Structural Analysis, Modeling and Forecasting of Electricity Prices of the Iberian Electricity Market

Jenice Ramos[1], Ângela Ferreira[2] and Paula O. Fernandes[3]

[1] Polytechnic Institute of Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Polytechnic Institute of Bragança, Bragança, Portugal
[3] Applied Management Research Unit (UNIAG), Polytechnic Institute of Bragança, Bragança, Portugal
ramosgonjenice@gmail.com, apf@ipb.pt, pof@ipb.pt

Abstract. The Iberian Electricity Market (MIBEL) resulted from a cooperation process developed by the Portuguese and Spanish administrations aiming to promote the integration of the electrical systems of both countries. With the liberalization of the electricity markets, price forecasting has become fundamental to the process of decision-making and strategy development by market participants. The unique characteristics of electricity prices such as non-stationarity, non-linearity and high volatility make this task very difficult. For this reason, instead of a simple timely forecast, market participants are more interested in a causal forecast that is essential to estimate the uncertainty involved in the price. This work analyses the impact of external variables on energy prices such as Per Capita Consumption, Heating Degrees-day, Cooling Degrees-day, Hydroelectric Productivity Index and Industrial Productivity Index, using a Multiple Linear Regression Model. From the models' application, it was observed adjusted coefficients of determination for the energy prices in Portugal of, approximately, 54% and 28%, for 2017 and 2018 years, respectively. Under the same time frame, adjusted coefficients of determination for the Spanish energy market are 52%, for 2017 year and 29%, for 2018.

Keywords: Energy Pricing Forecast, MIBEL, Multiple Linear Regression Model.

1 Introduction

The Iberian Electricity Market-MIBEL outcomes from a cooperative process developed by the Portuguese and Spanish governments aiming at promoting the integration of the electrical systems and markets of both countries. The MIBEL is organized in two poles, the Spanish Iberian Market Operator (OMIE), which provides the contracting of the Daily and Intraday Market and the Portuguese Market Operator (OMIP), which ensures the derivative markets.

With the MIBEL' implementation, the Iberian electricity market was moved to a liberalized market regime, being also an important step in the consolidation of the European Electricity Market. In this sense, it became possible for any Iberian consumer to acquire electricity from any producer or marketer operating in Portugal or Spain, under a regime of free competition [1].

The market price is established through a process in which the electricity price is the lowest one that guarantees the satisfaction of the demand by the supply [3]. Due to the its liberalized nature, electricity prices acquire volatile and uncertain characteristics, since they are obtained through proposals of supply and purchase of energy. In this competitive environment, it is imperative to predict the future price of energy, aiming the definition of a dispatch strategy, increasing the profit of energy producers and assisting a decrease in the electricity price for consumers.

The main objective of this work is the construction of statistical (or casual) models to forecast energy prices, in a monthly basis, in the time span of 2017 and 2018 years, through the Multiple Linear Regression Model (MRLM).

The paper is organized as follows: session 2 presents the main factors which may contribute to the variability of energy prices; session 3 introduces and discusses the forecasting methodology, while session 4 presents and discusses its application in the Iberian countries. Finally, session 5 draws the main conclusions and outlines the future work.

2 Key Factors Affecting Electricity Prices

One of the most notorious characteristics of energy prices is their extreme volatility. This means that, throughout the year, there may be sharp price changes. Therefore, it is necessary to analyse the variables that are able to explain, even though partially, the variability of prices.

Unique features of energy pricing such as non-stationarity, non-linearity and high volatility make the forecast of energy prices task difficult. For this reason, instead of a simple one-off forecast, market players are more interested in a causal forecast able to estimate the uncertainty involved in price.

From a large number of external variables able to affect the electric energy prices, the ones that have demonstrated a higher correlation with the electric energy price are Consumption Per Capita (CPC), Heating and Cooling Degrees-days (HDD and CDD, respectively), Hydroelectric Productivity Index (HPI) and Industrial Productivity Index (IPI).

Sudden changes in electric energy consumption can lead to spikes in energy prices. The energy demand is interrelated with meteorological conditions, for instance heating and cooling requirements, here accessed through technical indexes based on weather conditions, HDD and CDD variables [4], which describe the requirements of the energy demand for heating and cooling (air conditioning) of buildings. These variables are derived from meteorological observations of the air temperature and interpolated in regular networks with a resolution of 25 km in Europe. These variables present a complementary characteristic throughout the year, that is, they are a quantification of the degrees Celsius required for heating in the winter months, and cooling in the summer months.

The availability of the hydric resource, due to its high penetration in the Iberian electricity market also impacts in the energy prices. The Hydroelectric Productivity Index (HPI) [5], reckons the deviation of the total amount of electric produced from hydric resources in a given period, related to the one which would occur if an average hydrological regime occurred. The latter is evaluated taking into account 30 historical hydrological regimes. If HPI is higher than 1, the period under analysis is considered wet, and if HPI is lower than 1, from the hydrological point of view, it is considered dry.

The Industrial Productivity Index (IPI) [6] measures changes in the volume of production of goods at short and regular intervals, relative to a period taken as a reference. Under the assumption of stability of technical coefficients, this index also measures the trend of value added in volume. Doing so, its relation to the energy demand also affects the energy price.

3 Forecasting Research Methodology

Numerous methods of forecasting energy prices have been proposed over the last years. Typically, these models fall into three types of time horizon, short term (hours to days), medium term (months) or long term (years). However, there is no consensus on the limits of each models.

From the several developed models of forecasting, multi agent simulation models, statistic and Artificial Intelligence based models stand out. It is also noteworthy the growing use of hybrid models, combining those methodologies [3].

The present work is based on the use of statistical models. These models help to produce forecasts for a dependent variable, which in the case under study, is the variable energy price, as a combination of independent variables that may have influence on the energy price. In this context, the model chosen for energy pricing is the Multiple Linear Regression Model (MLRM).

3.1 Multiple Linear Regression Model

The MLRM is a statistical model that assumes there is a linear relationship between a variable *Y* (the dependent variable) and *X* independent variables. The independent variables exogenous, explanatory, non-stochastic and observable, used to explain the variation of the variable *Y*. A model that comprises more than one independent variable is a multiple regression between a dependent variable and a set of n+1 independent variables assuming a linear form and stochastic because it includes an error term [7]. The Multiple Linear Regression Model is given by [8]:

$$\begin{aligned} Y_t &= b_0 + b_1 X_{1t} + b_2 X_{2t} + b_j X_{jt} + \ldots + b_k X_{kn} + u_t \\ t &= 1, 2, \ldots, n, \; j = 1, 2, \ldots, k \end{aligned} \tag{1}$$

where b_0 is the y-intercept, b_j represents the parameters of the model, and u_t is the error term.

A casual association is not assumed between dependent and independent variables. In this sense, the dependent variable, Y, depends on a set of n+1 known factors and an unknown factor, being an endogenous variable, explained, stochastic or random and observable [8].

Typically, the linear regression model uses the following assumptions [12]:

- The regression mode is linear as proposed in (1);
- The regressors are assumed to be fixed or non stochastic in the sense that their values are fixed in repeated sampling;
- Given the values of the independent variables, the expected value of the error term is zero;
- The variance of each error term, given the values of independent variables, is constant or homoscedastic;
- There is no correlation between two error terms, i.e., there is no autocorrelation;
- There are no perfect linear relationships among the dependent variables, i.e., there is no multicollinearity;
- The regression model is correctly specified.

Based on the assumptions mentioned above, the most popular method for parameters estimation, the Ordinary Least Squares (OLS), provides estimators which have several desirable statistical properties, such as [13]:

- The estimators are linear, which means that they are linear functions of the dependent variable, *Y*;
- The estimators are unbiased, which means that, in repeated applications of the method, on average, they are equal to their true values;
- The estimators are efficient, which means that they have minimum variance.

3.2 Measures of Forecasting Accuracy

The main purpose of the whole modelling and forecasting process is to clearly discern the future values of the dependent variable, and the most important criterion of all is how accurately a model does this. The most familiar concept of forecasting accuracy is evaluated through the error magnitude accuracy, e_t, which relates to forecast error with a particular forecasting model. This is defined as [9]:

$$e_t = A_t - F_t \tag{2}$$

being A_t the actual value and F_t its forecast in the time period, t.

Although there are a number of forecasting errors that can be used for accurate evaluation, in this work it is used the Mean Absolute Percentage Error (MAPE), expressed in generic percentage terms and it is computed by [12]:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{3}$$

4 Electric Energy Price Modelling and Forecasting

4.1 Data sample and generic model

The modelling methodology used the historical data from January 2010 till December 2015, in a total of 72 observations. The year 2016 was used to validate the model, and the years 2017 and 2018 used to produce the forecasts and to build the models, based on the previous validation for 2016, already working with 84 observations (January 2010 till December 2016).

To produce all the results, it was used GRETL statistical software (Gnu Regression, Econometrics and Time-series Library) for Windows. It should be noted that the data of the variables presented above were collected from the Eurostat [4], REE [14], REN [15] , INE of Spain [10] and INE of Portugal [10] databases.

The model to be created is no more than a representation of the relations between the variables at the same moment of time according to (1). Energy Prices (EP) modelling and forecasting, for the Portuguese and Spanish markets, used the following econometric model:

$$EP_t = a + b_0 CPC_t + b_1 HDD_t + + b_2 CDD_t + + b_3 GDA_t + \\ + b_4 HPI_t + b_5 IPI_t + e_t \quad (4)$$

It should be noted models of Portugal and Spain interrelate the electric energy price with explanatory variables for each country.

4.2 Energy Prices Modeling for Portugal

The results obtained for Portugal with the Multiple Linear Regression Model, estimated by the application of the Ordinary Least Squares Method for 2017 are presented in Table 1.

Table 1. Performance Measures of the Estimated Model for Portugal for 2017.

	Coefficient	Error	ratio-t	*p-value*	Significance	VIF
Const	−25,2669	23,1584	−1,091	0,2786		
IPIP	−0,191297	0,0853043	−2,243	0,0278	**	1,436
CPCP	0,122396	0,0326736	3,746	0,0003	***	3,802
HDDP	−0,0870782	0,0188241	−4,626	1,46E-05	***	5,758
CDDP	−0,0900926	0,0486428	−1,852	0,0678	*	2,69
HPIP	−15,1130	1,67964	−8,998	1,09E-13	***	1,022
Mean var. dependent		*44,40357*	***D.S. var. dependent***		*10,56537*	
White Test (TR²)		*21,778043*	***Durbin-Watson***		*1,041234*	
R²		*0,565662*	***R² adjusted***		*0,53782*	
F (5, 78)		*20,31672*	***p-value (F)***		*6,31E-13*	

*Notes: *, Significance of 10%; **, Significance of 5%; ***, Significance of 1%.*

The obtained coefficient of determination is 0.56562, which indicates that the variables Per Capita Consumption, Heating/Cooling Degree Days, Hydroelectric Productivity Index and Industrial Productivity Index and main variables markets, explained about 57% of the variations that occurred in the electric energy prices in Portugal. The adjusted coefficient of determination is 0.53 which indicates that about 53% of the changes in electricity prices were explained by the variations in the independent variables.

From obtained results shown above, in Table 1, it is possible to conclude:

- The autonomous component indicates that -25.2669 of the electricity prices for Portugal are not explained by independent variables. However, this variable does not reveal a statistically significant value.
- If the variables Industrial Productivity Index, Heating/Cooling Degree Days and the HPIP varies by one unit, the Energy Price variable decreases, and it is evident that they have a negative relation between them. All these variables are statistically significant.
- The variable Per Capita Consumption Portugal, has a positive relation with the Energy Price: if the first one varies one unit, the Portuguese Price variable increases by approximately 0.122396 units. This variable is statistically significant, with a significance level of 1%.
- Regarding the *F* statistic (5.78), there is sufficient statistical evidence to verify that there are variables that assume values other than zero and, as previously mentioned, the variables included in the model explain satisfactorily the changes in electricity prices in Portugal.
- From the analysis of the violation of the basic hypotheses of the model, in terms of multicollinearity and based on the values of the Variance Inflation Factor (VIF), there is no violation of the basic hypothesis of multicollinearity, since the VIF values, for all variables, are not higher than 10. It can be concluded that there is no dependence on explanatory variables.
- The test of the normality of the residue made through the test statistic (χ^2 = 0.579, with test value = 0.7486, which means that this model follows a normal distribution at a significance level of 1%, so this hypothesis is not violated. the mean is equal to μ = -4,22946e-0177,1827; this value is approximately zero then the zero-mean hypothesis is also not violated E (μ) = 0;
- For homoscedasticity, a constant variance of the error term was verified by the White test for heteroskedasticity and the test statistic TR^2= 21.778043 with test value (χ^2 (20)> 21.778043) = 0.35675. As the test value is higher than 10%, it can be concluded that there is no violation of homoscedasticity, that is, the variance is constant observation for observation. There is no loss of the characteristics of OLS estimators, since they remain BLUE;
- The Durbin-Watson statistic = 1.041234 lies in the zone of positive autocorrelation of the errors. Then, it can be concluded that there is an infringement of the independence of the error term and that this model suffers from autocorrelation of the errors. In order to correct the infraction hypothesis, the Cochrane-Orcutt test was applied.

Accordingly, the following statistic was obtained: Durbin-Watson = 2.025485, which is now in the zone of independence of the errors.

In order, to be able to model and predict prices for 2018, it was necessary to create a trend line from the price of electricity for Portugal and create 12 dummies (dm) or Periodic Auxiliary Variables that represent each of the months of the year of 2018. Subsequently, the least squares method is applied. These auxiliary variables were created as aids to the model, due to the absence of data from the independent variables referring to the year 2018, from May 2018.

Table 2. Performance Measures of the Model with Periodic Auxiliary Variables for 2018.

	Coefficient	**Error**	**ratio-t**	***p-value***	**Significance**
Time	0,05542	0,033433	1,658	0,1012	
dm1	42,6982	3,49385	12,22	3,03E-20	***
dm2	36,4877	3,50774	10,4	1,05E-16	***
dm3	32,5611	3,52189	9,245	2,11E-14	***
dm4	31,5256	3,53631	8,915	9,67E-14	***
dm5	39,099	3,55097	11,01	6,64E-18	***
dm6	44,9773	3,56589	12,61	5,45E-21	***
dm7	46,4994	3,58106	12,98	1,09E-21	***
dm8	45,9265	3,59648	12,77	2,76E-21	***
dm9	48,3898	3,61214	13,4	1,87E-22	***
dm10	49,1344	3,62804	13,54	1,00E-22	***
dm11	45,7027	3,64418	12,54	7,45E-21	***
dm12	49,2064	3,49941	14,06	1,13E-23	***
Mean var. dependent	*45,4093*		***D.P. var. dependent***		*10,589*
White Test (TR^2)	*35,935126*		***Durbin-Watson***		*0,6786*
R^2	*0,3679*		***R^2adjusted***		*0,2765*
F (12, 83)	*4,0258*		***value P(F)***		*0,0001*

It is also necessary to verify that the model for 2018 does not violate the infractions in order to be able to validate it. Based on the information presented in the previous table, the following can be concluded regarding the violation of the model hypotheses for 2018:

- All auxiliary variables are statistically significant at a significance level of 1%.
- Based on the Inflation Factor of the variance, it is verified that there is no violation of the basic hypothesis of multicollinearity.
- The test of normality of the residue made through the test statistic χ^2= 1.077, with test value = 0.58369, means that this model follows a normal distribution at a significance level of 1%, so this hypothesis is not violated.

- Constant variance of the error term, by White test for heteroscedasticity, as the value of evidence higher than 10%, from which, there is no violation of homoscedasticity.
- The Durbin Watson statistic = 0.6778602 was found in the zone of positive autocorrelation of the errors, meaning that there is an infraction to the independence of the error term. To overcome the previously verified infraction, the Cochrane-Orcutt test was applied. Accordingly, the following statistic was obtained by Durbin-Watson = 1.994749, which translates in independence of the errors.

4.3 Energy Prices Modelling for Spain

Following the same methodology as described in previous section, the model obtained for Spain, in 2017 (Table 3), presents a coefficient of determination of 0.550964 and indicates that the variables Per Capita Consumption, Heating/Cooling Degree Days, Hydroelectric Productivity Index and Industrial Productivity Index explain 55.1 % changes in energy prices in Spain in 2017. The adjusted coefficient of determination is 0.52, which indicates that about 52% of the changes in energy prices in Spain are explained by the independent variables.

Table 3. Performance Measures of the Estimated Model for Spain for 2017.

	Coefficient	Error	ratio-t	*p-value*	Significance	VIF
Const	14,346	22,6092	0,6345	0,5276		
IPIE	−0,178757	0,112919	−1,583	0,1175		1,621
CPCE	0,074336	0,03064	2,426	0,0176	**	4,036
HDDE	−0,0422069	0,015252	−2,767	0,0071	***	6,562
CDDE	−0,0834428	0,057153	−1,460	0,1483		6,493
HPIE	−13,7937	1,65802	−8,319	2,26E-12	***	1,026
Mean var. dependent	*44,30881*	***D.S. var. dependent***		*10,18166*		
White Test (TR^2)	*27,338403*	***Durbin-Watson***		*0,859692*		
R^2	*0,550964*	***R^2 adjusted***		*0,522179*		
F (5, 78)	*19,14107*	***value***		*2,23E-12*		

Note: *, Significance of 10%; **, Significance of 5%; ***, Significance of 1%.

Based on the results obtained and presented in the table above, it can be concluded that:

- The autonomous component shows that 14,346 of electricity prices in Spain are not explained by the independent variables. This variable is not a statistically significant variable.
- There is an inverse relationship between variables Industrial Productivity Index, Heating/Cooling Degree Days, Hydroelectric Productivity Index and energy price in Spain. Only variables Consumption per Capita, Heating Degree Days and Hydroelectric Productivity Index are statistically significant.

- Per Capita Consumption has a positive relation with the energy price, being statistically significant, at a significance level of 5%.
- As for the statistic of F (5,78) = 19,14107, with a test value lower than 1%, there is sufficient statistical evidence that there are variables that assume values different from zero and as previously mentioned, the variables included in the model explain in a satisfactory way the variations occurred in the prices of electric energy in Spain.
- Regarding the analysis of the infraction to the basic hypotheses of the model, considering the VIF, it is verified that there is no infringement of the basic hypothesis of multicollinearity, since values for any of the variables are less than 10. There is no correlation between the explanatory variables.
- The test of normality of the residue performed through the test statistic χ^2= 0.228, with test value = 0.8925, means that this model follows a normal distribution at a significance level of 1%, so this hypothesis not is not violated. The mean value is approximately zero, so the zero-mean hypothesis is also not violated E (μ) = 0.
- Constant variance of the error term, through the White test for heteroskedasticity and the test statistic, is higher than 10%, so it can be concluded that there is no violation of homoscedasticity.

Table 4. Performance Measures of the Model with Periodic Auxiliary Variables for 2018.

	Coefficient	**Error**	**ratio-t**	***p-value***	**Significance**
Time	0,061663	0,031961	1,929	0,0571	*
dm1	42,9797	3,34004	12,87	1,80E-21	***
dm2	36,4318	3,35332	10,86	1,29E-17	***
dm3	32,3001	3,36685	9,594	4,25E-15	***
dm4	31,1422	3,38063	9,212	2,46E-14	***
dm5	38,8043	3,39465	11,43	1,01E-18	***
dm6	44,4514	3,40891	13,04	8,60E-22	***
dm7	45,991	3,42341	13,43	1,59E-22	***
dm8	45,4193	3,43815	13,21	4,14E-22	***
dm9	47,6314	3,45312	13,79	3,47E-23	***
dm10	48,2735	3,46833	13,92	2,06E-23	***
dm11	45,0456	3,48375	12,93	1,38E-21	***
dm12	49,2064	3,49941	14,06	1,13E-23	***
Mean var. dependent		*45,29708*	***D.S. var. dependent***		*10,24108*
White test (TR^2)		*36,087775*	***Durbin-Watson***		*0,712452*
R^2		*0,382413*	***R^2 adjusted***		*0,293124*
F (12, 83)		*4,282844*	***Value P(F)***		*0,00003*

As in the case of Portugal, for 2018, it was necessary to create a trend line, based on the electricity price for Spain by using 12 dummies (dm) representing each of the months of the year.

- In the statistical tables for 5 independent variables dL is equal to 1.5219, dU equal to 1.7732, 4-dU equal to 2.2268 and finally 4-dL is equal to 2.48. It was obtained the following Durbin-Watson statistic = 0.858692, which lies in the zone of positive autocorrelation of the errors, meaning that there is an infringement of the independence of the term of error. Following the application of Cochrane-Orcutt test, a Durbin-Watson statistic = 1.996358 is obtained, which satisfies the independence of the errors.

From the information presented in Table 4, it can be verified that the model for 2018 does not violate the infractions, which validates it. All auxiliary variables are statistically significant with a significance level of 1%. Additionally,

- Regarding the analysis of multicollinearity, considering the VIF, it is verified that there is no violation of this hypothesis.
- The test of normality of the residue made through the test statistic = 0.645, with test value equal to 0.7243, which means that this model follows a normal distribution at a significance level of 1%, so this hypothesis is not violated.
- White test has a test value higher than 10%, so it can be concluded that there is no violation of homoscedasticity;
- The Durbin-Watson statistic = 0.712452 was obtained. This value is in the zone of positive autocorrelation of the errors, being necessary to analyse further, using the test of Cochrane-Orcutt to verify if that the infraction can be solved. With the Cochrane-Orcutt test the following Durbin-Watson statistic = 0.2019850 was obtained and, consequently, there is independence of the errors.

4.4 Forecast Results for Portugal and Spain

This section presents the forecasts for the electric energy price, for each of the countries under analysis integrating MIBEL, for the years 2017 and 2018, based on the models created and described in the previous sections.

To evaluate the accuracy of the prediction, it will be used the Absolute Percent Error (APE) and Mean Absolute Percent Error (MAPE). The assumed confidence interval to produce forecasts is 95%. The results obtained for the two models selected with the methodologies used and for the respective statistical measures/indicators are presented in Table 5 and Table 6, for Portugal and Spain, respectively.

Regarding Portugal and year 2017, it can be observed that the difference between the actual and expected value is € 1.65 and the MAPE is 12.61%. Forecasts for 2017 follow the behaviour of real historical prices. For 2018, and considering the known prices, that is, between January 2018 and May 2018, it is notorious that the forecast for 2018 follows the same behaviour. As for the MAPE, in this year, since it can only be calculated for 5 months, a MAPE of 11.52% has been obtained.

From the analysis of the data of average energy price for the Portuguese Market, considering the period of analysis from January 2017 to May 2018, it is verified that this indicates maximum values in the winter months, where variables such as CPC and HDD are higher which may justify the increase in prices. Extrapolating this analysis to the remaining periods, it is possible verify that the energy price registers low values for the summer months, where the CPC is lower. The minimum values in the energy price are in the months of March and April in both years under analysis. This decrease in price is justified when the months have a very high HPI, from which higher-cost energy sources can be withdrawn from service, thus contributing to the decrease of the energy price.

Table 5. Forecast of Energy Prices for Portugal for 2017 and 2018.

Months	2017			2018		
	Real Price, €/MWh	Forecast Price, €/ MWh	APE	Real Price, €/ MWh	Forecast Price, €/ MWh	APE
January	71,52	56,38	21,17%	49,98	53,16	6,36%
February	51,39	46,44	9,63%	54,88	44,76	18,44%
March	43,95	49,99	13,74%	40,18	39,44	1,84%
April	44,18	54,44	23,22%	42,67	37,52	12,07%
May	47,12	52,78	12,01%	54,92	44,54	18,90%
June	50,22	55,68	10,87%	-	50,08	-
July	48,6	58,19	19,73%	-	51,41	-
August	47,43	51,91	9,45%	-	50,73	-
September	49,16	54,37	10,60%	-	53,14	-
October	56,97	59,71	4,81%	-	53,88	-
November	59,36	54,41	8,34%	-	50,45	-
December	59,49	54,91	7,70%	-	54,43	-
Mean values	***52,45***	***54,10***	***12,61%***	***48,53***	***48,63***	***11,52%***

Table 6. Forecast of Electricity Prices for Spain for 2017 and 2018.

Months	2017			2018		
	Real Price, €/MWh	Forecast Price, €/MWh	APE	Real Price, €/ MWh	Forecast Price, €/ MWh	APE
January	71,49	60,97	14,72%	51,63	52,58	1,84%
February	51,74	44,33	14,32%	54,98	44,3	19,43%
March	43,19	51,83	20,00%	39,75	39,11	1,61%
April	43,69	52,54	20,26%	42,66	37,3	12,56%
May	47,11	53,91	14,43%	55,08	44,58	19,06%
June	50,22	56,99	13,48%	-	50,01	-
July	48,63	57,22	17,66%	-	51,43	-
August	47,46	58,22	22,67%	-	50,81	-
September	49,15	56,19	14,32%	-	53,02	-
October	56,77	57,43	1,16%	-	53,68	-
November	59,19	55,83	5,68%	-	50,48	-
December	57,94	51,61	10,93%	-	54,69	-
Mean Values	***52,22***	***54,76***	***14,14%***	***48,82 €***	***48,50 €***	***10,90%***

As regards electricity prices for Spain (Table 6), it is observed that the forecast values for 2017 are higher by € 2.54. The MAPE obtained for 2017 was 14%, higher than that of Portugal and for the year 2018, about 11%. Analysing the year 2017, it can be verified that the predictions follow the same behaviour of the original series, which allows trusting the model. With reference to the forecast of the mean energy price for the Spanish market, maximum values are also found in winter months, where variables such as CPC, HDD are higher. Similar to the results obtained for Portugal, it can be verified that the energy price registers low values in summer months, when the CPC is lower.

5 Conclusion

The objective of this work was the development of a model for forecasting energy prices in the Iberian Market. MIBEL's electricity prices show great volatility, with spikes, which limits the performance of the models.

The analysis of the relation of the electric energy price with the external variables allows to conclude that the same influence significantly the properties of the distribution of the price. The Hydroelectric Productivity Index is largely responsible for price volatility, with scenarios in which this renewable resource production is high, leading to a decrease in prices.

With respect to Per Capita Consumption, it reveals to be a crucial variable that partly justifies the behaviour of the energy price. Being intrinsically related to the consumption variable, HDD and CDD also have a considerable impact on price variability.

All the variables presented in the model for the Portuguese market are statistically significant variables. As far as the Spanish market is concerned, the variables are not all statistically significant: only the variables CPC, HDD and HPI are statistically significant.

From the analysis of the performance of the models created, the model for Portugal for the year 2017, presents better results than the model applied to Spain. Regarding the forecast models for the year 2018, the model created for Spain presents the best performance and the lowest MAPE. For the year 2018, the model that produced the best results was the model built for the Spanish market, although the difference is not significant.

The establishment of a reference model presents itself as an innovative idea, with the objective of understanding the inexplicability associated to the forecast models. However, as good as some forecasting models may be, there will always be a forecasting error associated with factors and causes that do not depend on model inputs and occur more frequently at certain times of the day, week, month or year. The modeling presented also reflects that factors that influence the Portuguese market may not be the same factors that influence the Spanish market, even they belong to the same energy market.

The quality of the estimated models validates the use of statistical or causal methods, such as the Multiple Linear Regression Model, as a plausible strategy to obtain causal forecasts of energy prices.

References

1. J. Borges, “MIBEL and the Market Splitting - 5th International Conference on the European electricity market,” in 2008 5th International Conference on the European Electricity Market, 2008.
2. Homepage OMIE, “Mercado Diário.” [Online]. Available: http://www.omie.es/pt/principal/mercados-e-produtos/mercado-da-electricidade/os-nossos-mercados-de-eletricidade/mercado-di. [Accessed: 05-Feb-2018].
3. R. F. da C. Moreira, “Previsão Probabilística dos Preços de Energia Elétrica do Mercado Ibérico de Eletricidade,” 2015.
4. Homepage EUROSTAT, “Cooling and heating degree days by country - monthly data.” [Online]. Available: http://ec.europa.eu/eurostat/web/products-datasets/product?code=nrg_chdd_m. [Accessed: 15-May-2018].
5. Homepage ERSE, “Índice de Produtibilidade Hidroeléctrica (IPH).” [Online]. Available:http://www.erse.pt/pt/glossario/Paginas/glossario.aspx?folder=23324d4f-8a61-4074-a170-6f8373f1827e&master=ErsePrint.master. [Accessed: 15-May-2018].
6. Homepage INEPT, “Índice de Produtividade Industrial” [Online]. Available: https://www.ine.pt/xpotal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009411&contexto=bd&selTab=tab2. [Accessed: 15-Jun-2018].
7. C. Cristina, E. Maciel, P. Guimaraes, and J. C. Ribeiro, Instrumentos estatísticos de apoio à economia. Lisboa, 2000.
8. N. S. M. dos Santos and P. O. Fernandes, “Modelação da procura turística em Portugal,” Rev. Tur. Desenvolv., 2012.
9. D. N. Gujarati and D. C. Porter: Essentials of Econometrics. McGraw Hill (2010).
10. Homepage INEES, “Índice de Produtividade Industrial.” [Online]. Available: http://www.ine.es/jaxiT3/Datos.htm?t=26061. [Accessed: 05-May-2018].
11. D. Gujarati. Econometrics by example. UK: Palgrave Macmillan. 2015.
12. H. White, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” Econometrica, 1980.
13. M. Verbeek. A Guide to Modern Econometrics. UK: A John Wiley & Sons. 2012.
14. Homepage Rede Elétrica Espanha, “National statistical series.” [Online]. Available: http://www.ree.es/en/statistical-data-of-spanish-electrical-system/statistical-series/national-statistical-series. [Accessed: 05-May-2018].
15. Homepage Rede Elétrica Nacional, “Estatística Mensal - SEN” [Online]. Available: http://www.centrodeinformcao.ren.pt/PT/InformacaoExploracao/Paginas/EstatisticaMensal.aspx. [Accessed: 05-Feb-2018].
16. Homepage INEES, “Indice de Produtividade Industrial.” [Online]. Available: http://www.ine.es/jaxiT3/Datos.htm?t=26061. [Accessed: 05-May-2018].