International Journal of Energy Economics and Policy
Vol. 4, No. 4, 2014, pp.744-758
ISSN: 2146-4553
www.econjournals.com

Electricity Price Forecast: a Comparison of Different Models to Evaluate the Single National Price in the Italian Energy Exchange Market

Andrea Cervone
Department of Electrical Engineering –SAPIENZA University of Rome
Via Eudossiana 18, 00184 Rome, Italy. Email: andrea.cervone@uniroma1.it

Ezio Santini
Department of Electrical Engineering –SAPIENZA University of Rome
Via Eudossiana 18, 00184 Rome, Italy. Email: ezio.santini@uniroma1.it

Sabrina Teodori
Department of Electrical Engineering –SAPIENZA University of Rome
Via Eudossiana 18, 00184 Rome, Italy. Email: sabrina.teodori@uniroma1.it

Donatella Zaccagnini Romito
Department of Electrical Engineering –SAPIENZA University of Rome
Via Eudossiana 18, 00184 Rome, Italy. Email: donatella.zaccagnini@uniroma1.it

ABSTRACT: In the last decades, electricity markets throughout the Eurozone have undergone substantial changes. The deregulation of electricity markets stimulated investments in the production and distribution of energy, but there are large risks associated with these investments due to price volatility. The paper in the introduction describes the algorithm that governs the operation of the Day-Ahead Market in the Italian Power Exchange and proposes an econometric model for short-term forecasting (six months or a year) of the daily Single National Price (*Prezzo Unico Nazionale,* PUN) of electricity. The model includes constants, regressors, moving averages, weekly and seasonal dummies, autoregressive and heteroschedastic variables. The results show a significant decrease in error of the short-term forecast of the analyzed time series, in comparison with the method of linear least squares, traditionally used in literature. An analysis on the influence of different variables on PUN such as brent, solar radiation and weather has been reported. A comparison of the different models with specific indices have been performed and discussed.

Keywords: Electricity prices; Day-Ahead Market; Italian Power Exchange; ARMA–GARCH model; Forecasting.
JEL Classifications: C5; C51; L; L1; L11

1. Nomenclature

J is the number of generation plants
K_j is the maximum Energy output of plant j
λ_j is the marginal cost of Energy production of plant j
$d_j(t)$ is the availability of production of plant j at time t. It is a stochastic variable in the range [0, 1]
$Y_j(t)$ is the actual energy produced from plant j at time t.
$D_i(t)$ is the Energy demand
$w_i(t)$ is a parameter capturing climatic conditions
π_i is the profit function of consumer i
$p_i(t)$ is the Energy purchase price of consumer i
$\vec{Z}(t) = \langle Z_1(t), \dots, Z_k(t) \rangle$ is the vector of the energy flow
$L(t)$ are the transmission losses
$Z_{min,k}$, $Z_{max,k}$ are the maximum and minimum flow rates for each k transmission line

$b_k(t)$is the actual availability of transmission of the *k-th* transmission line. It is a variable in the range [0, 1].
$\mu_j(t)$ is the coefficient that shows how much energy a plant j must produce compared to the maximum production capacity,.
$\eta_k(t)$ is the coefficient that depicts the shadow price of an additional unit of capacity of the line k.
$\theta(t)$ is the shadow price associated with an additional unit of demand anywhere in the network. This is the optimal price with respect to an arbitrarily node, chosen s the points of measurement
$E_i(t)$ is the Energy price of i consumer at time t.
PUN is the daily price of energy in the Day-Ahead Market
PB is the daily price of a barrel of Brent

2. Introduction

The creation of a true internal energy market is a priority for the European Union, which has been implemented throughout the Community in several steps since 1999. This has the aim to offer all of the consumers' European Union, citizens or businesses, real freedom of choice, create new business opportunities and more cross-border trade. This will lead to greater efficiency, competitive prices, higher service levels, and contribute to security of supply and sustainable development. This paper represents a first step in the forecast of electricity prices in the European energy exchange markets and, in a broader sense, in all electricity markets worldwide.

Since the transition to the deregulated energy market of electricity in Italy occurred, the notion of predictability has acquired considerable importance for investors in the energy market. Predicting the trend of energy prices in the near future with the lowest possible margin of error is crucial in order - for market participants - to set the ask and bid prices of energy (Gianfreda et al., 2010), thereby decreasing the risk associated with foregone transactions (Braun and Lai, 2006). The Italian power exchange is subject to such risk, due to price volatility, competition and congestions. That is why the paper focuses on the problem of estimating future energy prices in the Italian Day-Ahead Market. In literature is possible to find different authors that analyzed the problem of Day-Ahead price forecast such as Contreras et al. (2003), Garcia et al. (2005) and Bowden and Payne (2008) but few contributions can be found about the Italian Power Exchange: Bosco et al. (2008) conducted an analysis of the time series of daily average prices, Gianfreda (2010) evaluated the volatility of electricity market and with Grossi (2012) used a ARFIMA models with exogenous variables to forecast the zonal prices. In the present paper different models of Single National Price forecast are compared applied to Italian Day-Ahead Market. By analyzing the time series by means of a model characterized by autoregressive behavior, moving average and periodicity, the proposed models assured the dynamic features that are necessary to have a more consistent and accurate forecast of the value of energy prices. Due to these characteristics, it is possible to evaluate the PUN's volatility and to decrease the error making the exchange of energy in the electricity market a safer investment, with lower risk. Girish et al. (2014) reported a review analysis of the methods that can be used to forecast the electricity price in deregulated electricity market.

3. Structure of the Optimal Price Algorithm in thr Italian Day-Ahead Market

The optimization problem, formulated by Bohn et al. (1984) is known as the "nodal spot pricing" and can be used for the operation of the Day Ahead Market. The model assumes the maximization of consumer welfare and considers the following inputs: power, producible energy and a stochastic variable capturing whether each of several generators operates. Each generator sets the price at which it sells energy to a number of consumers, whose demand has a stochastic pattern that depends on several factors, as described below.

The model is based on four principles: demand and supply are spatially located; electricity travels through a fixed network, whose operation is of a stochastic type; both demand and generator availability are stochastic variables; the market clearing price adjusts itself instantly.

To describe a unique algorithm that considers all variables, the system is divided into three macro-zones, according to the functionality that each component has within the system: generation, transmission or demand.

The generation is always composed of J plants that operate at a certain time, the availability of energy is a partially stochastic variable, whose value range is [0, 1]. This value limits the amount of energy that can be transferred through the network at any given time. For each plant, the inequality (1) is always verified

$$0 \leq Y_j(t) \leq K_j * d_j(t) \tag{1}$$

The energy demand is independent from generation. It depends on many variables such as weather conditions, time, day, type of user, price of electricity, and many other factors. For each consumer i of energy, a function F is considered which represents the added value that a consumer places on the use of energy. F is strictly dependent on $D_i(t)$ and $w_i(t)$, at time t.

This function, after deduction of the energy costs sustained by a consumer who buys energy, represents the profit (π) of the consumer, as shown in Eq.(2).

$$\pi_i = F_i(D_i(t), w_i(t)) - p_i(t)D_i(t) \tag{2}$$

The profit function has to be maximized, since the purpose of the model is to find the intersection point between supply and demand, given maximum gain value for all users operating in the electricity market, as shown in Eq. (3) and subsequent manipulations (4) (5).

$$\max \pi_i = \frac{\partial[F_i(D_i(t), w_i(t)) - p_i(t)D_i(t)]}{\partial D_i(t)} = 0 \tag{3}$$

$$\frac{\partial F_i(D_i(t), w_i(t))}{\partial D_i(t)} = p_i(t) \tag{4}$$

$$D_i(t) = D_i(p_i(t), w_i(t)) \tag{5}$$

The value of demand is not necessarily positive, but may also be negative. In fact, a consumer may also be, at times, a producer of energy stipulating a contract with the distributor.

The boundaries of the model of "nodal spot pricing" are not only dependent upon supply and demand, but it is also necessary to analyze energy transmission. Transmission occurs through K nodes. The losses throughout the network are a function of a vector, which considers the flow of energy, as shown in Eq. (6).

$$L(t) = L\left(\vec{Z}(t)\right) \tag{6}$$

The energy balance can be written as in Eq. (7):

$$\sum_{j=1}^{J} Y_j(t) = L(t) + \sum_{i=1}^{I} D_i(t) \tag{7}$$

A violation of this balance leads to loss of stability of the generators and to the possible collapse of the network. The power lines give another boundary: the capacity of each line has a minimum and a maximum value. The inequality (8) must be verified.

$$Z_{min,k}b_k(t) \leq Z_k(t) \leq Z_{max,k}b_k(t) \tag{8}$$

The model uses standard welfare maximization, by maximizing the profit of consumers and producers while satisfying the boundary conditions analyzed above. The boundary conditions depend on the number of generating plants, the transmission lines from the reserve and stochastic exogenous variables (climatic variables, working of lines and generators, fuel price).

At time t, given the presence of real and unavoidable boundary, the complete model has the following formulation (9), expressed in Figure 1.

The basic principle of the search for the shadow price is that market receives bids for purchase and sale, for a given day and a given time. Under the assumptions made by the Transmission System Operator (TSO), the electricity market operator GME (*Gestore del Mercato Elettrico*) collects the proposals and creates a supply curve of energy; then a fixed increment of the sales price per kilowatt hour (kWh) of energy is determined. For each price range it has to be evaluated how much of the total energy is for sale at a price lower than the predetermined price. Similarly, a curve is created for buyers.

The intersection point between the two curves represents the shadow price ($\vartheta(t)$).
In contrast to the algorithm of Bohn, Day-Ahead Market of electric energy in Italy has a different assumption since the "nodal spot price" supposes the algorithm to impose how much energy plants must produce, in order to find a maximizing function leading to the lowest possible price of energy.

Differently, in a liberalized market producers determine demand and supply and the market determines the price, in order to guarantee the plurality of participants.

$$\max_{Y_j(t),p_i(t)} \sum_{i=1}^{I} F_i(D_i(t), w_i(t)) - L(t) - \sum_{j=1}^{J} \mu_j(t)[Y_j(t) - K_j d_j(t)] - - \sum_{i=1}^{I} D_i(t)\,] \sum_{k=1}^{K} (Z_k(t) - Z_{k,max}(t))\eta_{k,max}(t)) + \sum_{k=1}^{K} (Z_{k,min}(t) - Z_k(t))\eta_{k,min}(t) \quad \textbf{(9)}$$

Each node of the grid communicates with other nodes through the lines that can reach their limits of capacity determined by equation (8). TSO must ensure that the prescribed limits are not exceeded and if is not respected an efficient and reliable transfer of energy between different zones, the GME algorithm provides a division of markets in different areas.

The algorithm of the "nodal spot price" is applied to each individual area, different submarkets are originated in which a single equilibrium price is determined. Each zone looks for his shadow price between energy produced and sold through the inlet of the surplus of energy required by the communication networks with the other zones, until reaching the technical limit of the lines.

Figure 1 shows the six areas in which the Italian Day-Ahead Market is divided. In addition, the figure shows the maximum transmission capacity between related areas and areas and electrical connections to foreign countries.

The shortage of transport capacity of the lines causes the equilibrium prices between different areas to be different since higher cost offers of generation have to be accepted in the in import area. The local shadow price is the rate of electricity sales to bidders, but not the purchase price to consumers.

To determine the price of energy purchases in the free market it is necessary to calculate the average price of each area, weighted by the energy consumption of a single zone, as shown by (10).

$$PUN = \frac{\sum_i E_i(t)\theta_i(t)}{\sum_i E_i(t)} \quad \textbf{(10)}$$

The value obtained is defined PUN, Single National Price.

Figure 1. Zonal Distribution of Italian Grid (Terna)

4. Construction of a Mathematical Model for the PUN Forecast

To define a single econometric model to estimate the daily price of electricity in the Day-Ahead Market, data of the equilibrium price of electricity are used. Data include the average daily national prices (hereafter simply PUN) of energy from January 1, 2007 to December 31, 2011, for a total of 1826 observations (source: GME). There are many models that approximate a time series and the root mean square error allows comparing their performances.

The Root Mean Square error (RMSE) is a measure of the differences between values predicted by a model or an estimator and the values actually observed (10)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} (X_{measured,i} - z_{estimed,t})^2}{n}} \quad \textbf{(11)}$$

where *n* is the number of observations.

Some preliminary considerations have to be discussed firstly. The series of PUNs presents clearly a not-stationary trend since it does not oscillates around any average value. In contrast, there is a growing trend in energy prices, although it is not known whether it is deterministic or stochastic.

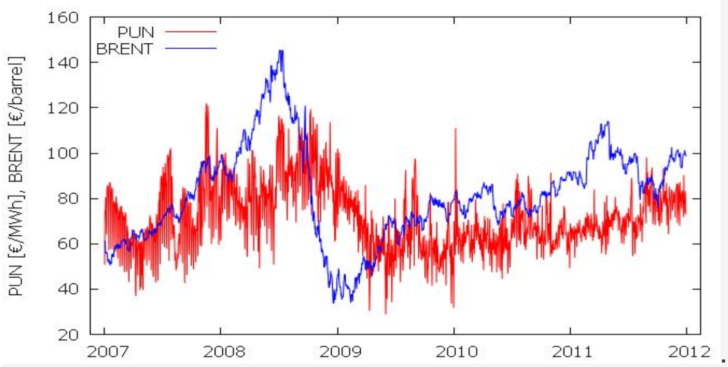
Given the dependence of PUN on many external variables that affect the daily pattern, the analyzed series can be studied through the composition of these variables. The chosen regressor is the price of a barrel of Brent of North Sea (source EIA) as in Rastegar et al. (2009).

Since the world oil exchange, differently to the Italian electricity market, does not allow trading on public holidays, the price is considered constant in these days and fixed at the price on the prior day.

Figure 2 shows the trend of the two series considered for the period 2007-2011. By inspecting the graphs it can be noted that there is a dependence of energy prices in the Italian market on the price of oil, since a change in the price of oil corresponds to a change of the price of PUN, but with greater volatility (Regnier, 2007).

To perform a regression, excluding the white noise, the PUN must be properly analyzed to understand how future values can be computed. For this reason, it is employed the Dickey-Fuller test (ADF).

Figure 2. Comparison between development of PUN and Brent price



The Dickey-Fuller test can determine whether a time series is stationary. It tests the null hypothesis of unit root against the alternative hypothesis of an autoregressive process. In (12) is shown the null hypothesis:

$$H_0 : \varphi = 1 \tag{12}$$

If the null hypothesis is not ruled out, the series is a Random Walk with Drift (RWD), and its performance is not stable. However, if the null hypothesis is rejected in favor of the alternative hypothesis (13), the series has an autoregressive behavior. In the case where there is also a drift, the series is approximated to an ARMA.

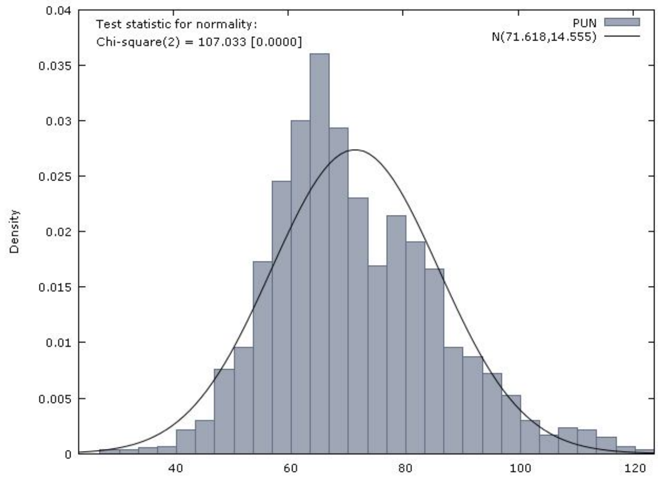
$$H_1 : |\varphi| < 1 \tag{13}$$

In the case in question, the ADF test give as result a p-value of 0.038 that is lower than the critical value of 0.05. Consequently, the hypothesis of unit root is rejected. The PUN is not a simple random walk with drift, dependent on prior value, a trend and an error. Rather the model is approximated to an ARMA with an autoregressive term and a moving average.

To accurately estimate the performance of a time series, a normality test must be performed. The Jarque-Bera test examines the skewness and kurtosis of the distribution of a series. The null hypothesis is that data are normality distributed; the alternative hypothesis is that the distribution is uneven. In Figure 3 the results are reported.

To verify that the series is normal, the p-value is discussed. Its value ($1.10106\ 10^{-23}$) is much lower than 0.05 so the distribution of the values of the price of energy cannot be considered normal. The PUN series, consequently, is not a normal series. This result leads to confirm the volatility clustering (Wang and Xiang, 2011), the inability to control the energy price since it presents element of unpredictability.

Figure 3. Normality Test of Jarque-Bera (Gretl)



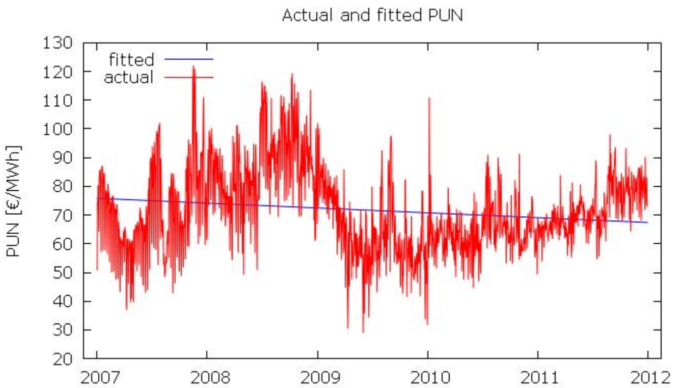
A. Ordinary Least Squares with trend

The PUN series shows a trend of growth resulting from increases in the price of fuel used in many generation plants in Italy. A consistent model of energy price can be built considering a first-order polynomial constituted by a constant plus a linear time function; this model is shown in (14).

$$y(t) = C_1 + C_2 t \qquad \textbf{(14)}$$

This feature, that uses only a deterministic trend, gives rise to the simplest model. By applying the ordinary least squares method to this function, the results in Figure 4 are obtained. The root mean square error of this model is equal to 14.345 €/MWh, which corresponds to a percentage error of 20.02 % compared to the average of the values of the PUN in the considered period. The low value of the determination index indicates the poor ability of this regressor to model the data in the short term.

Figure 4. Fitted and actual PUN with OLS with trend and constant



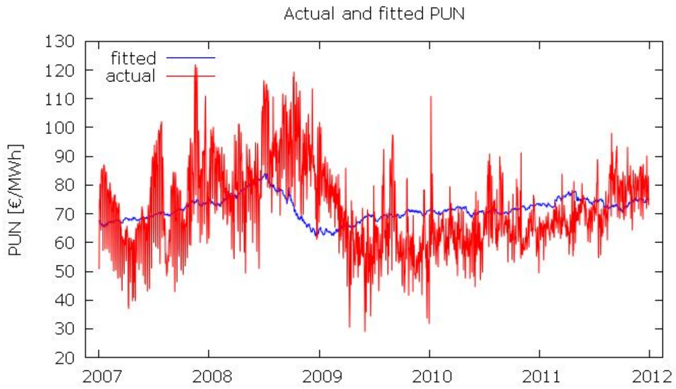
B. Ordinary Least Squares with regressor

To reduce the error in order to obtain a better estimation of the time series, the time variable can be replaced with the price of Brent. The function with regressor is written in (15).

$$y(t) = C_1 + C_2 PB \qquad \textbf{(15)}$$

The results provided by are shown in Figure 5. The RMSE is 13.964 €/MWh. This value represents an error of 19.49% compared to the considered series, and an improvement of 1.5% compared to the previous model. In this case, the method of the least squares with regressor is a more accurate method of forecasting the value of PUN in the Italian stock exchange.

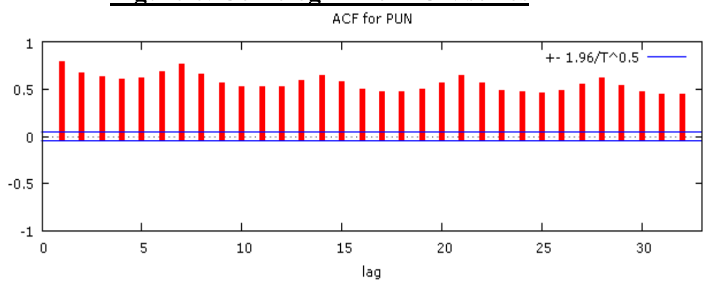
Figure 5. Fitted and actual PUN with OLS with regressor and constant



C. *Ordinary Least Squares with regressor, trend and weekly dummies*

To improve the model it is necessary to determine which data in the series must be taken into account as additive variable. The first step is to introduce a model with moving average (MA). The method of Box-Jenkins is used to obtain information by means of the use of a correlogram. Considering the case of the PUN series, the test of autocorrelation gives rise to the graph of Figure 6.

Figure 6. Correlogram of PUN series



This graph shows in x-axis the number of Lag and in y-axis the autocorrelation function (ACF) index associated to the time delay. The ACF graph allows an assessment of the degree of drift of the series. Since a correlation can be noticed for each of the twenty-five lags, the series has a very high degree of drift. However an important aspect can be deduced by noting that there is a subdued oscillation every seven days. The introduction of specific dummies can improve the accuracy of the model reducing the periodicity given by the weekly oscillations. The dummies are binary additive terms that assume value 1 on the analyzed day of the week and 0 on other days.

The model is described in (16) and the results are reported in Figure 7. The RMSE is equal to 12.938 €/MWh that is the 18.06% of the average price.

$$.y(t) = C_1 + C_2 PB + C_3 t + \sum_{i=1}^{7} C_{4i} d_i \qquad \text{(16)}$$

Figure 7. Fitted and actual PUN with OLS with regressor, trend and weekly dummies

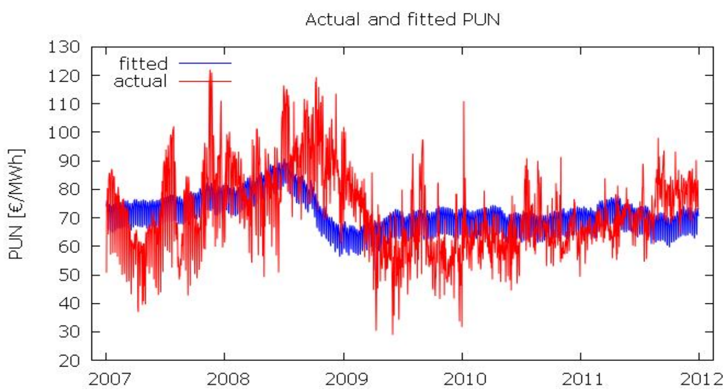
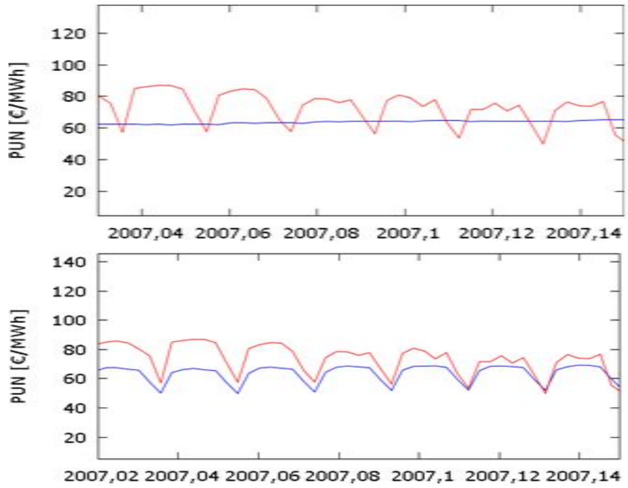


Figure 8. Zoom of fitted and actual PUN from the second to fourteenth week of 2007

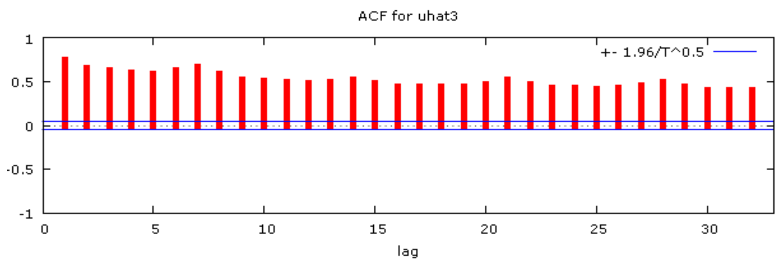


In addition to the reduction of the RMSE of about 1.5% compared to the previous model, the improved model gives a better estimation in the days of low power demand and low price, as show in Figure 8.

D. *Autoregressive model with regressor, trend, weekly dummies*

By the method of Box-Jenkins with the correlogram ACF additional information can be obtained considering the PUN series without trend and seasonally adjusted. The correlogram analyzes the residual information and the results are reported in Figure 9.

Figure 9. Correlogram of PUN series seasonally adjusted



The series still shows a weakly oscillation, but less pronounced if compared to the model analyzed in the previous section. The use of an Autoregressive Moving Average Model (ARMA) can lead to an improvement of the accuracy of the model. The Order of AR is chosen comparing the models with the Akaike Information Criterion (AIC). This paper analyzes the AIC for AR(p) and compare it with AR(p+1), starting from AR(1). If AIC for AR(p) is less of AR(p+1), the paper consider a good approximation of model the AR(p).

The index of Akaike Information Criterion is calculated as (17):

$$AIC = 2k - 2\ln(L) \qquad \textbf{(17)}$$

Where k is the number of parameters of model and L is the maximum value of the likelihood of the model. The results of output are shown in Table I.

TABLE I. AIC INDEXS FOR AR(1) AND AR(2)

AR(1)	AR(2)
6713,456	6715,243

The descriptive function of the model is shown in (18).

$$y(t) = C_1 + C_2PB(t) + C_3t + \sum_{i=1}^{7} C_{4i}d_i + C_5AR(1)(t) \qquad \textbf{(18)}$$

To make forecasts by means of the considered model, the method of Cochrane-Orcutt is necessary. The function $\mathrm{y(t) - Ly(t-1)}$allows the calculation of the coefficient of lag, L, in an iterative process on $\mathrm{u_t}$. Known L, the model assumes function (19).

$$y(t) = C_1 + C_2PB(t) + C_3t + \sum_{i=1}^{7} C_{4i}d_i + \mathrm{Ly(t-1)} \qquad \textbf{(19)}$$

The analyzed model has the RMSE equal to 8.1317 €/MWh. This value corresponds to an error of the average values of the series PUN during the period amounted to 11.35%.

E. *Autoregressive model with regressor, trend, weekly dummies, seasonally dummies*

Thanks to the information about the trend of the Italian electricity demand it is possible to construct a more accurate model (Petrella and Sapio, 2009). As can be seen from the graph of Figure 10 that shows monthly request of energy in Italy, there is a considerable seasonal component. Like in the previous case, dummies that take the value of 1 in the month considered and 0 otherwise are introduced. To the model a new parameter is added, arriving to the form in Eq. (20).

$$y(t) = C_1 + C_2PB(t) + \sum_{i=1}^{7} C_{4i}d_i + \sum_{i=1}^{12} C_{5i}S_i + Ly(t-1) + \varepsilon_t \qquad \textbf{(20)}$$

By applying the model of Cochrane-Orcutt for the calculation of coefficient associated with the lag, the results that are arrived are resumed in Figure 11.

Figure 10. Monthly report on Italian electricity consumption (Terna)

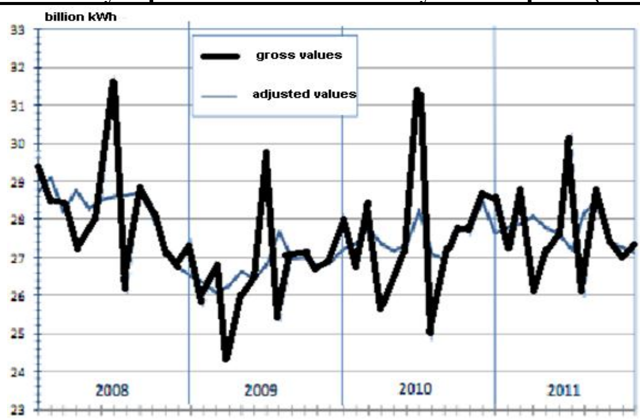
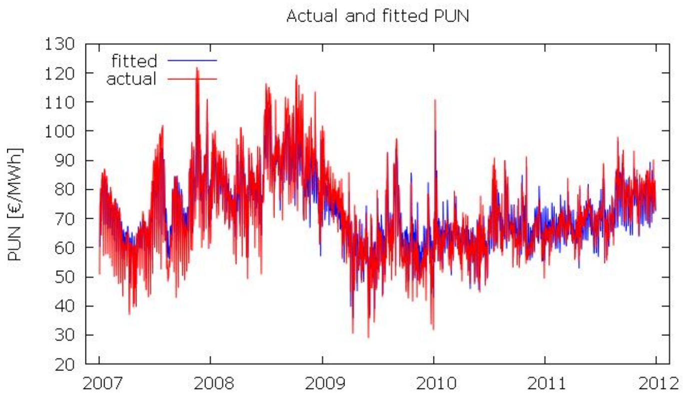


Figure 11. Fitted and actual PUN with ARMA with regressor, trend and weekly and seasonally dummies



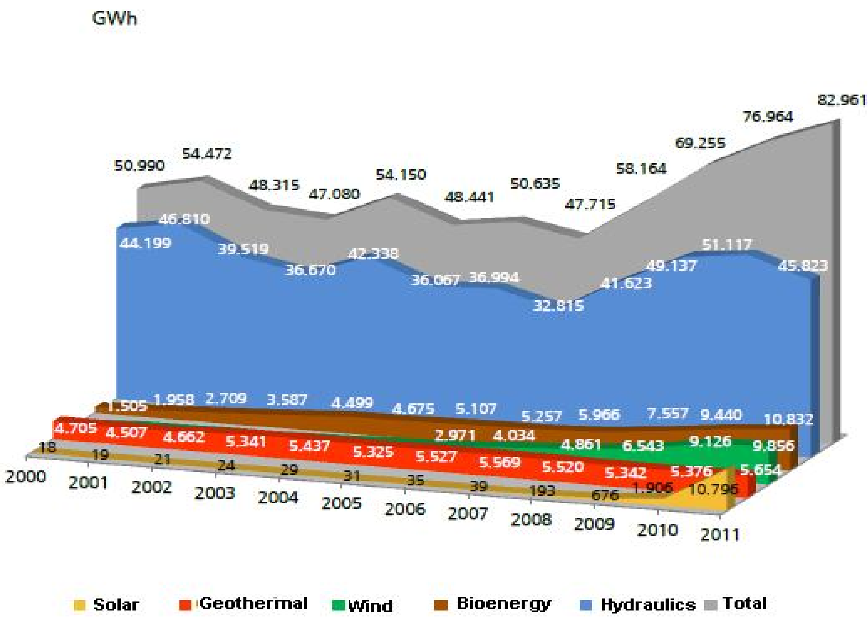
The model gives a RMSE equal to 8.0246 €/MWh, which corresponds to an error on the average PUN equal to 11.20%.

F. *Autoregressive model with trend, weekly dummies, seasonally dummies, solar radiation, regressor and weather*

An analysis of other variables that can influence PUN has been conducted introducing solar radiation and weather conditions. The solar radiation can affect PUN since it influences loads and photovoltaic production. Energy consumptions are strictly dependent on external temperature, which is influenced by solar radiation. Then, the electrical energy demand depends also on solar radiation.

The photovoltaic production depends on solar radiation as well. When the irradiance increases, the power produced by photovoltaic power plants grows. Photovoltaic energy has been increasing in the recent years reaching in 2011 more than 10 TWh as can be seen in Figure 12.

Figure 12. Trend of Renewable Energy in Italia from 2000 to 2011 (Statistical Report GSE, 2011)



The influence of the solar radiation on PUN series has been evaluated in two analysis: taking into account the value of solar irradiance and using a parameter that quantifies the effective radiation of a given day in comparison to the maximum radiation in the considered day. The former has been considered in the following as *radiation* influence, the latter as *weather* influence.

In both cases the results obtained showed no significant correlation and the same sentence can be deduced by the next correlation analysis whose results are reported in TABLE II.

In the correlation analysis has been investigated another relevant aspect: the influence of the degree of the polynomial that relates PUN to the different variables. The models presented so far consider a linear relationship between PUN and Brent. However, the historic analysis of series could not exclude that the correlation can be nonlinear.

For this reason, the relationship between PUN, Brent and radiation is analyzed also with degrees of polynomial greater than one. The index of correlation of Paerson is utilized to perform this analysis. This index, under the hypothesis of normality of analyzed historical series, compares the covariance of two series with the standard deviation of each series, as summarized in (21).

$$\rho_{xy} = \frac{\sigma_{xy}}{\delta_x \delta_y} \tag{21}$$

The correlation between PUN and independent variables is visible in Table II. As shown, the index decreases for Radiation and Weather greater than one and grows in the relation between PUN and Brent. It is possible to state that it could be useful increase just the exponent of Brent and so the influence of irradiance and weather can be neglected.

It is important to know the exponent of Brent so that the best value of interpolation is found without burdening the model. For this reason, different models with growing exponent are analyzed. To compare the models, the Akaike Information Criterion (AIC) is utilized. In fact, this criterion considers three parameters: number of observations, number of independent variables and quadratic error between the real and the approximated series. Moreover, this criterion follows the rule of “smaller is better”, or rather the best model is the one that has the smaller AIC value.

TABLE II. CORRELATION INDEX OF PAERSON FOR RADIATION, WEATHER AND BRENT

Variables	PUN
Radiation	-0,0449
Radiation^2	-0,0322
Radiation^3	-0,0218
Weather	0,0165
Weather^2	0,0234
Weather^3	0,0266
Brent	0,2811
Brent^2	0,3279
Brent^3	0,3531
PUN	1

TABLE III. AKAIKE INFORMATION CRITERION FOR BRENT FROM FIRST TO TWENTIETH EXPONENT

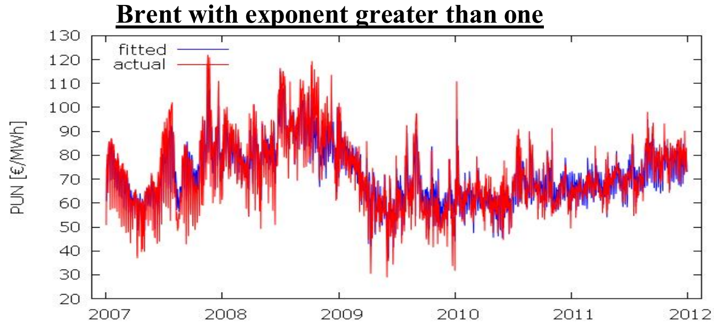
Maximum Exponent of Brent	AIC
1	10006,46
2	9990,12
3	9985,85
4	9986,10
5	9984,92
6	9985,03
7	9984,96
8	9985,79
9	9987,51
10	9987,75
11	9987,71
12	9989,71
13	9991,65
14	9993,65
15	9995,52
16	9997,38
17	10006,48
18	10075,02
19	10126,24
20	10118,78

The used series are explained in (22), while the results of AIC are shown in Table III.

$$y(t) = C_1 + \sum_{i=1}^{n} C_{2i}PB(t)^i + C_3 t + \sum_{i=1}^{7} C_{4i}d_i + \sum_{i=1}^{12} C_{5i}S_i + \mathrm{Ly(t-1)} \qquad \underline{\textbf{(22)}}$$

n represents the exponent of Brent to which the Brent series is truncated. As shown in Figure, the smaller value of AIC is for the approximated series with exponents from one to five. The model leads to the results in Figure 13.

Figure 13. Fitted and actual PUN with ARMA with regressor, trend, weekly and seasonal dummies, Brent with exponent greater than one



The results of this model are better than those obtained from the previous one: the RMSE is 7.8802 €/MWh, 11% of the mean value of PUN and this RMSE is less than the previous case of 0.15 €/MWh. Another index to evaluate the accuracy and to confront different models can be used. It is R^2 and measures the ratio between the variance of the regression and the variance of time series to be analyzed. It can assume values between zero and one, where zero represents complete dispersion of data and one the perfect reproduction of the series. For the model proposed in this section the R^2 is 70.65%.

G. GARCH model

From the analysis of the distribution of historical series of PUN, it was possible to deduce the presence of volatility clustering and leptokurtosis. As known, these two proprieties reveal that linear and autoregressive models present significant errors. It is possible to use different models for the interpolation and the extrapolation of PUN. In particular, the leptokurtosis assumes the propriety of heteroschedasticity of historical series PUN.

It is possible to test the presence of an ARCH of $\boldsymbol{q}$ in this following way (White, 1980).

- Estimating the model under consideration by Ordinary Least Square and save the squared residuals;
- performing and auxiliary regression where the squared residuals are regressed on a constant and on q retards;
- determining the TR2 (sample size multiplied by R^2) for the auxiliary regression;
- comparing the TR2 value with the distribution χ^2 with q degrees of freedom and, if the p-value is "enough small", refuse the null hypothesis of omoschedasticity, in favor of the alternative hypothesis of the existence of an ARCH(q) process.

This test is implemented and if the TR2 value of the auxiliary regression has a p-value less of 0,10. By means of the White test, the previous model is analyzed in order to find the eteroschedasticity. As shown in Table IV, there is convergence of the model with the first value of p-value.

TABLE IV. WHITE CRITERION FOR ETHEROSCHEDASTICITY

Delays	*P-value*
Alpha(0)	$2{,}88 \cdot 10^{-15}$
Alpha(1)	$2{,}51 \cdot 10^{-5}$
Alpha(2)	0,1543
Alpha(3)	0,0888
Alpha(4)	0,2110
Alpha(5)	$9{,}92 \cdot 10^{-10}$
Alpha(6)	0,08151
Alpha(7)	0,0221

The results of White test confirm the heteroschedasticity. The complete equation for the GARCH model is shown in (23). In this model the historical series of Brent is limited to the third exponent to meet convergence since if the exponent of Brent increases the software could have problem of convergence.

To complete the analysis, it is important to evaluate the values of $\boldsymbol{q}$ and $\boldsymbol{p}$ to find the best model. As the AR models, AIC is used with an empirical method to determine the values of $\boldsymbol{p}$ and $\boldsymbol{q}$ for the best model. For $\boldsymbol{q}$=2, there is no convergence. The minimum value of AIC is found with $\boldsymbol{p}$ and $\boldsymbol{q}$ equal to 1, as shown in Table V. The complete result of this model is shown in Figure 14.

$$y(t) = C_1 + \sum_{i=1}^{3} B_i\, PB(t)^i + C_3 t + \sum_{i=1}^{7} C_{4i} d_i + \sum_{i=1}^{12} C_{5i} S_i + C_6 * IRR(t) + C_7 \sum_{i=1}^{n} C_{8i} * \varepsilon(t-i)^2 + \sum_{i=1}^{n} C_{8i} \sigma(t-i)^2 \quad \textbf{(23)}$$

Figure 14. Fitted and actual PUN with GARCH Model

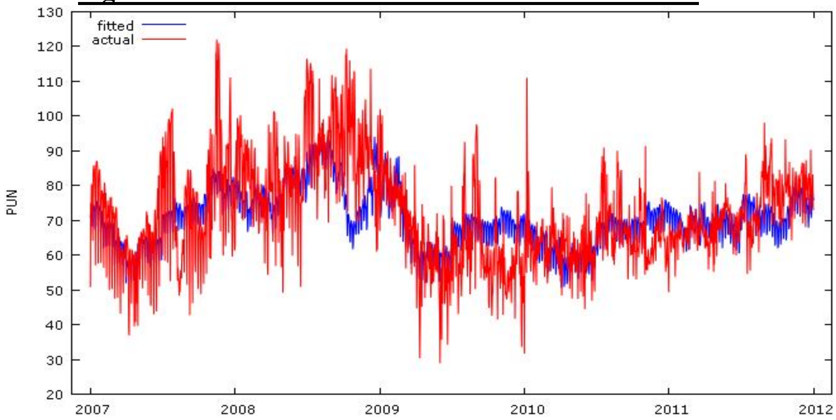


TABLE V. AIC TO CHOOSE THE BEST P AND Q

	q = 1	q = 2
p = 1	10456,14	No convergence
p = 2	10458,91	No convergence

5. Choosing the Best Model for the Interpretation

Table VI summarizes the models and sets; for each model, the values of the mean square and R^2 are listed. As can be seen from Table VI, the Autoregressive model with cost, trend, weekly dummies, season dummies and Brent with exponent from one to five, is the most reliable for modeling the series PUN. From the obtained results, it is possible to state that models having autoregressive components are much more reliable than linear models.

Table VI. R^2 and REQM for Analyzed Models

MODELS	R^2	*REQM*
Cost, Trend	2.80%	14.345 €/MWh
Cost, Regressor	7.89%	13.394 €/MWh
Cost, Regressor, Trend, Weekly Dummies	20.94%	13.398 €/MWh
AR with Cost, Regressor, Trend, Weekly Dummies	*68,75%*	*8.1317 €/MWh*
AR with Cost, Regressor, Trend, Weekly and Seasonal Dummies	69.56%	8.0246 €/MWh
AR with Cost, Trend, Weekly and Season Dummies, Brent with Exponent More than One	70.65%	12.6666 €/MWh
GARCH Model		11,154 €/MWh

6. Choosing the best Model for the Prevision

This conclusion does not establish, however, that the ARMA model above is also the best model for predicting the Single National Price of daily electricity. The ability to predict the smallest margin of error has still to be tested. For this purpose, an experimental method is used. Knowing the functions and the coefficients of the models, it is possible to predict the price of energy in the first ten months of 2012 and to compare it to the actual value. For this comparison, the ARMA model with linear, quadratic and cubic Brent and the GARCH model are related to the forecast prices of early 2012. These results are shown in Table VI.

To choose the best model for predicting the price estimators based on ARMA and GARCH, standard deviation of the real series is considered. In particular, for a comprehensive assessment of the two models: the sum of absolute errors, the sum of the modules of absolute errors, the minimum error, the maximum error of the RMSE and the average daily error are used.

The sum of absolute error is shown in Eq. (24)

$$SEA = \sum_{i=1}^{n} y(t) - \hat{y}(t) \qquad \underline{\textbf{(24)}}$$

The sum of modules of absolute errors is shown in Eq. (25)

$$SMEA = \sum_{i=1}^{n} |y_i(t) - \hat{y}(t)| \quad \text{(25)}$$

For minimal and maximal error, as in Eq. 26:

$$Em = MIN(|y_i(t) - \hat{y}(t)|)$$
$$EM = MAX(|y_i(t) - \hat{y}(t)|) \quad \text{(26)}$$

The daily average error has the form shown in Eq. (27)

$$EG = \frac{\sum_{i=1}^{n} |y_i(t) - \hat{y}(t)|}{n} \quad \text{(27)}$$

$y(t)$ and $\hat{y}(t)$ are, respectively, the actual and the forecast value of PUN at time t. $\bar{y}$ is the medium value of actual PUN and n is the number of observation.

The absolute errors are expressed as a percentage of the total sum of the values of the PUN in the considered period. The minimum and maximum errors are expressed as a percentage of the value of real PUN during that day. The average daily percentage error and the percentage of RMSE are obtained by the respective absolute values divided by the mean value of the real time series.

Table VI shows the results in the two cases considered. All estimators present a lower value in the case of GARCH respect to the ARMA model. The GARCH model presents the lower minimum and maximum error, demonstrating that the GARCH model performs better even in the case of a punctual estimate the price of electricity.

TABLE VII. R^2 AND REQM FOR ANALYZED MODELS

	SUM OF ABSOLUTE ERRORS	
ARMA	443,239 €/MWH	1,82%
GARCH	318,501 €/MWH	1,34%
	SUM OF MODULES OF ABSOLUTE ERRORS	
ARMA	2929,271 €/MWH	12,39%
GARCH	2592,671 €/MWH	10,97%
	MINIMUM ERROR	
ARMA	0,03 €/MWH	0,05%
GARCH	0,012 €/MWH	0,015%
	MAXIMUM ERROR	
ARMA	55,046 €/MWH	67,39%
GARCH	54,586 €/MWH	39,93%
	ROOT MEAN SQUARE ERROR	
ARMA	12,953 €/MWH	16,72%
GARCH	11,692 €/MWH	15,09%
	DAILY MEDIUM ERROR	
ARMA	9,604 €/MWH	12,40%
GARCH	8,501 €/MWH	10,97%

The results also show that in the proposed GARCH model, during the forecast period, the average daily produces an error of 8.5 €/MWh, which corresponds, in relative values, to the 10,97% of the average daily price over the forecast period considered. This value is 1.6 percentage points higher than the ARMA model and 5 points greater than the forecast model of the series of the PUN of first order least squares, which is the simplest model of regression used in this work.

7. Conclusions

The paper analyzed different models to assess and forecast the PUN in order to reduce the risk associated to price volatility. Some ARMA and GARCH models are implemented and compared to ensure an estimation of the more accurate models. The hypotheses of consider the PUN a simple random walk with drift and a normal series have been rejected respectively by means of the Dickey-Fuller and Jarque-Bera tests.

First of all an ordinary least square model with trend has been implemented and evaluated. Then the model has been improved inserting a regressor and weekly dummies. The method of Box-

Jenkins also showed the possibility of using an AR model which order has been chosen comparing the models with Akaike Information Criterion. AR models have been implemented also with seasonal dummies and polynomial relation with brent. Also the influences of other variables such as solar radiation and weather condition have been evaluated with low correlation results.

A GARCH model has been used after the performing of the White test to confirm eteroschedasticity of the residuals. A comparison based on the RMSE on the regression of the models has been conducted. Then also the forecasting performance of the models has been compared by means of error indices. These results lead to the conclusion that, while the ARMA is a better model for the regression of the time series of PUN, the GARCH model is a better estimator to predict the PUN. On the one hand, given the structure of ARCH, with coefficients following a trend, cyclical or based on error propagation, the model is useful especially for short term predictions. On the other hand, the series features the characteristics of randomness, stochasticity and volatility and therefore it is senseless to even suggest a model that can find the price of energy with great precision in long time. In this case, it may be more convenient to assume a monthly, or even quarterly, model, but remembering the fact that random variables in the period profoundly affect or alter the results.

References

Bohn, R., Caramanis, M., Schweppe, F. (1984), Optimal Pricing in Electrical Networks Over Space and Time, Rand Journal of Economics, 15(3), 360-376.

Bosco, B., Parisio, L., Pelagatti, M. (2007), Deregulated wholesale electricity prices in Italy: an empirical analysis, International Advance Economic Research, 13, 415–432.

Bowden, N., Payne, J.E. (2008), Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models, Energy Economics, 30(6), 3186-3197.

Braun, H., Lai, L. (2006), Insurance issues for energy risk, IEEE Power Engineering Society General Meeting

Contreras, J., Espinola, R., Nogales, F.J., Conejo, A.J. (2003), ARIMA models to predict next-day electricity prices, IEEE Transactions on Power Systems, 18(3), 1014–1020.

Garcia, R.C., Contreras, J., van Akkeren, M., Garcia, J.B.C. (2005), A GARCH forecasting model to predict day-ahead electricity prices, IEEE Transactions on Power Systems, 20(2), 867–874.

Gestore dei Mercati Energetici (GME) (Italian National Board for Energy markets) www.mercatoelettrico.org

Gianfreda, A. (2010), Volume and volatility effects in European electricity markets. Economic Notes, 39, 47–63.

Gianfreda, A., Grossi, L. (2012), Forecasting Italian electricity zonal prices with exogenous variables, Energy Economics, 34(6), 2228–2239.

Gianfreda, A., Grossi, L., Olivieri, D. (2010), Volatility structures of the Italian electricity market: An analysis of leverage and volume effects, 7th International Conference on the European Energy Market (EEM)

Girish G.P.M, Vijayalakshmi, S., Ajaya K.P., Badri N.R. (2014), Forecasting Electricity Prices in Deregulated Wholesale Spot Electricity Market: A Review, International Journal of Energy Economics and Policy, 4(1), 32-42.

Petrella, A., Sapio, S. (2009), A time series analysis of day-ahead prices on the Italian power exchange, Energy Market, EEM 2009, 6th International Conference on the European.

Rastegar, M.A., Guerci, E., Cincotti, S. (2009), Agent-based model of the Italian wholesale electricity market, IEEE, 6th International conference on the European electricity market.

Regnier, E. (2007), Oil and energy price volatility, Energy Economics, 29(3), 405–427.

Statistical Data on electricity in Italy – 2011, Terna.

Statistical Report 2011 Renewable-Energy Power Plants, Gestore Servizi Energetici (GSE).

U.S. Energy Information Administration (EIA), www.eia.gov

Wang, H., Xiang, F. (2011), Volatility characteristics of electricity price based on the normal distribution probability density function, International Conference on Electronics, Communications and Control (ICECC), September 23-25, Chengdu (RPC).

White, H. (1980), A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, Econometrica 48(4), 817–838.