is an explanation for the proportionally smaller evening peak (figure 4, household #2).

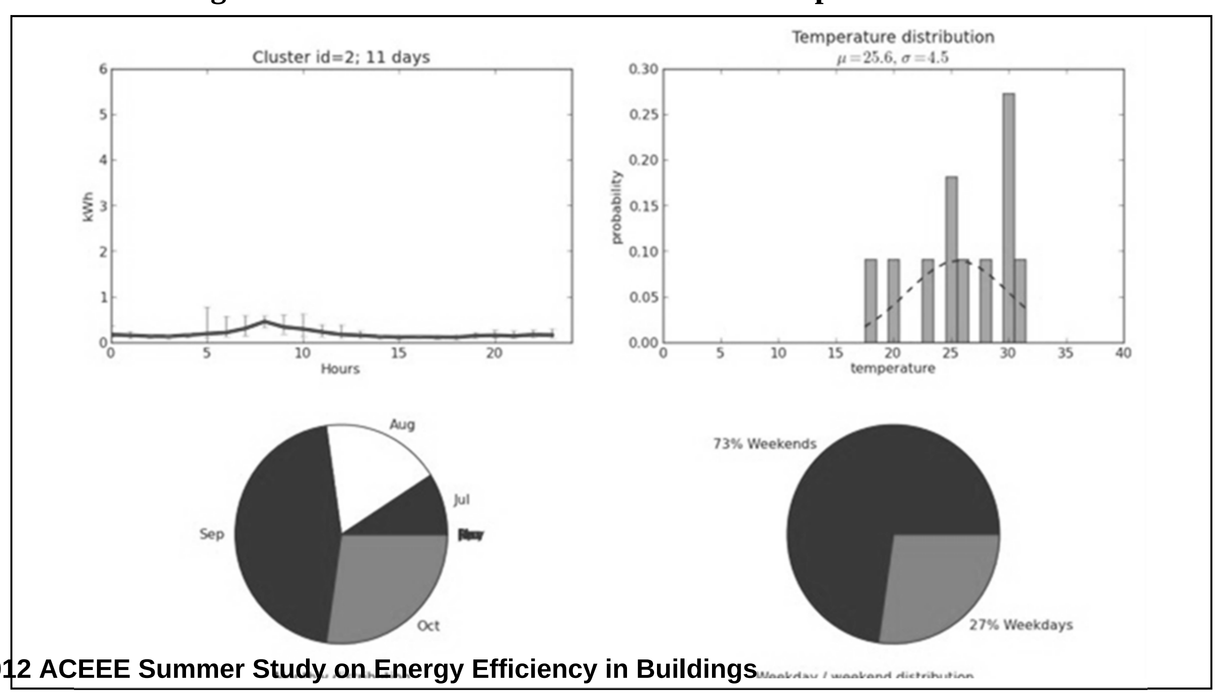
Figure 4: Load profile for #f2: Warm days distributed between weekend and weekdays



Small Datasets

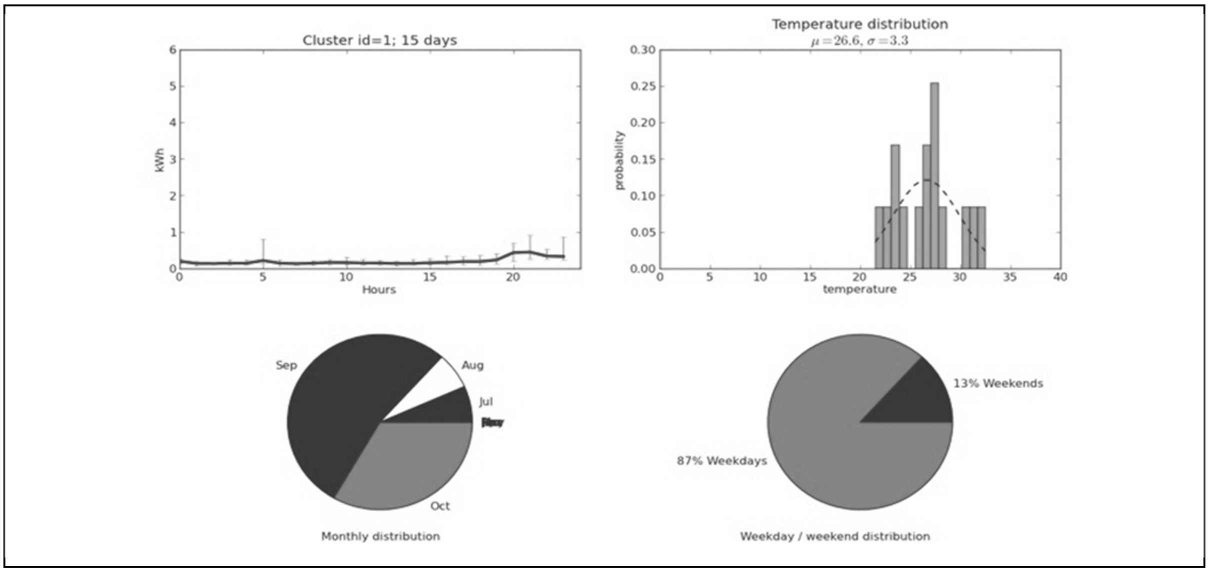
Household #13 illustrates the effect of extraction of baselines from small datasets. Clustering thresholds were 8 min. and 30 max. Cluster (Figure 5) indicates typical weekend behavior for September and October.

Figure 5: Load Profile for f#13 – Weekends September / October



While figure 6 shows the behavior of the same family, for the same months, during working days. The observation of the results for small datasets clearly shows that it is possible to discriminate better week cycles.

Figure 6: Load Profile for f#13 – Working days September / October



Conclusions

The research conducted and exemplified in this paper, shows that it is possible to semi-automatically extract specific baselines based upon actual consumption behavior. These baselines reflect household routines for specific weather and season characteristics. As such, for the majority of the households, it is possible to identify the profiles that correspond to:

- *Baseload* when the occupants are not at home. This baseline reflects the impact of refrigerators and freezers, appliances set in standby or off modes.
- *Hot working days* when the temperatures are high and happening more frequently during the working week
- *Temperate days*. This baseline happens more frequently in working days, for moderate temperatures when the households don't need to turn the electrical heaters.
- *Cold working days* happening during the winter, which reflect electrical heating.

- *Cold weekend days* happening during the winter, which reflect electrical heating and occupation during the day.

The following describe usefulness of the integration of the method in automatic algorithms:

a) Discrimination of electricity consumption routines for individual households may be a useful for electricity service companies, in particular for automatic service personalization, since the routines reflect the lifestyle of the families. This could be attained by the integration of the algorithm in the "energy box" of the household, and improve feedback, and the setting of goals. For example, given the predicted routines in respect to forecasted weather conditions, the software could be designed to issue alerts for out of the ordinary events. Other uses could include those that rely on actual lifestyles to optimize electrical vehicle charging, for example.

b) Segmentation of the customer base is useful for utilities and retailers. Even for the same customer service region, differences in geography, tradition, affluence or whether the population is urban or rural, can influence level of consumption and routine profiles. In this respect, a business intelligence software tool, that can highlight specific customer profiles, on top of electronic metering, can give utilities and retailers the capacity to recognize specific differences and offer services triggered to the actual characteristics of their customers, and help improve the general efficiency of the system

References

Abreu, J.M., Câmara Pereira, F. and Ferrão, P. (2012) Using pattern recognition to identify habitual behavior in residential electricity consumption. *Article in Press Energy and Buildings.*

Cohen, D., & Farley, T. a. (2008). Eating as an automatic behavior. Preventing chronic disease, 5(1), A23. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2248777&tool=pmcentrez&rendertype=abstract

Conchado, A., & Linares, Pedro. (2010). The Economic Impact of Demand-Response Programs on Power Systems . A survey of the State of the Art . The economic impact of demand-response programs on power systems . A survey of the state of the art. Power.

Heijs, W. (2006). Household Energy Consumption. In W. Heijs (Ed.), Household energy comsumption. Habitual Behavior and Technology. Springer.

Matos, M. a, Fidalgo, J. N., & Ribeiro, L. F. (2005). Deriving LV Load Diagrams for Market Purposes u (pp. 105-110). IEEE. doi:10.1109/ISAP.2005.1599248sing Commercial

Information. Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems.

Rochlin, C. (2009). The Alchemy of Demand Response: Turning Demand into Supply. The Electricity Journal, Elsevier, 22(9), 10-25. doi:10.1016/j.tej.2009.09.004

Vasconcelos, J. (2009). Energy regulation in Europe : regulatory policies and politics of regulation. European Review, 3(3), 1-12.