DELEVOPMENT OF A METHODOLOGY FOR FORECASTING ELECTRICITY-PRICE SERIES TO IMPROVE DEMAND RESPONSE INITIATIVES

Antonio GUILLAMÓN
U.P. de Cartagena – Spain
antonio.guillamon@upct.es

Mari Carmen RUIZ
U.P. de Cartagena – Spain
maricarmen.ruiz@upct.es

Antonio GABALDON
U.P. de Cartagena – Spain
antonio.gabaldon@upct.es

Sergio VALERO
U. Miguel Hernández – Spain
svalero@umh.es

Carlos ALVAREZ
IIE-UPV – Spain
calvarez@die.upv.es

Mario ORTIZ
U. Miguel Hernández – Spain
mortiz@umh.es

ABSTRACT

The main objective of electricity regulators when establishing electricity markets is to decrease the cost of electricity through competition. However, this scenario can not be achieved without the full participation of the electricity demand. The aim of this paper is to propose a procedure, through the detection of electricity-price patterns based on what happened in the energy markets the previous day, which could help customers and aggregators to take decisions in Electricity Markets. In this way, the capacity of a methodology (Statistical Ward's Linkage) to classify and forecast high electricity market prices is analyzed. Besides, some price-patterns were found in the abovementioned clusters. The knowledge contained within these patterns supplies customers with market-based information on which to focus its energy use decision to improve the usefulness of Demand Response Initiatives.

INTRODUCTION

The participation of customers in electricity markets is a basic concern for achieving a better market operation. The market will not be complete until demand and supply sides could compete on an equal footing and have "similar" possibilities and products to participate both in energy and ancillary markets [1].

U.S. Department of Energy, reported [2] that "the most important benefit of demand response initiatives (DRI) is improved resource-efficiency of electricity generation and transmission due to closer alignment between customer electricity prices and the value they place on electricity". This increased efficiency should create a variety of benefits: participant benefits (the bill savings and costs earned by customers that adjust their electricity demand in response to prices), market-wide financial benefits (to mitigate the ability to exercise market power by raising supply power prices) and reliability benefits (DRI lowers the likelihood and consequences of outages).

Unfortunately, small and medium sized customers face serious barriers to participate in electricity markets: the lack of real price information which makes difficult any investment or change in energy patterns [3], the minimum size of demand for DRI eligibility (sometimes hundreds of kW), the requirements for metering and communication resources, the complexity of market rules, and finally how and why to change their demand if the price varies. In this way, the role of "energy aggregators" becomes a necessity in actual markets to help customer response.

This situation promotes the need for "simple" tools that allow "aggregators" to identify customer groups in the market and determine the products that best suit each customer segment. From an economic point of view this can be carried out by the use of elemental data, such as daily demand profiles [4], simulation of the load [5], and the results of market prices in the past. Moreover, the DR portfolio should be easy to understand and apply with the help of enabling technology [4].

The first step of the proposed methodology classifies customers according to demand and response. The bibliography presents methodologies to perform this task: Self-Organizing Maps (SOM), Fuzzy Techniques or Statistical Tools [6].

The second step, and first concern of this paper, is to develop a process oriented to filter high energy-price days and subsequently to find "price patterns" in Day-Ahead (DA) and Real-Time (RT) energy markets. Here Non-Parametric Estimation (NPE) is used to find patterns in Real-Time energy markets and forecast high-price events in the short-term (< 24h).

There are many papers concerning the forecast of DA price series including no spikes, but it requires data information from nearly 45 previous days in order to forecast one or two weeks of daily prices (see [7]).

We propose an alternative for forecasting energy price series based on what happened in the markets the previous day. The procedure of clustering proposed in this paper can be viewed as follows:

- We consider two-consecutive-day price series, that is, a series of size 48 corresponding to the price of the 48 hours of the two consecutive days. In the first stage, we classify and extract patterns of the two consecutive days from some annual RT Locational Marginal Price (LMP) database, corresponding to the previous year.
- In the second stage, we identify each daily-price series (current year) with one of the price patterns obtained in

the first stage (previous year): at the end of a day, the daily price series (size 24) is identified with one of the price patterns comparing the first 24 hours. Then the price series of the following day (day-ahead forecasting) is estimated by hours 25 to 48 of the price pattern selected.

Obviously, the objective of the method is not to provide an accurate estimation for each day, but helping customers to take decisions.

Finally, through the knowledge of these patterns and price forecasts, the behavior of a medium customer is simulated to evaluate the benefits of each response [2]. With these tools, demand-side players can consider the hours with higher prices, the length of price valleys to accomplish energy recovery, and the possibility to face multi-day critical periods with a minimum energy cost if demand is not appropriately rescheduled.

MAIN CHARACTERISTICS OF DEMAND RESPONSE INITIATIVES

Nowadays, a lot of Independent System Operators (ISO) has developed DR Programs with the aim of changing the demand of large power users. Indeed, small customers have been exposed to prices (CPP tariffs, California, USA [8] or "Tempo-tariff" in France). Besides, some ISO encourage the possibility of demand aggregation through commercializer entities [9] to reach the minimum demand level stated for the participation of users.

The present problem is how (remember that we have selected the customer segments which are more suitable to participate in DR policies [6]) to detect the most effective loads in order to achieve a minimum change in demand level to comply with energy reductions in short-term markets with a quick response. The methodology proposed here is geared to find opportunities in Real-Time Price Markets (RTPM) and prevent customer exposure to peak prices, with the help of Semi-Automated or Fully Automated-Demand Response mechanisms (Auto-DR Project 2005/07 in California [4]).

Notice that this response, if any elasticity is possible [1], can be also obtained through on-site small generation or changing the technology of the loads (for instance, by the use of dual-fuel-gas-electricity loads).

Customers enrolled in DRI agree to achieve a certain level of load reduction at the discretion of the ISO in a short term notice period (from 30 to 120 minutes) involving demand control, backup or self-generation. In Price Response Programs the peak price periods are announced at the end of the preceding day [8] although, in some occasions, the notification can be as brief as four hours ahead (CPP-Variable in California, USA). The peak period varies in length from 1 to 5 hours.

STATISTICAL CLUSTER METHODOLOGY

In order to participate in short-term programs it is necessary to know customer's response and probable curtailment periods. This information can be used to define and understand when demand should be reduced (energy offers) and when the energy recovery is allowed (energy bids).

To extract price patterns from some annual Real Time LMP series, New England (NE, USA) market prices in 2006 were used (internal hub). The idea is to filter weekdays with the highest price periods, i.e. to find the more suitable days for calling Load or Price Response Initiatives.

Description of Ward´s Linkage

The Ward's Linkage Method based on Euclidean distance is a widely used technique of hierarchical cluster analysis. The method begins with isolated individuals or cases and progressively combines them into clusters until each individual is in the same cluster. The linkage function specifying the distance between two clusters is computed as the increase in the "error sum of squares" (SSWG) after fusing two clusters into a single one. Ward's Method seeks to choose the successive clustering steps minimizing the increase in SSWG on each step.

After the clustering process, in next section each individual cluster will be analyzed further putting each price-pattern through a non-parametric estimation.

The final procedure classifies the Real Time LMP series from the New England market database with regard to their price patterns.

Determination of the Number of Clusters

As every hierarchical cluster method, the Ward's Linkage provides the classification into 1, 2,…, "n" clusters. Therefore, the optimal number of clusters to consider should be determined previously. In this context, the Dunn Index [10] and the RMSSTD (Root-mean-square standard deviation) [11] can be used to establish which value of n_c (number of clusters) is the most parsimonious.

- **Dunn Index**: This index attempts to identify compact and well separated clusters. The index is defined by:

$$D(n_c) = \min_{i=1,\dots,n_c} \left\{ \min_{j=i+1,\dots,n_c} \left(\frac{d(C_i, C_j)}{\max_{k=1,\dots,n_c}(diam(C_k))} \right) \right\}$$

(1) Where $d(C_i , C_j)$ is the dissimilarity function between two clusters **C** and **C** defined as:

$$d(C_i, C_j) = \min_{x \in C_i, y \in C_j} d(x, y)$$

and *diam(C)* is the diameter of a cluster, which may be considered as a measure of dispersion of the clusters and it is given by:

$$diam(C) = \max_{x, y \in C} d(x, y)$$

Large values of the Dunn index indicate the presence of compact and well-separated clusters.

- **RMSSTD Index** (Root-mean-square standard deviation index):

This index is used to determine the number of clusters inherent to a data set. It measures the homogeneity of the resulting clusters. Thus, the index value must be as low as possible to obtain a good clustering.

$$RMSSTD_k = \left[\frac{W_k}{\upsilon(N_K - 1)}\right]^{1/2} \quad k=1,\ldots,n_c$$

Where **n** is the number of clusters, υ is the number of variables (data dimensions), **N** is the number of observation in the k^{th} cluster (C_k) and $W_k = \sum_{i \in C_k} \|x_i - \bar{x}\|^2$ is the sum of squares on cluster k^{th}.

The results obtained using the Dunn and RMSSTD Indexes are given in table I:

TABLE I
DUNN AND RMSSTD INDEXES

	n_c=2	n_c=3	n_c=4	n_c=5	n_c=6
Dunn	0.103	**0.124**	0.124	0.124	0.124
RMSSTD	16.5	**10.9**	13.5	12.2	17.8

Therefore, according to Dunn´s and RMSSTD criteria, the optimal number of clusters for the classification is n_c= 3. Figure 1 shows the dendogram obtained using the Ward's clustering method.

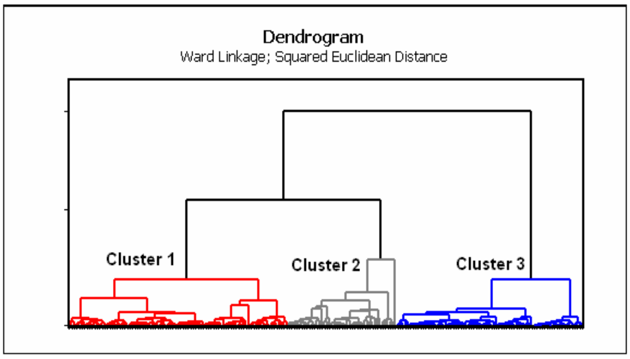

Figure. 1. Dendrogram Hierarchical Clustering.

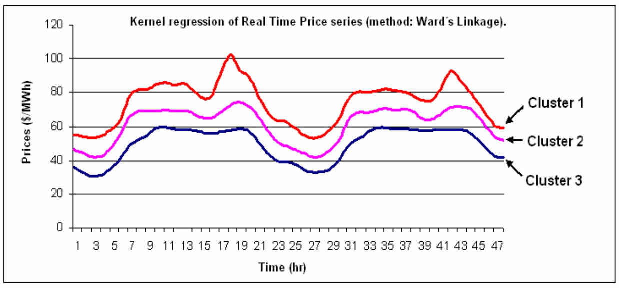

Figure. 2. Average values and Kernel regression of Real Time Price series (Ward´s linkage).

The price-patterns results are shown in figure 2 and the number of price-day in each cluster is done in table II.

TABLE II
WARD LINKAGE CLUSTERING RESULTS

	Cluster size (Ward's)
Cluster 1	43 series
Cluster 2	87 series
Cluster 3	74 series

Once obtained the classification of the 48 hours (two consecutive days of LMP series) in three clusters, in next section we attain the associated pattern for each cluster by means of kernel regression method (see [12]).

ESTIMATION OF PATTERNS BY MEANS OF KERNEL REGRESSION

Non-Parametric curve estimation (NPE) methods provide a powerful statistical tool for exploring the underlying structure in a data set. The parametric option has been rejected because the evolution of the series does not present an easy form to fit the parametric models, that is to say, a way that cannot raise the adjustment with a simple model throughout the considered period.

Taking into account that the series have been evaluated at the same time pace (hourly data), a fixed regression model can be fitted by considering an equally spaced design on [0, 48], that is:

$$Y_{l,t} = m_l(X_t) + \varepsilon_{l,t} \quad t=1,2,\ldots,48, \quad l=1,2,\ldots,204$$

Where m_l is the trend component of series l and $\varepsilon_{l,t}$ denotes the random component.

Kernel regression was derived independently by Nadaraya and Watson [12]. The mathematical basis was provided by Parzen's earlier work on kernel density estimation [13], based on the idea that the conditional expectation $m_l(x_t)$ $=E(Y_t/X=x_t)$ can be obtained by means of :

$$m_l(x_t) = \frac{\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x_t - x_i}{h}\right) Y_{l,t}}{\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x_t - x_i}{h}\right)}$$

This procedure implies the use of a function *K(x)* to assign weights to near observations. The function *K(x)* is the kernel one, which is traditionally chosen from a wide variety of symmetric density functions [13] and the parameter *h* is called the bandwidth or smoothing parameter.

After the classification of the price series into three clusters, nonparametric kernel regression based on the Gaussian density

$$K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right)$$

was applied in order to obtain the "price pattern" of each

group of series.
On the other hand, the smoothing parameter selection is carried out by means of a data-based method for local linear regression given in Ruppert-Sheather-Wand work [14]. The results obtained are shown in Table III.

TABLE III
RANGE AND OPTIMAL SMOOTHING PARAMETER

	Cluster size (Ward´s)	Min. value estimated	Max. value estimated	Opt. smoothing parameter (h)
Cluster 1	43 series	52.94	102.42	0.5385
Cluster 2	87 series	41.58	74.05	0.5230
Cluster 3	74 series	30.23	59.72	0.6215

The kernel estimation values (solid line) and confidence bands (dashed lines) for price series clusters 1, 2 and 3, obtained through Ward's method are shown in figure3. Notice that the confidence bands are narrow and it can be stated that a characteristic pattern can be determined for each of the three clusters.

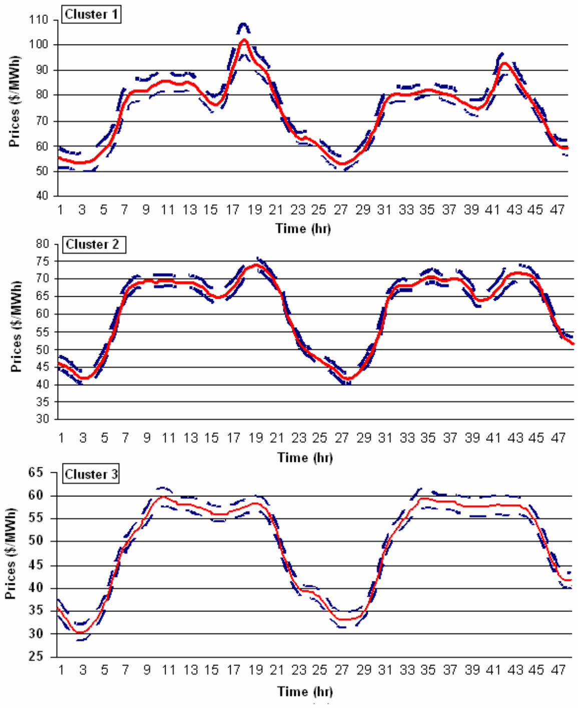

Figure. 3. Kernel estimation values and confidence bands for Real Time prices (respectively clusters 1, 2 and 3).

Moreover, we can state that cluster 1 is related to winter periods while cluster 3 refers to summer periods (see Table IV).

TABLE IV.
SEASON DISTRIBUTION ON CLUSTERS.

	Spring	Summer	Autumn	Winter
Cluster 1	7 (16.3%)	4 (9.3%)	10 (23.2%)	**22 (51.2%)**
Cluster 2	27 (31.1%)	13 (14.9%)	26 (29.9%)	21 (24.1%)
Cluster 3	23 (31.1%)	**30 (40.5%)**	15 (20.3%)	6 (8.1%)

However, it has been proven that all week-days are presented in the three clusters homogeneously (see Table V).

TABLE V
WEEK-DAYS DISTRIBUTION ON CLUSTERS.

	Tuesday	Wednesday	Thursday	Friday
Cluster 1	12 (27.9%)	10 (23.2%)	9 (20.9%)	12 (27.9%)
Cluster 2	20 (23.0%)	23 (26.4%)	26 (29.9%)	18 (20.7%)
Cluster 3	19 (25.7%)	18 (24.3%)	16 (21.6%)	21 (28.4%)

APPLICATION OF PRICE PATTERNS IN CUSTOMER PRICE RESPONSE INITIATIVES

Load and Price Response programs

The results of some international surveys on DR programs from Industrial and Commercial Loads [15] show that Load Response programs are more common than Price-Response ones. This lack of balance is even more significant in the European Union. Due to this fact, the application of price pattern forecasting will be applied in price response, but obviously load-response is another alternative for the customer where patterns can play an important role. For example, pilot projects such as "Auto-DR" in California [4] report the increase in participant monthly utility bills as a major concern for some of the sites of this project. A predictive tool that outlines the number of peak price (Critical Peak Price events) could help eliminate these concerns.
In New England, up to 21% of the MWs [16] (43% of customers) joins Price-Response programs. HVAC, refrigeration and lighting loads accounts for the 80% of price-response customers. The load reduction methods are the use of Energy Management Systems (EMS, 50%) or manual increase (decrease) in thermostat temperatures (1-2 °F). It is also clear that enabling technology (EMS, the development of specific internet servers,..) increases the response [8]. The evaluation of this response and the design of specific policies to be implemented in Auto-DR systems need "methods to estimate peak demand savings for short duration …new dynamics models are needed based on knowledge of weather data, peak load shapes and HVAC systems and controls…"[4]. These models are the so called Physically-Based Load Modelling methodology, and a specific software was developed and has been applied by authors in this section [5].

Requirements for RTP Price Response in ISO-NE.

These programs require in ISO-NE individual or group response with a minimum reduction of 100kW. The ISO notifies customers if wholesale prices exceed $100/MWh either the night before or in the morning (Load-Response must be accomplished in a shorter term: 30min to 2h). In aggregated loads, the forecasting of spike prices gives the aggregator more time to react to time and price response. Two important concerns are energy recovery after the load control period and the presence of multiple day-events.

From an economic point of view, the ISO pays the customer the greater of Real-Time Wholesale Price or Guaranteed Minimum of $100/MWh. The price response windows open as early as 7am and remain open until 6pm, but in practice [16] and depending on NE load zone price response, events start at 2pm until 6pm from October to April, and from 12am to 6pm during the rest of the year. The metering requirements are low (hourly meter), but it should be noted here that the cost of an automated system (EMS, internet communications, …) is expensive (some authors estimated quantities from $3000 to $5000 for each site, [4]).

Customer profiles and end uses

Two different customers have been selected for simulation purposes: an Office Building (medium customer) and a University Campus (large customer). All the customers correspond to real customers in Europe. Figure 4 shows the winter load (notice 51% of peak price patterns appear in winter days, see table IV) and main end-uses. Office have natural gas heating, whereas the University uses electricity for HVAC.

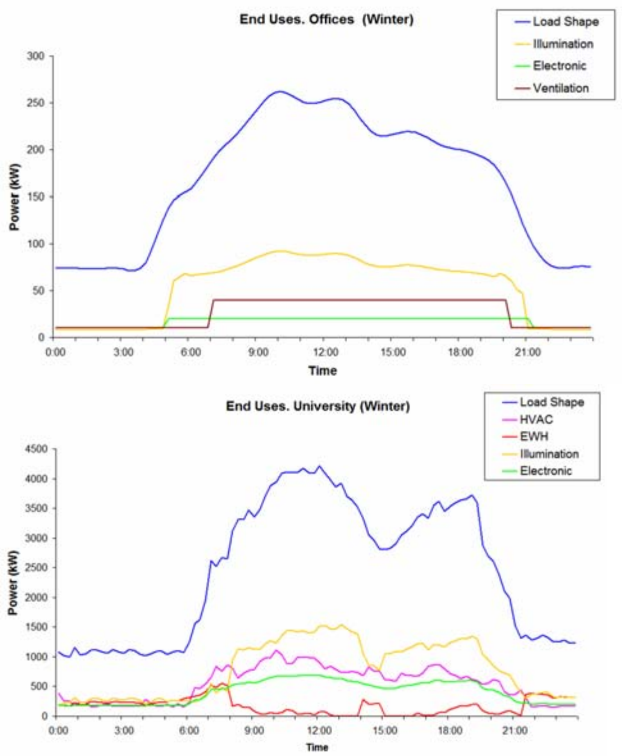

Figure. 4. Daily Load Curves (winter) for the customers.

The load size and end-uses give different possibilities for each customer. For instance, the Office customer can not participate directly in Price or Load Response programs because they need an aggregator. The aggregator can use lighting and ventilation loads, but to a limited extent (the user can reduce air exchange or can profit from daylight but these actions do not switch-off the 100% of load but only a partial fraction: 30-60%). On the contrary, the University can reduce large amounts of load due to the use of electrical heating, ventilation and lighting, and perhaps because of emergency generators. Ventilation and lighting loads have another characteristic: they do not produce energy recovery, but HVAC loads need some energy recovery to reach comfort conditions after a control period to prevent complaints from students or university staff.

Simulation results

The results of some price-response policies in Office Building are shown in figures 5. Due to the higher prices from 14 to 18h and to the orientation of the building (south), the aggregator should consider sending a signal to reduce both end uses (ventilation and lighting) at the same time. The simulation results obtained through the specific software described in [5] (see figure 5), shows that the customer responds with an average reduction of 45kW. Obviously, our "aggregator" needs to find additional customers (medium or small customers) to fill the minimum level of response (100kW).

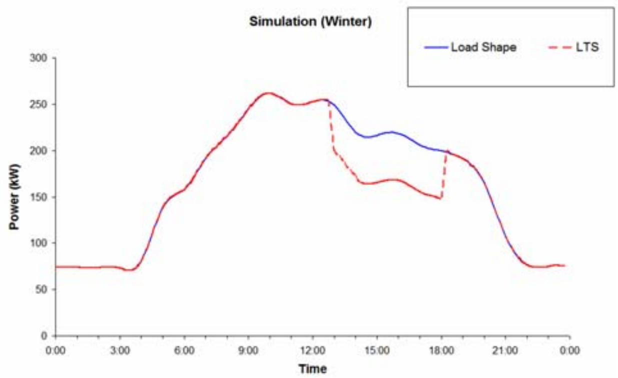

Figure 5. Office load (winter),before and after the control.

In the case of the university, the objectives of the simulation are different. We assume that the engineering staff of the university knows through the methodology explained in previous sections, the possibility of having a peak-day in the short term. In this case, they evaluate the possibility of notification of a response event from the ISO. Thus, the customer simulates a higher reduction in the morning (heating up to 300kW + ventilation 150kW) and a later price-response from 14h to 18h. To face both events, HVAC load needs a recovery period after the response event (the event is called from 30min to 2h in advance and usually lasts for 2-3 hours). A shorter notice of the event with an adequate response will allow them a substantial increase in the revenues from \$0.35/kWh to \$0.5/kWh. Due to this recovery period (from 12 to 15h), the price-response is lower but the economic revenues are considerably higher (from \$0.1/kW to \$0.5 /kWh). Results are shown in figure 6.

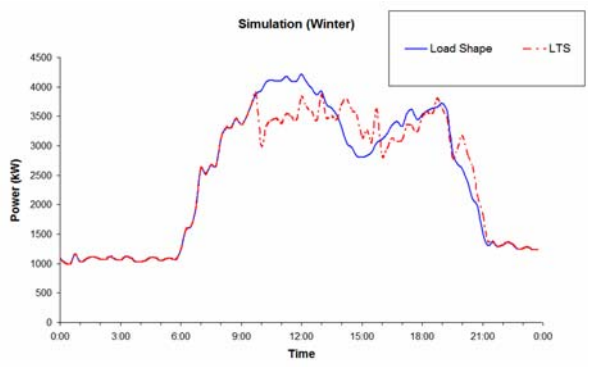

Figure 6. University customer (winter season). Load and Price responses.

Economic evaluation of price-response initiatives

A simple evaluation is performed for the two customers. Through the use of the price pattern (cluster 1) and load simulation, we obtain: average price values of Real-Time LMP in NE market (for simplicity, capacity and other costs components are evaluated as 12% of the wholesale market price), the number or price events, and the load and power reduction/recovery. In the case of load response events we have considered 8 events a year (a conservative number). The electricity tariff is supposed \$0.6/kWh [16]. Results are shown in table VI.

TABLE VI
Real Time Price-Response for the Office and University Customers

Item	Office	University
Demand reduction range (kW)	47.5	210-350
Energy reduction per event (kWh)	182	651
Event days (winter)	22	22
Hours per event	4	4
Guaranteed minimum energy payment (\$/kWh)	0.100	0.100
Real-Time Price during event (\$/kWh, average)	0.106	0.106
Electricity Supply price (\$/kWh)	0.06	0.06
Demand response payment (winter season)	\$443	\$1518
Additional energy purchased due to recovery	0.0	\$780
Avoided Electricity purchase (winter season)	\$250.1	\$859
TOTAL BENEFIT (winter season)	**\$693.1**	**\$1597**

Notice that power demanded by the university fluctuates during the control period and so the space heating load needs some energy recovery after each control period (see figure 6 in the noon and in the evening).

CONCLUSIONS

Customers can obtain interesting benefits from the energy trading with the help of an aggregator acting as an enrolling participant in ISO Response Programs. Unfortunately, it is difficult to directly participate in the electricity market due to its complexity. Through the proposed Statistical Ward's Linkage and Non Parametric Estimation tool, the customer or a third party agent (aggregator) can identify, simulate and derive a plan to maximize its own benefit using the electricity purchased in RTPM. Each price-pattern suggests different demand strategies for achieving the improvement of system operation, while customers manage and reduce their energy costs. So, price-response initiatives are necessary in real markets, and these programs should be promoted by ISO, Load Serving Entities, utilities and Energy regulators, because the reduction of supply-side market power benefits all the customers, not only those involved in response policies. Moreover, the possibility of Load Response initiatives would multiply the benefit presented in table VI for the customer and besides to all the society due to price elasticity.
Finally, the methodology presented in the paper provides a new way through the effective integration of Demand and Supply resources for a better market operation

REFERENCES

[1] E. Hirst, "Reliability benefits of price-responsive demand", IEEE Power Eng. Review, vol 22, n 11, 16-21, 2002.
[2] US Department of Energy, "Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them (A Report to the United States Congress pursuant to Section 1252 of The Energy Policy Act of 2005)", February 2006.
[3] Neenan Associates, Lawrence Berkeley Nat. Lab, "How and Why Customers Respond to Electricity Price Variability", LBNL Report 52209, January 2003
[4] Piette, M.A., Watson D., Motegi N., Kiliccote S. (Lawrence Berkeley National Laboratory). "Automated Critical Peak Pricing Field Tests: 2006 Pilot Program Description and Results". California Energy Commission, PIER Energy System Integration Program. CEC-500-03-026
[5] C. Alvarez, A. Gabaldón, "Assessment and Simulation of The Responsive Demand Potential in End-User Facilities: Application to a University Customer" *IEEE Trans. On PWRS*, vol. 19, n 2, 1223-123, 2004
[6] S. Valero, M. Ortiz, Fco. García. A. Gabaldon, "Classification, Filtering and Identification of Electrical Customer Load Patterns Through the Use of SOM Maps", *IEEE Trans. On PWRS*, vol.21, n 4, 1672-1682, 2006.
[7] Conejo, A., Conteras, J., Espínola, R. y Plazas, M, Forecasting electricity prices for a day-ahead pool-

based electric energy market, *International Journal of Forecasting*, vol 21, pp 435-462, 2005.
[8] Faruqui, A., George, S.S. "California´s Statewide Pricing Pilot. Overview of Key Findings", MADRI Advanced Metering Infrastructure Workshop, May 2005.
[9] ISO New England Load Response Program Manual [Online].Available:http://www.isone.com/rules_proceds/isone_mnls/index.html June 2005.
[10] J. C. Dunn, "Well separated clusters and optimal fuzzy partitions," *Journal of Cybernetics*, vol. 4, pp. 95–104, 1974.
[11] Subhash Sharma, *Applied Multivariate Techniques*, John Willxy & Sons, 1996.
[12] E.A. Nadaraya. *Nonparametric Estimation of Probability Densities and Regression* Curves. Kluwer Academic Publishers, Dordretch, 1989.
[13] Wand, M. P. and Jones, M. C. *Kernel Smoothing*. Chapman and Hall, London, 1995.
[14] Ruppert, D., Sheather, S. J. and Wand, M. P., "An effective bandwidth selector for local least squares regression". *Journal of the American Statistical Association*, 90, 1257–1270, 1995.
[15] Rocky Mountain Institute, "Demand Response: An Introduction. Overview of programs, technologies and lessons learned", for Southwest Energy Efficiency Project, April 2006. Available on line at http://www.rmi.gov.
[16] Reports of New England´s wholesale electricity markets (USA). Available on line: https://www.iso-ne.com/markets/mkt_anlys_rpts/.

Acknowledgments

This work was supported by the Spanish Government (Ministerio de Educación y Ciencia) under Reasearch Project ENE2007-67771-C02-01&02/CON and Bancaja-UMH 2007 IPOG06 Grants