Structural Breaks, Price and Income Elasticity, and Forecast of the Monthly Italian Electricity Demand

Dicembrino, Claudio and Trovato, Giovanni

Department of Strategic Planning, Enel Spa, Rome (Italy), Department of Economics and Finance, University of Rome Tor Vergata", Rome (Italy)

14 June 2013

Structural Breaks, Price and Income Elasticity, and Forecast of the Monthly Italian Electricity Demand

Claudio Dicembrino[*], Giovanni Trovato[†]

June 14, 2013

DRAFT VERSION DO NOT QUOTE WITHOUT PERMISSION

Abstract

Insights about electricity demand dynamics is fundamental for investment capacity, optimal energy policies, and a balanced electricity system. This paper presents an empirical analysis of the monthly Italian electricity demand since January 2001 to June 2012. In the first section we conduct the analysis of structural breaks in the electricity demand finding that the series has two structural breaks in August 2002 and August 2004 as market liberalization effects on consumption. In the second part of the paper we estimate demand price elasticities both for residential and industrial sector. As expected from the electricity economics literature concerning elasticities estimates, we find that the long run price and income elasticities are more price elastic than the short run both in industrial and residential consumption. In the third and last section, we compare two different forecasting models: the Hidden Markov Models (HMM) and the Holt Winters (H-W) seasonal smoothing method. Considering the Mean Absolute Percentage Error (MAPE), the HMM approach seems to show a superiority in forecasting the monthly electricity demand compared to the H-W methodology.

Keywords: *Electricity Demand, Price and Income Elasticity, Hidden Markov Model, Holt-Winters Seasonal Filter Smoothing.*

Jel Numbers: *C53; Q41; Q47; R2.*

[*]Department of Strategic Planning, Enel Spa, Rome (Italy).

[†]Department of Economics and Finance, University of Rome "Tor Vergata", Rome (Italy).

Introduction

The theoretical and empirical motivation of this analysis stems from the unpredictable and unstable electricity demand path in Italy. Studying the electricity consumption dynamics represents a key asset to drive decisions on capital-intensive investments both for government agendas and companies business strategies (Hamzacebi, 2007)[24]. In this context a reliable forecasting model might represent a crucial asset in programming the necessary actions concerning supply security, environmental quality and other important aspects of energy policy. This paper offers an exhaustive understanding of the Italian electricity demand specification analzying structural breaks, income and price elasticity and its monthly forecast. It is composed by three sections. The first one will investigate the trend, seasonality factors and structural breaks. The second section analyzes price and income demand elasticity comparing the results obtained with the related academic literature. Finally, last section compares two original forecasting methods that are not usually implemented by energy operators *Holt Winters Seasonal Filter Smoothing* (H-W) and *Hidden Markov Model with finite mixture* (HMM) shedding some lights on the most appropriate forecasting mechanism of electricity demand.
The first task to accomplish is to test if the electricity demand series shows difference in time for mean and variances, secondly if it has been subjected to some fundamental changes (breaks[1]) and, lastly, if these breaks have permanent effects on series dynamic or, alternatively, the effects are just destined to "vanish" in the short time. Moreover once investigated on breaks existence we want to test if it is possible to derive a sorted timing: from the first to the last change. Beside structural breaks just mentioned, calendar effects and weather temperatures represent two more important elements affecting the electricity demand. Engle et al. (1986)[13], Filippini (1995) [17], Henley and Peirson (1997[25], 1998[26]) Considine (2000)[11], Johnsen (2001)[28], Valor et al. (2001)[48], Pardo et al. (2002)[38] among the others, provide important contributions on the impact of seasonal weather variations on the electricity consumption fluctuations. Nevertheless Sailor and Munoz (1997)[43] or Yan (1998)[51] have used several meteorological factors such as humidity, wind speed, cloudiness, rainfall and solar radiation for taking into account all the climate variables related to the electricity demand.

The second section of the paper aims to measure the consumer demand reactivity to price and income changes. An accurate perception of price elasticities to electricity consumption is of crucial importance both for planning

[1]A structural break can be defined as an unexpected shift in a time series. It has a strong relevance in the economic theory since a structural break may change past trends or theories regarding the issue investigated.

the expected electricity demand and, more important, for policy makers decisions on the appropriate capacity for future electricity consumption levels. Although it emerges a diversified and somewhat fragmented estimate of income and price elasticity depending on data, geography and sector, generally speaking there is a shared literature consensus on a low elasticity in the short run and a more elastic demand to price and income changes in the long run. Taylor (1975)[46], has completed one of the first review of the electricity demand literature. He estimated a short run price elasticity between -0.9 and -0.13. Conversely, long-run price elasticities ranged from -2 to near 0. For commercial sector, he found a short-run price elasticity of -0.17 and a long run elasticity of -1.3. Bohi and Zimmerman (1984)[6] found a short run residential electricity price of -0.2, and -0.7 in the long run[2]. Garcia-Cerruti (2000)[20] estimated that the price elasticities for residential consumers in California was in mean -0.17, with a minimum of -0.79 and a maximum of 0.01. Bernstein and Griffin (2005)[4] have found a quite inelastic relationship between electricity demand and price, noticing that it has not changed significantly between 1977 and 2004 for residential sector and from 1977 to 1999 for commercial sector in 48 US States.

Further, Labandiera at al. 2012[30] use a random effects model for panel data in Spain finding that the electricity price elasticity in urban sector is -0.11, and in rural sector is -0.2; while the income elasticity is -0.29. Blazquez et al. (2012)[5] indicate that the short-run price elasticity is approximately -0.11 in the short run and -0.24 in the long run, while the income elasticity is 0.14 and 0.30 respectively for the short and long run. As claimed in Espey et al. (2004)[15] price elasticities reported in the literature range from 0.07 to -2.01 for the short run, and from -0.07 to -2.5 for the long run. The low levels of elasticities are explained by the scarce degree of substitutability and from the fact that in front of an electricity price increases, the consumer may react, but not instantaneously, buying a more efficient new appliance characterized by a less-expensive energy use. Further Bernstein and Griffin (2005)[4] and Blazquez et al. (2012)[5] agree on the inelastic demand in the short term and a more elastic demand for the long run, although they say that in areas where the costs of substitutes are competitive, price elasticities may be larger. Moreover, Reiss (2005)[42] states that beside the different econometric techniques in finding the elasticities estimates, the nonlinear structure of tariff schedules and aggregation of the single user consumption add complexity to the relation between marginal prices and consumption. Concluding we can say that, although there is a shared consensus on the inelastic demand in the short run and more elastic

[2]Bohi and Zimmerman (1984)[6] investigated on the responsiveness of the elasticity demand during the oil prices shock of 1974 and 1979 finding that the estimated price elasticities did not differ substantially before and after the abrupt price changes happened during the Seventies

demand in the long run, the estimates reported by the most relevant literature on this issue, vary in function of the geographic area and the data analyzed[3].

The third and last goal of our analysis concerns the forecast of the monthly italian electricity demand, comparing the seasonal H-W filter and the HMM forecasting technique. The H-W approach originally presented in Holt (1959)[27] and Winters (1960)[50] is an univariate method, already implemented by academics as such as Taylor (2003)[45] and professionals for simple forecasting approaches without necessarily fitting a proper econometric model. Cipra and Romera (1997)[9] have developed a more robust version, respect to the one of the first attempts proposed by Cipra (1992)[8], implementing the Kalman filter for H-W robust forecasts. His main advantage relies on the fact that the forecasts are automatically updated by each new incoming observation. This approach is particularly suited to forecast economic and financial time series containing changing trends and seasonal correlation. However, generally speaking the principal limitation of the H-W filter smoothing implemented to obtain robust forecasts is strictly related to the presence of outliers in the sample which can result in biased smoothing values and not reliable predictions for the short, medium and long run (Croux et al. (2008)[21]). The second forecasting methodology used in this paper is the HMM. These models are widely used in many applications as economics (forecasts of financial time series or portfolio strategy management), psychology (learning process or social interactions), biology (DNA sequences) or in speech recognition (among the others Chomsky, (1963)[34], Wickens, (1982)[49], Langeheine and Van de Pol, (1990)[31] Rabiner, (1989)[41], Schmittmann, Visser, and Raijmakers, (2006)[44], Kim (1994)[33] and Ghysels (1994)[22], Miller and Rainer (2000)[35], Fruhwirth-Schnatter, 2006[19]). Strictly speaking Hidden Markov Models are based on the idea that the data generation process has been affected by two main facts: (i) a state of the world, (ii) the transition between states over time. In section (5.2) we will examine in detail how the methodology works.

1 The Data and Summary Statistics

In order to analyze the electricity demand in Italy we consider a series of monthly variables within the period from January 2000 to June 2012 (Figure 1). For the arc of time analyzed, the monthly values of electricity

[3]Further, as underlined in Lee, Yi-Bin (2011)[32] *"An increase in electricity price has a negative or no influence on electricity consumption (...) the estimated elasticities of time dynamic indicate that electricity demand is income inelastic, price inelastic and temperature inelastic".*

consumption "*log de*" were obtained by Terna that provides the official Italian statistics about electricity consumption[4]. The industrial production data come from the Istat database (Figure 1)[5]. The industrial "$log\ price_{ind}$" and residential "$log\ price_{res}$" electricity prices were extracted from Enerdata[6] (Figure 2). For our aim we transform the annual series into a monthly series simulating a linear incrementing value from January to December of each year[7]. The GDP per capita income variable comes from the International Monetary Fund database.

The weather temperature (t_j) come from the *weather platform database* of Bloomberg and in particular the average temperature registered in Italy for each month during the time analyzed. So, considering 18C as a threshold for cold and heating degrees, we specify the weather factors in monthly terms as:

$$CDD = \sum_{j=1}^{nd} max(0, t_j - 18)$$

$$HDD = \sum_{j=1}^{nd} max(0, 18 - t_j)$$

We segment temperature variations in terms of heating (HDD) and cooling (CDD) degrees days[8].

All variables are transformed to natural logarithms. To determine the time series of the electricity demand, GDP, income, industrial production, residential and industrial electricity prices, the conventional unit root tests (the Augmented Dickey Fuller (ADF) [12] and the Phillips-Perron (1998) [39] unit root tests) are applied to the natural logs of the series. We choose

[4]Terna is a leading energy transmission grid operator. The Terna Group is responsible in Italy for energy transmission and dispatching throughout the entire territory and therefore for the safe management of energy flows in Italy, guaranteeing the balance between electricity supply and demand.

[5]Istat is the Italian National Institute of Statistics in charge of registering and providing official dates on industrial production.

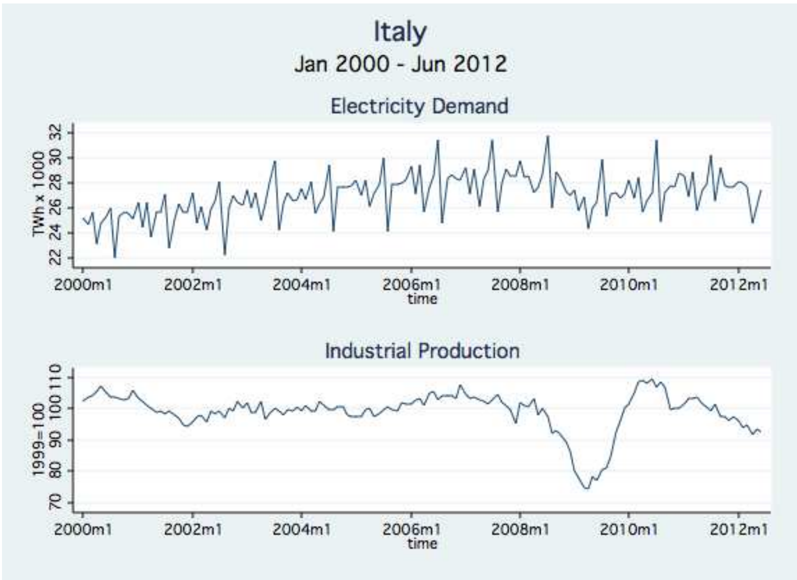
[6]The prices data are taxes excluded and in US$ 2005. We have opted to transform the yearly observations in monthly observations (for the variables that we do not have directly the monthly variations as the objective of our analysis is to analyze the monthly electricity demand).

[7]We compare the data obtained from Eurostat dataset (half year observations) and the IPEX (Italian Electricity Exchange) price observing that monthly price data obtained through our linear transformation from Enerdata show the same incremental trend.

[8]An alternative estimation procedure to consider the seasonal fluctuations is to specify the model considering 15C as the threshold for heating and 22C for cooling degrees. This approach considers the estimated comfort area between $15 - 22$C in which no heating or cooling degrees will take place. However, although we recognize the importance of this approach, we do not consider it suitable for our task.

the Akaike's information criterion (AIC) and the Schwarz's Bayesian information criterion (SBIC) to decide the variable lags. As it can be seen from table (8), for "*log de*" "*log gdp*" "*log* $prod_{ind}$" "*log* $price_{res}$" and "*log* $price_{ind}$" series the null hypothesis of a unit root cannot be rejected even at 10% significance level. Therefore we may conclude that the series are integrated of order one or I(1).

Figure 1: Electricity Demand and Industrial Production

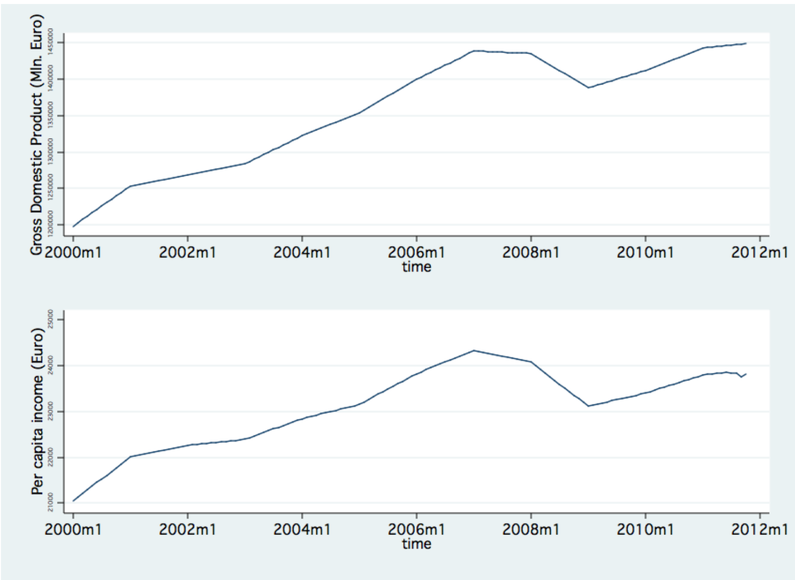


Source: Authors elaboration on Terna and Istat data

Figure 2: Industrial and Residential Electricity prices

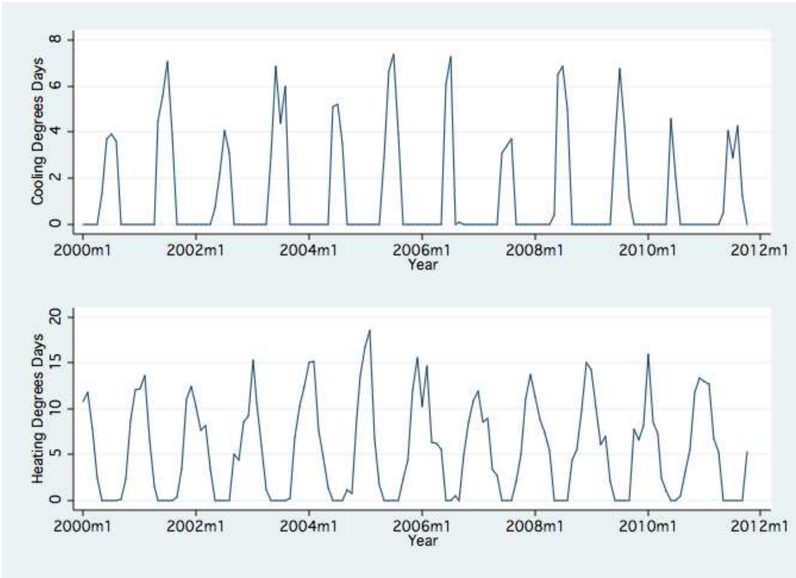
Source: Author's elaborations on Enerdata data (prices taxes excluded, US$ '05)

Figure 3: Gdp and per capita income



Source: Authors elaboration on IMF data

Figure 4: Cooling and Heating degree days



Source: Authors elaboration on Bloomberg weather platform data

2 Trend, Seasonality and Structural Breaks

2.1 Seasonal adjustment for the Electricity Demand

The electricity consumption shows certain behavioral "*components*" that repeats itself any n periods. In this brief section we provide an analysis to decompose the electricity demand of its seasonal and trend components. To filter these components, we proceed with a de-seasonalization technique using a multiplicative approach composed by a multiplicative seasonal factor that increases (decreases) the variable by the same percentage every month. To make the series de-seasonalized requires that a set of seasonal dummies ("$mseas$") be created by defining the elements of the set with a specification including for instance $January = 1$ when all the other months assume 0 value: $February = 1$ when all the other months assume 0 and so on. The regression run to evaluate the importance of seasonal factors is shown in table (7). Taking into account all the above mentioned effects, the model developed to eliminate the seasonality fluctuations is finally given by:

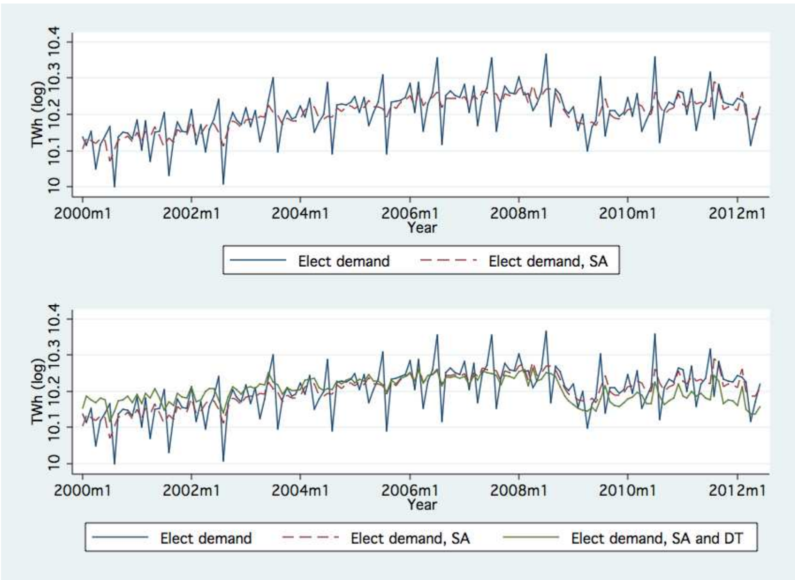
$$log\ Ed_t = \alpha_0 \sum_{t=1}^{12} mseas_t + \alpha_2 CDD + \alpha_3 HDD + e_t : \qquad (1)$$

In order to investigate for "long term" movements (i.e. trend) in electricity demand time series, we estimate the equation (2) adding a trend element (t) in the equation:

$$log\ Ed_t = \alpha_0 \sum_{t=1}^{12} mseas_t + \alpha_2 CDD + \alpha_3 HDD + t + e_t : \tag{2}$$

The result are reported in the table (7) while the graphs below plot the historical data of electricity consumption, its seasonal adjusted path, and the de-trended series.

Figure 5: Real path, seasonal adjusted (SA) and de-trended (DT) evolution of the electricity demand.



Source: Author's elaborations on Terna data

2.2 Checking for Structural Breaks

One of the most thorny issues dealing with economic time series concerns the behavior of the series. A structural break can be defined as a series change caused and reflecting a result of institutional, legislative or technical changes. In some cases, it can also reveal deep economic policies changes or large economic shocks (i.e. oil crisis 1973). Thus the presence of a structural break in the series analyzed would bias the test towards a non-rejection of the null hypothesis explaining the results obtained. In particular the problem arises if the series shows a time changing behavior both in mean and variance, conducting to biased results for tests based on OLS assumptions[9]. For instance, if the electricity demand shows a stationary path, than operators have a limited uncertainty about future values (rational expectations?) with a regular and limited in time fluctuations. It could be controlled by regulators, and any effect of interventions will not be permanent. This feature causes the intrinsic weakness of unit root tests which have an I(1) series as a null hypothesis. Therefore, we subsequently run developed tests for structural change in univariate time series that do not erroneously accept the unit root hypothesis in presence of breakpoints.

Among the most used methodologies for checking structural breaks there is the Zivot-Andrews (1992)[2] test that allows for the presence of a single structural break and then performs a DF test on the series inclusive of the estimated breakpoints. The null hypothesis of an I(1) process without an exogenous structural break is tested against that of a trend-stationary series with a break which occurs at an unknown time. One of the most important weakness of the Zivot-Andrews test is *(i)* its inability to deal with more than one break in the series; *(ii)* the test inability in capturing endogenous breaks. Thus if we deal with time series showing several up and down swings during the arc of time analyzed, the Zivot test might be inappropriate to capture all the breaks impacting the series. For all these reasons we do not consider the Zivot-Andrews test and we take in consideration the Clemente Monténes and Reyes (1998)[10] (hereafter CMR test) test for both one and two breaks. Addressing this problem with the CMR methodology, the test would allow us to check for more than one structural break.

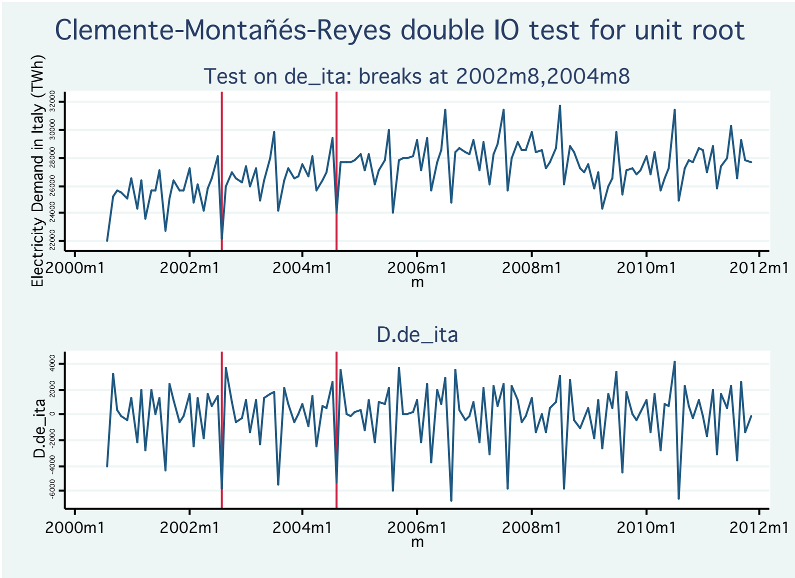
This test has the desirable property of being implemented to search for an unknown break date, which may occur under both the hypotheses of stationarity or nonstationarity. Secondly, if the series actually exhibits a break, CMR test exploits this information to improve the power of the test itself. The tests devise level-shift models, changing-growth models and "mixed" models, allowing for shifts in both the level and slope. Furthermore, their test verifies the existence either of an additive outlier (AO), which captures a sudden change in the series due to a transitory shock or to an anomaly in the data, or alternatively of an innovational outlier (IO), which implies a gradual shift in the series mean time. Not having any reason to restrict ourselves to either level or slope shifts, we implemented both IO and AO model.

Looking at results an IO model seems more appropriate, since persistent

[9]In econometrics we define as unit root variables those have changes (in mean or variance or both) depending on time

shocks which influenced the variables of interest for a longer time period seems more likely in this context. The test conducted points to the existence of two (significant) structural breaks for the series.

Figure 6: Clemente-Montanes-Reyes approach for structural breaks.



Source: Author's elaborations on Terna data. Both graphs are given by Clemente Monténes and Reyes methodology. The upper graph is generated with the electricity demand expressed in levels. The lower graph shows the electricity demand growth rate.

Following the test results, the 2002:8 and 2004:8 breaks can be picked out as years of structural change by the IO model. One can state that the effects of the new regulations, the renewables energies entrance and a more competitive electricity market have started to show between 2002 and 2004 causing these important changes in the electricity demand path. In particular the most relevant changes concerned *(i)* the possibility for the end user to choose the electrical energy provider; *(ii)* the implementation of the Directive 2003/54/CE aiming at increasing the market competition in the EU countries; *(iii)* the construction of several new CCGT power plants for a total of 20.000 MW of installed capacity.
The breaks found through the CMR test will be considered as additional information for the h-step ahead forecast estimate in the HMM methodology (section 5.2).

3 A Partial Adjustment Model for Price and Income Elasticity Estimates

According to the traditional economic theory, electricity demand is supposed to fall when it occurs an increase in the electricity price (holding all the other conditions constant), and conversely it is expected to arise when prices fall. The consumers reactivity to these price oscillations has known as *price elasticity* which can be defined as the ratio between the percentage change in demand and the percentage change in price. On the other side, generally speaking, electricity demand is supposed to arise when it occurs an increase in the disposable income and it is expected to decrease when there is an income reduction. The reactivity to these income variations is called *income elasticity*.

Following the Erdogdu (2007) [14] methodology our aim is to analyze both the short and long run income and price elasticity demand through a *"partial adjustment model"*. Let us specify the electricity demand as:

$$lnEd_t = \alpha + \beta_i lnP_{i,t} + \gamma_i lnY_{i,t} + u_t \tag{3}$$

where Ed_t is the electricity demand, $P_{i,t}$ is the real industrial and household price of electricity, $Y_{i,t}$ is the gdp (which it will be used in analyzing the industrial sector) and income level (which will be used in analyzing the residential sector), u_t is the error term, the subscripts t represents time, α is the intercept and β_i and γ_i represent the coefficients of price and income elasticities of demand. Since the static formulation proposed in (4) does not take into account the dynamic and it does not make a distinction between short and long run elasticities, we use the following formulation to measure the long run elasticities. The fundamental assumption in this model is that consumption levels do not adjust immediately to price and income changes, but gradually converges toward a long run equilibrium relationship. If we suppose that Ed'_t is the equilibrium electricity level given by the following expression:

$$lnEd'_t = \alpha + \beta_i lnP_{i,t} + \gamma_i lnY_{i,t} + u_t \tag{4}$$

and the adjustment to the equilibrium demand level from income and price variations can be expressed by the parameter λ with ($\lambda > 0$) as follow:

$$lnEd_t - lnEd_{t-n} = \lambda(lnEd'_t - lnEd_{t-n}) \tag{5}$$

Substituting the (5) into the (6) after a few algebraic substitutions and solving for $lnEd_t$, we achieve the following expression:

$$lnEd_t = \lambda\alpha + \lambda\beta_i lnP_{i,t} + \lambda\gamma_i lnY_{i,t} + (1-\lambda)lnEd_{t-n} + \lambda u_t \qquad (6)$$

To make the notation in (7) more comprehensive let us simplify it as:

$$lnEd_t = \zeta_i + \phi_i lnP_{i,t} + \varphi_i lnY_{i,t} + \psi_i lnEd_{t-n} + \epsilon_t \qquad (7)$$

where $\zeta_i = \lambda\alpha$; $\phi_i = \lambda\beta_i$; $\varphi_i = \lambda\gamma_i$; $\psi_i = (1-\lambda)$; $\epsilon_t = \lambda\mu$

Since we have aggregated with $P_{i,t}$ the industrial and residential price elasticity and with $Y_{i,t}$ the gdp level and income (i.e. gdp per capita) elasticities, let us disaggregate the equation (7) into two equations sectoral (industrial and residential) equations:

$$lnEd_t = \zeta_i + \phi_1 lnP_{res,t} + \varphi_1 lnY_{inc,t} + \psi_1 lnEd_{t-n} + u_t \qquad (8)$$

$$lnEd_t = \zeta_i + \phi_2 lnP_{ind,t} + \varphi_2 lnY_{gdp,t} + \psi_2 lnEd_{t-n} + u_t \qquad (9)$$

For a better understanding of the long and short price and income elasticities in (7), (8) and (9) see table (1) and (2) and the following results presented in table (3) and (4):

Table 1: Residential sector coefficients entering into the analysis

Residential sector	Long Run	Short Run
Income Elasticity	γ_1	φ_1
Price Elasticity	β_1	ϕ_1

Table 2: Industrial sector coefficients entering into the analysis

Industrial sector	Long Run	Short Run
Income Elasticity	γ_2	φ_2
Price Elasticity	β_2	ϕ_2

The price and income coefficients have corrected signs according with the economic theory stating that there is an inverse relationship between demand and price, and a positive relation between demand and income. In (7), (8) and (9) all the coefficients are significant rejecting at 2% significance level the null hypothesis that one of the coefficients is zero. The Durbin-Watson test is 2.16 in the equation (8) and 2.2 in the equation (9). The existence of serial correlation in the residuals has been resolved at the beginning since we run our regression through the Prais-Winsten (1954) [40] estimation method[10].

So, considering that: $\phi_1 = \lambda\beta_1$= -0.013; $\phi_2 = \lambda\beta_2$= -0.018; $\varphi_1 = \lambda\gamma_1 =$ 0.041; $\varphi_2 = \lambda\gamma_2$=0.06;

and that $\psi_1 = (1 - \lambda_1) = 0.9$; and $\psi_2 = (1 - \lambda_2) = 0.9$, it is immediate to see that $\lambda = 0.1$;

therefore, $\beta_1 = -0.014$, $\beta_2 = -0.018$, and $\gamma_1 = 0.42$, $\gamma_2 = 0.62$.

Table 3: Industrial Sector (equation 8) - Price and Income Elasticity in the Short and Long Run

Industrial sector	**Long Run**	**Short Run**
Income Elasticity	$\gamma_2 = 0.62$	$\varphi_2 = 0.06$
Price Elasticity	$\beta_2 = -0.018$	$\phi_2 = -0.018$

Source: Authors estimates of a change in the independent variables (gdp and industrial price) over the dependent variable (electricity demand).

Table 4: Residential sector (equation 9) - Price and Income Elasticity in the Short and Long Run

Residential sector	**Long Run**	**Short Run**
Income Elasticity	$\gamma_1 = 0.42$	$\varphi_1 = 0.041$
Price Elasticity	$\beta_1 = -0.014$	$\phi_1 = -0.013$

Source: Authors estimates of a change in the independent variables (income and residential price) over the dependent variable (electricity demand).

From the results of income and price elasticities, also reported in table (3) and (4), it emerges that the long run demand is more elastic compared to the short run demand both in the residential and industrial price analysis. Furthermore, income variations have a stronger effect on the demand level compared to price oscillations. In detail for the residential sector we can

[10] Prais-Winsten use the generalized least-squares method to estimate the parameters in a linear regression model in which the errors are serially correlated

state that a 100% increase in real income determines a 42% increase in the long run and only a 4% in the short run. A 100% increase in real price has a negative effect on the electricity demand making the demand decreasing by a 1,4% in the long run and 1.3% in the short run. Our findings are perfectly aligned with the economic theory since we have a negative sign for the industrial price variable and a positive one for the gdp variable. The magnitude is a bit higher respect to the residential sector. In the industrial sector we have that a 100% increase in real income produces a 62% increase in the long run and only a 6% in the short run. A 100% increase in real price has a negative effect on the electricity demand making the demand decreasing by a 1,8% in the long and short run. It basically means that the long term demand is more elastic compared to the short term demand. These findings are consistent with the related literature on the electricity demand previously mentioned in the introduction.

4 Forecasting the Electricity Demand: H-W vs HMM

4.1 The Holt Winters methodology

Following the approach presented in Croux et al. (2008) [21] and Gelper et al.(2008) [21] let's model the electricity demand process Ed_t with $0 < t < T$.

The classical approach to the smoothed series is expressed as the solution of the following optimization problem:

$$where \quad \tilde{Ed}_t = \underset{\theta}{\mathrm{argmin}}(\sum_{i=1}^{t}(1-\xi))^{t-1}(Ed_i - \theta)^2, \tag{10}$$

Considering that the electricity demand path shows a trend, let's add to the equation (11) a local level a_t and a local linear trend F_t. The local level and trend are the solution of the following optimization problem:

$$\tilde{Ed}_t = \underset{\theta}{\mathrm{argmin}}(\sum_{i=1}^{t}(1-\xi))^{t-1}(Ed_i - (a+F_i)^2),\ with\ 0 < \xi < 1 \tag{11}$$

Here, the smoothed value at time t, $\tilde{Ed}_t$, then equals the local level k_t:

$$\tilde{Ed}_t = k_t \tag{12}$$

The smoothed series at time m and the trend is respectively given by the fitted value at m, and the trend by the fitted slope parameter:

$$Ed_t = \widehat{\alpha}_0 + \widehat{\beta}_0 m \quad and \quad F_m = \widehat{\beta}_0 \tag{13}$$

The *h-steps ahead* forecast is described by the following expression:

$$\widehat{Ed}_{t+h|t} = \tilde{E}d_t + hF_t \quad for \;\; t = m, ..., T \tag{14}$$

In section (5.3) we show the empirical estimates of the forecast.

4.2 Hidden Markov Models methodology

Let's suppose that it is possible to identify two different states in the economic situations (good and bad) and that we also know the cyclical sequence of cold and hot days. We aim at modelling the dynamics of electricity demand hypothesizing that we could have six possible (not observable directly) electricity demand volumes[11]. These states start from 1 to 6 with the state 1 corresponding to a very high demand intensity, 2 an high intensity, 3 for moderate high, 4 for just moderate low, 5 is when the demand is low and 6 when is extremely low. The market dynamic over time can be, then, expressed by some hidden states sequence that identify the variable in the realization of each possible state of nature i.e.: 1,1,2,2,1,3,5,5,6...etc.

Assuming now that the electric market dynamic (and its hidden corresponding state) can be affected by observable economic situations (e.g. in economic boosting periods industries and families demand more energy for production or consumption) and weather cyclical conditions (e.g. in summer the demand for electric energy reaches higher levels respect to spring and fall season, since it is pushed by an intense air conditioners consumption. In winter also, residential heat electricity consumption pushes the electricity demand to higher levels). The problem is that these different *electricity states* are not observable directly ex-ante, then we suppose that they are the hidden states (that we want to estimate).

In our model, we assume that the electricity demand is not directly observable, but that wheatear and economic conditions influence its future consumption. In other words, we estimate the model imposing that weather and economic conditions affect the distribution of the a priori transition probablity. In a dependent mixture model, the joint likelihood of observations $\mathbf{Y_t} = (\mathbf{Y_1}, \ldots, \mathbf{Y_T})$ and latent states $S_t = (S_1, \ldots, S_T)$, given model parameters θ and matrix of covariates $\mathbf{z_T} = (\mathbf{z_1}, \ldots, \mathbf{z_T})$, can be written as:

[11] The number of states of nature has been estimated by the bayesian theory

$$P(Y_T; S_T|\theta, z_T) = \pi_i(z_1) b_{S_1}(Y_1|z_1) \prod_{t=1}^{T-1} a_{ij}(z_t) b_{S_t}(Y_{t+1}|z_{t+1}) \quad (15)$$

in which we identify:

- S_t is an element of $S = 1, \ldots, n$, the set of the n latent classes or states for the electricity demand.
- $\pi_{z_1} = P(S_1 = i|z_1)$, giving the probability of class/state i at time $t = 1$ with covariate z_1.
- $a_{ij}(zt) = P(S_{t+1} = j|S_t = i, z_t)$, provides the probability of a transition from state i to state j with covariate z_t, where the specific covariate is the previous estimated status.
- b_{S_t} is a vector of observation densities $b_j^k(z_t) = P(Y_t^k|S_t = j, z_t)$ providing the conditional densities of observations Y_t^k associated with latent state j and covariate z_t, $j = 1 \ldots, n$, $k = 1, \ldots, m$.

For the example above, b_j^k could be a distribution function for the electricity demand (time varying variable), and a distribution for the weatear variable. In our model the transition probability functions a_{ij} and the initial state probability functions π depend on covariates z_t. From the above mentioned scheme, it derives that the log-likelihood of the probability of state to observe a specific volume of electricity demand y_t corresponding to the specific state S_t can be written as:

$$\ell(\cdot) = \prod_{t}^{T} log\phi_t \quad (16)$$

in which $\phi_t = P(Y_t|Y_1, \ldots, Y_{(t-1)})$.

For that it is possible to derive the forecast estimation of the state and of the corresponding value of the electricity demand for the time t.

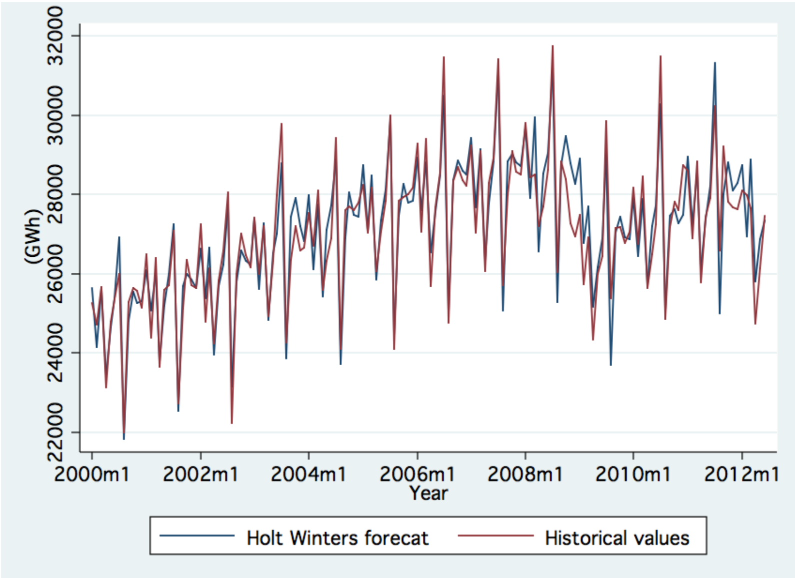
$$P(Y_{t+h} = y|\mathbf{Y_T}) = \phi_t \mathbf{\Gamma}^h \mathbf{P}(y) \mathbf{1'} \quad (17)$$

In the above equation the Γ matrix identifies the transition matrix with element $a_{ij} = f(\mathbf{z})$.

4.3 Empirical Results: H-W vs HMM

In both methodologies we have implemented the same models as expressed in (3). In figure (7) we plot the forecasted electricity demand obtained through the H-W methodology. It shows a path highly closed to the real evolution of the electricity demand time series. We produce our forecast for the last six months (2012:1 - 2012:6). In table (10), we report and compare the analytical estimate both for the H-W and the HMM forecasts.

Figure 7: Electricity demand with Holt-Winters filter smoothing approach.



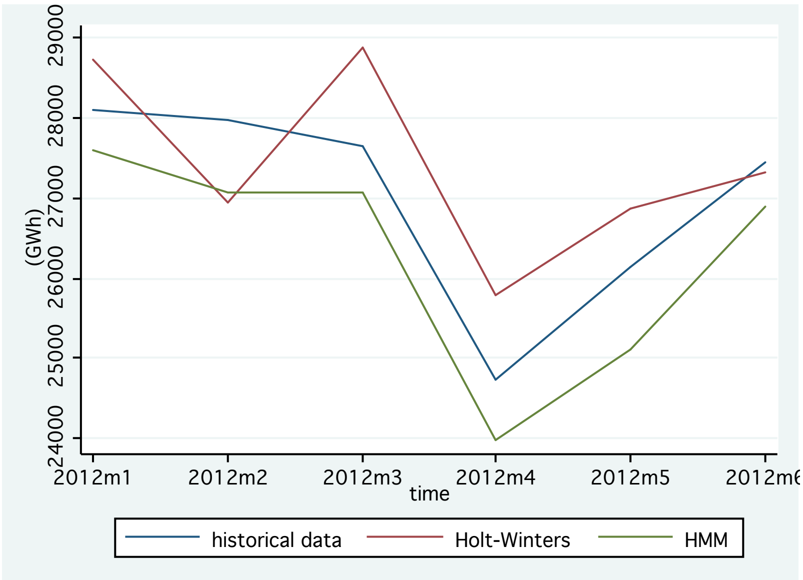
Source: Author's elaborations on Terna data. Twh on the y axes

The estimates obtained through the HMM approach seems to be consistent with the data generation process. Table (10) shows that the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) of the series obtained through the two approaches. As showed in table (10) the H-W approach shows a lower RMSE (27.2) compared to the HMM (49.9), meaning that the H-W approach shows a better forecasting performance compared to the MAPE.

However, in the forecasting econometric literature there is a shared consensus among researchers and praticitioners (e.g. Newbold, 1983[37], Thompson 1990[47], Armstrong et al.(1992)[3], Fildes [16] among the others), considering the RMSE as not reliable measure of the forecast goodness. It is principally for two reasons. First because unit-free measures are necessary for comparisons among forecast methods and RMSE is not unit-free (Ahlburg 1992)[1]; second because as clearly claimed in Goodwin et al. (1999)[23] the MSE is inappropriate since the major variations in the scale of observations between series risk to dominate the comparisons and thus unit free measures represent the most reliable measure of this kind.

Thus considering the MAPE as a forecasting comparison instrument, the HMM approach underestimates $(1,15\%)$ the forecasted electricity series, while the H-W methodology overestimates $(-2,13\%)$ the forecast showing a lower quality forecast for the database analyzed.

Figure 8: Comparing the electricity demand historial values with the forecasts obtained through the Holt-Winters filter and Hidden Markov Models.



Source: Author's elaborations on Terna data

5 Concluding Remarks

Considering the unstable and trend-changing path of the Italian electricity demand, the need for developing and use more sophisticated and effective tools and methods has recently emerged as never before for estimating: *(i)* endogenous breaks; *(ii)* reactivity to price and income changes on the consumption level; *(iii)* the ability to offer reliable future demand forecasts.

Two structural breaks in the electricity demand have been found in 2002:8 and 2004:8, mostly caused by the electricity market liberalizations effect on the demand.

Further our results provide insight into the electricity demand for Italy shedding lights on the relation between both price and income variations and the impact on the electricity demand. The originality of the paper lies principally in: *(i)* using monthly time series to study the Italian electricity demand while all the previous contributions are focused on yearly observations; *(ii)* implementing the HMM as forecasting methodology that can be extended into other specific business data analysis and forecasting.

The elasticities analysis has demonstrated that the price elasticity in residential and industrial sector are aligned with the magnitudes showed in the electricity demand literature with an higher elasticities in the long run, and a very low elasticity in the short run for both income and price variations. Forecasting performance was evaluated by analyzing different types of errors for the first semester of 2012. MAPE ranged from -2,13 (in the H-W case) to 1.15 (HMM). Even if there is a superiority in forecasting the electricity demand by the HMM, the difference between MAPEs is small and indicates that the two models have been quite successful in explaining and forecasting most of the oscillations and variability in the seasonal fluctuations.

We are confident that the findings showed in this paper represent a valid attempt both to extend the knowledge about the monthly Italian electricity demand and to produce useful insights for industrial strategic planning operations, energy government policies and further academia research.

6 Descriptive Statistics and Results

Table 5: Complete list of variables used in the empirical analysis

Variable (unit)	**Sample**	**Frequency**	**Source**	**Acronym**
Electricity demand (TWh)	2000m1 - 2012m6	Monthly	Terna	ed ita
GDP (Mln. Euro)	2000m1 - 2012m6	Monthly	IMF	gdp
Population (Mln.)	2000m1 - 2012m6	Monthly	IMF	pop
Industrial Production (2005=100)	2000m1 - 2012m6	Monthly	Istat	ind prod
Industrial price (cent/Euro al kWh, tax excl.)	2000m1 - 2012m6	Monthly	Enerdata	price ind
Residential price (cent/Euro al kWh, tax excl.)	2000m1 - 2012m6	Monthly	Enerdata	price res
Cooling degrees days	2000m1 - 2012m6	Monthly	Bloomberg Weather Platform	CDD
Heating degrees days	2000m1 - 2012m6	Monthly	Bloomberg Weather Platform	HDD

Source: Authors elaborations on Terna, IMF, Enerdata, Bloomberg Weather Platform.

.

Table 6: Descriptive statistics

	ed ita	**gdp**	**pop**	**ind price**	**res price**	**CDD**	**HDD**	**ind prod**
min	22003	1198292	56924	8.18	11	0	0	80.48
p1	22226	1202908	56927	8.33	11	0	0	80.57
p5	24239	1230601	56946	9.08	11	0	0	83.01
p10	24790	1256222	56967	9.25	11.08	0	0	84.70
p25	25798	1282844	57307	10	11.5	0	0	90.34
p50	27066	1390606	58722	12.9	13.26	0	5	100.80
p75	28174	1426215	60026	14.9	14.6	2.1	9.9	103.43
p90	29192	1438197	60549	15.5	15.3	4.6	13.4	106.08
p95	29814	1445132	60620	15.6	15.4	6.5	15	107.04
p99	31489	1449430	60964	15.7	15.5	7.2	13.4	108.88
sd	1768	74523	1338	2.51	1.64	2.15	5.26	7.98
skewness	-0.06	-0.49	0.19	-0.44	0.13	1.52	0.49	-0.74
kurtosis	3.46	1.83	1.58	1.36	1.36	3.93	1.97	2.19
N	150	150	150	150	150	150	150	150

Source: Authors elaborations on RTE, Terna and Bloomberg database.

Table 7: Regression to evaluate the importance of seasonal factors. Dependent variable: Electricity demand

	(1)	(2)
mseas1	0.022	0.028***
	0.019	0.013
mseas2	-0.0325*	-0.025**
	0.020	0.014
mseas3	0.038*	0.033*
	0.023	0.016
mseas4	-0.041	-0.056***
	0.028	0.020
mseas5	-0.004	-0.004
	0.033	0.023
mseas6	0.049	0.006
	0.039	0.028
mseas7	0.131**	0.085***
	0.040	0.0288
mseas8	-0.069**	-0.109***
	0.368	0.026
mseas9	0.040	0.015
	0.031	0.022
mseas10	0.046**	-0.028
	-0.0269	0.019
mseas11	0.009	0.004
	0.020	0.014
mseas12	xxx	xxx
	xxx	xxx
CDD	0.014*	0.004**
	0.042	0.003
HDD	0.004*	0.001*
	0.002	0.001
t		0.007***
		0.007
Observations	144	144
R^2	0.556	0.781

Source: Authors estimates Terna data.

Table 8: Augmented Dickey-Fuller and Philips Perron tests for unit roots.

Independent Variables	**ADF**	**Number of Lags**	**PP**	**Number of Lags**	**Integration Order**
log ed ita	-1,824	12	-283,052	12	I(1)
log gdp	-2,122	2	-2,006	2	I(1)
log prod ind	-0.886	12	-3.506	12	I(1)
log price res	-0,546	2	-1,389	2	I(1)
log price ind	-1,227	2	-1,978	2	I(1)

Source: Augmented Dickey-Fuller (ADF) and Philips Perron (PP) tests for unit roots. Authors estimates on Terna, Enerdata and IMF data.

Table 9: Results of industrial and residential price and income elasticities regression. The t statistics are reported in parenthesis.

	(9)	(10)	
log price ind	-0.18***		
	(-3.53)		
log price res		-0.013***	
		(-2.88)	
log gdp	0.06***		
	(2.98)		
log income		0.041***	
		(2.05)	
log ed ita l.12	0.9***	0.9***	
	(25.69)	(25.89)	
Constant	-7.2***	0.06***	
	(-2.59)	(0.10)	
Observations	144	144	
R^2	0.98	0.98	
D-Watson	2.15	2.2	
F	5704.7	6269.4	

Source: Authors estimates.

Table 10: Historical data and Forecasts results with H-W, One Step Ahead and HMM.

	(Historial Data)	(H-W)	(HMM)
January 2012	28093	28732	27593
February 2012	27985	26943	27062
March 2012	27641	28879	27063
April 2012	24728	25795	23974
May 2012	26141	26868	25114
June 2012	27456	27335	26898
MAPE		*-2.13*	*1.15*
RMSE		*27.2*	*49.9*

Source: Authors estimates on Terna data.

References

[1] Ahlburg, D. A., "Error Measures and the Choice of a Forecast Method", *International Journal of Forecasting*, n. 1, vol. 8, pp 99-100, 1992.

[2] Andrews D. Zivot E., "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit-Root Hypothesis", *Journal of Business and Economic Statistics*, pp. 251-270, Vol. 10, 1992.

[3] Armstrong J.S. and Collopy F, "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons" *International Journal of Forecasting*, vol. 8, 1992

[4] Bernstein, M.A Griffin J., "Regional Differences in Price-Elasticity of demand for Energy", *The Rand Corporation Technical Report*, 2005.

[5] Blazquez L., Boogen N., Filippini, M., "Residential Electricity Demand for Spain: New Empirical Evidence Using Aggregated Data", *CEPE Working Paper*, pp. 1-23, 2012.

[6] Bohi D., Zimmerman M., "An Update on Econometric Studies of Energy Demand Behavior", *Annal Review Energy*, 119-39, 26, 1984.

[7] Borenstein S., "To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing", *Working Paper - University of California, Berkeley*, 2009.

[8] Cipra. T., "Robust Exponential Smoothing", *Journal of Forecast*, vol 11, pp. 57-69, 1992.

[9] Cipra. T., Romera. R., "Kalman Filter with Outliers and Missing Observations", *Test* vol. 6 n.2, pp. 379-395, 1997.

[10] Clemente, J., Montanes, A., Reyes, M., "Testing for a Unit Root in Variables with a Double Change in the Mean," *Economics Letters*, Vol. 59, pp. 175-182, 1998.

[11] Considine, J.T., "The Impacts of Weather Variations on Energy Demand and Carbon Emissions", *Resource and Energy Economics*, vol. 22, 295-314, 2000.

[12] Dickey, D.A., Fuller, W.A., "Distribution of the Estimators for Autoregressive Time Series With a Unit Root", *Journal of the American Statistical Association*, pp. 427-431, Vol. 74, 1979.

[13] Engle R.F. Granger C.W., Rice, John, Weiss, Andrew, "Semiparametric Estimates of the Relation between Weather and Electricity Sale", *Journal of the American Statistical Association*, n.394, vol. 81, pp. 310-320, 1986.

[14] Erdogdu E., "Electricity Demand analysis Using Cointegration and ARIMA Modelling: A Case Study of Turkey", *Energy Policy*, n. 2, vol.35, pp.1129-1146, 2007.

[15] Espey, J.A. Espey, M., "Turning on the Lights: A Meta Analsysis of Residential Electricity Demand Elasticities", *Journal of Agricultural and Applied Economics*, vol.36, pp. 65-81, 2004.

[16] Fildes R., "The Evaluation of Extrapolative Forecasting Methods", *International Journal of Forecasting*, vol. 8 pp 88-98, 1992.

[17] Filippini. M., "Swiss Residential Demand for Electricity by Time-of-Use", Resource and *Energy Economics*, n.3, vol.17, pp. 281-290, 1995.

[18] Filippini. M., "Swiss Residential Demand for Electricity", *Applied Economic Letters*, n.8, vol.6, pp. 533-538, 1999.

[19] Fruhwirth-Schnatter, S., "Finite Mixture and Markov Switching Models", *Springer Series in Statistics*, Springer, New York., 2006.

[20] Garcia-Cerutti, L. Miguel, "Estimating Elasticities of Residential Energy Demand from Panel County Data Using Dynamic Random Variables Models with Heteroskedastic and Correlated Error Terms", *Resource and Energy Economics*, vol.22, pp.355-366, 2000.

[21] Gelper S.E.C Fried R. Croux C., "Robust Forecasting with Exponential and Holt Winters Smoothing", *Journal of Forecasting*, vol. 29, pp. 285-300, 2010.

[22] Ghysels, E., "On the Economics and Econometrics of Seasonality", in C.A Sims (ed.), "Advances in Econometrics - Sixth World Congress of the Econometric Society," *Cambridge University Press*, Cambridge, 1994.

[23] Goodwin P, Lawton R, "On the Asymmetry of the Symmetric MAPE", *International Journal of Forecasting*, vol.15 pp. 405-408, 1999.

[24] Hamzacebi C., "Forecasting of Turkey's Net Electricity Energy Consumption on Sectoral Bases", *Energy Policy*, vol. 35, pp. 2009-2016, 2007.

[25] Henley, A., Peirson, J, "Non-linearities in Electricity Demand and Temperature: Parametric versus non Parametric Methods" *Oxford Bulletin of Economics and Statistics*, vol. 59, pp. 1149-1162, 1997.

[26] Henley, A., Peirson, J, "Residential Energy Demand and the Interaction of Price and Temperature: British experimental Evidence", *Energy Economics*, vol.20, pp. 157-171, 1998.

[27] Holt C., "Forecasting Seasonals and Trends by Exponentially Weighted Moving Average", *Office of Naval Research Memorandum*, vol. 52, 1969.

[28] Johnsen, Tor Arnt, "Demand, Generation and Price in the Norwegian Market for Electric Power, " *Energy Economics*, vol. 23, pp. 227-251, 2001.

[29] Labandiera, X., Labeaga Azcona, J., Rodriguez Mendez, M., "A Residential Energy Demand System for Spain" *Energy Journal*, vol. 27, pp. 87-112, 2006.

[30] Labandiera, X., Labeaga Azcona, J., Rodriguez Mendez, M., "Estimation of Elasticity Price of Electricity with Incomplete Information" n.3 vol. 34, *Energy* Economics, 2012.

[31] Langeheine, R., Van de Pol, F., "A Unifying Framework for Markov Modeling in Discrete Space and Discrete Time" *Social Methods and Research*, n.4, vol. 18, pp 416-441, 1990.

[32] Lee. C. C. Chiu. Yi-Bin, "Electricity Demand Elasticities and Temperature, " *Australian Conference of Economists*, 2011.

[33] Kim, C.-J., "Dynamic Linear Models with Markov-switching", *Journal of Econometrics*, vol. 60, pp 1-22, 1994.

[34] Miller, G. A., Chomsky, N., New York, Wiley, "Finitary models of language users" (chap 13), In R. Luce, R. R. Bush, E. Galanter (Eds.), Handbook of mathematical psychology, 1963.

[35] Miller, E.K. Rainer, G., "Neural ensemble states in prefrontal cortex identified using a hidden Markov Model with a modified EM algorithm" *Neuralcomputing*, vol. 32-33, pp. 961-966, 2000.

[36] Neeland H., "The Residential Demand for Electricity in the United States", *Economic Analysis and Policy*, n.2, vol. 39, 2009

[37] Newbold P., "The Competition to End all Competitions" *Journal of Forecasting*, vol.2, pp. 276-279, 1983.

[38] Pardo A. Meneu V. Valor E., "Temperature and Seasonality Influences on Spanish Electricity Load", *Energy Economics*, vol.24, pp. 55-70, 2002.

[39] Phillips P.C.B Perron P., "Testing for Unit Roots in Time Series Regression" *Biometrika*, vol.75, pp. 335-346, 1988.

[40] Prais, S.J., Winstein, C.B., "Trend Estimators and Serial Correlation", *Cowles commission discussion paper: statistics* no. 1954.

[41] Rabiner, L.R., “A tutorial on hidden Markov models and selected applications in speech recognition” *Proceedings of IEEE*, n. 2, vol.77, pp. 267-295, 1989.

[42] Reiss. P.C., “Household Electricity Demand, Revisited”, *Review of Economic and Studies*, vol.72, pp. 853-883, 2005.

[43] Sailor, D.J., Munoz, J.R, “Sensitivity of Electricity and Natural Gas Consumption to climite in the U.S.A: Methodology and Results for Eight States” *Energy*, pp. 987-998, vol, 22, 1997.

[44] Schmittmann, V. D., Visser, I., Raijmakers, M. E. J., “Multiple learning modes in the development of rule-based category-learning task performance”, *Neuropsychologi*, n. 11, vol. 44, pp. 2079-2091, 2006.

[45] Taylor J.W., “Short-term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing”, *Journal of the Operational Research Society*, vol. 54, pp. 799-805, 2003 .

[46] Taylor. L., “The Demand for Electricity”, *Bell Journal of Economics*, vol.6, pp. 71-110, 1975.

[47] Thompson P.A., ”A MSE statistic for comparing mean square forecast errors”, *Journal of Forecasting*, vol. 6, pp. 219-227, 1990.

[48] Valor, E., Meneu, V., Caselles, V., “Daily Air Temperature and Electricity Load in Spain”, *Journal of Applied Meteorology*, vol.40, pp. 1413-1421, 2001.

[49] Wickens, T. D., “Models for behavior: Stochastic processes in psychology” San Francisco: W. H. Freeman and Company., 1982.

[50] Winters P., “Forecasting Sales by Exponentially Weighted Moving Averages” *Management Science*, vol. 6, pp.324-342, 1960.

[51] Yan, Y.Y., “Climate and Residential Electricity Consumption in Hong Kong”, *Energy* n. 1, vol 23. pp. 17-20, 1998.