Day-ahead electricity market price forecasting using artificial neural network with spearman data correlation

João Nascimento
Energia Simples
Porto, Porugal
joao.nascimento@energiasimples.pt

Tiago Pinto, Zita Vale
Polytechnic of Porto (ISEP/IPP),
Porto, Portugal
{tmcfp, zav}@isep.ipp.pt

Abstract— **Electricity markets are complex environments with very dynamic characteristics. The large-scale penetration of renewable energy sources has brought an increased uncertainty to generation, which is consequently, reflected in electricity market prices. In this way, novel advanced forecasting methods that are able to predict electricity market prices taking into account the new variables that influence prices variation are required. This paper proposes a new model for day-ahead electricity market prices forecasting based on the application of an artificial neural network. The main novelty of this paper relates to the pre-processing phase, in which the relevant data referring to the different variables that have a direct influence on market prices such as generation, temperature, consumption, among others, is analysed. The association between these variables is performed using spearman correlation, from which results the identification of which data has a larger influence on the market prices variation. This pre-analysis is then used to adapt the training process of the artificial neural network, leading to improved forecasting results, by using the most relevant data in an appropriate way.**

Index Terms-- **artificial neural networks, day-ahead spot market, electricity price, forecasting, spearman correlation**

I. INTRODUCTION

The liberalization of the electric energy market has changed the paradigm of a sector whose organizational models were traditionally monopolistic, being mostly owned by the state [1]. This change has introduced the competitive factor in the sector, making it possible to reduce costs, improve quality and reliability of service, and even promote innovation in new products and services [2]. However, this also brought instability to the prices practiced and a greater risk to the market agents of this sector [3].

The characteristics of the post-liberalization electricity market make forecasting price developments very complex, but important for the protection of those involved, be they producers, traders or consumers. With this purpose, this paper studies energy prices in the spot market in the Iberian Electricity Market (MIBEL) [4]. Initially, a collection and treatment of historical data of electricity production by technology and consumption, such as meteorological data for Portugal and Spain is carried out. Afterwards, all the data

obtained were evaluated through the Spearman coefficient [5], allowing to understand the trend of the price variation of the energy in the MIBEL, compared to the other values. It was also carried out a study of the demonstrated seasonality, as much by the price of the electric power as by the load verified in the Portuguese and Spanish markets, since the load would be one of the data that most influenced the price variation. Therefore, was also constructed models for forecasting, in order to help reduce the error obtained for the price forecast.

Prediction models were developed using the computational intelligence-based method, Artificial Neural Networks (ANN) [6], developed in the nntool tool of Matlab®. The daily model, the one for which we have the most recent data, and a shorter time horizon, is the one for which the error tolerance is lower. This provides for the price of electricity for the next 24 hours, adapting to the spot market, more specifically the daily market. The weekly model, with a structure very similar to the daily one, forecasts the price for the next 168 hours, having a higher tolerance, given that the historical data is not so recent, and the time horizon gives rise to more errors. As for the monthly model developed, in order to not have a very high error (since if it were drawn similarly to the previous ones, it would have to predict about 720 hourly values), and due to the greater difficulty in obtaining accurate values for a time window (meteorological data for all the days of the following month, for example), it is designed to predict an average price for the following month, containing, for example, historical values of average price of the previous year. Due to the need for this type of information, the total data obtained, although comprehensively collected, would not be the ideal for minimally accurate long-term forecasts.

The performance evaluation is performed through the Mean Absolute Percentage Error (MAPE) [7] and Absolute Percentage Error (APE) [8], and it was concluded that the forecast models developed show a very acceptable accuracy. Nevertheless, to maximize the error reduction, a procedure was used to reduce the dispersion of the results and thus to ignore values that deviate from the mean, using the standard deviation of the values obtained.

II. PROPOSED METHODOLOY

A. Data correlation

In order to increase the forecast accuracy, a correlation analysis of the variables collected with the price of electricity is carried out. In this way, and when evaluating the degree of relationship between variables, it is possible to find out precisely how much one variable interferes with the result of another. The techniques associated with Correlation Analysis represent a fundamental application tool in the Social Sciences and Behavior, Engineering and Natural Sciences. There are several evaluation criteria for this relationship, some for variables that follow a normal distribution and others for variables that do not follow a known theoretical distribution. Several studies are based on Pearson's correlation coefficient [9]. However, there are situations in which the relationship between two variables is not linear, or one of them is not continuous, or the observations are randomly selected. In situations like these, other coefficient alternatives should be applied. The method chosen to study the correlation of the different variables present in the data collected in relation to the price was the Spearman Coefficient [5]. This method measures the binding force between two variables, being as great as the more monotonous the relative evolution of these variables are, reaching the maximum values of +1 or -1, if they are directly or inversely proportional, respectively. The correlation value indicative is represented by ρ, as in (1):

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{1}$$

The Spearman coefficient orders the values of the variables x_i and y_i separately, where i represents the value of stations 1, 2, 3, ..., n. Then the difference, d_i, of the positions of x_i and y_i, and the square of this value (d_i^2) is calculated. The number of samples is represented by n.

In the analysis of data collected from July 1, 2015 to June 30, 2017, corresponding to the Iberian Electricity Market (MIBEL) [4], seasonal patterns of variability were identified between them, whether daily, weekly or annual, although the selected data to develop the present forecasting model, will only use the second year of information (through July 1, 2016, to June 30, 2017). Daily patterns are easily identifiable when analysing how the price of electric energy varies throughout the day, being related to the human activity and, of course, to the consumption. This relationship is copied by load consumption, when comparing both for each of the horizons studied, be it for the relationship of these characteristics throughout the day or the week, mainly. This suggests that the load represents an important piece of data for the forecast of electricity prices. Fig.1 illustrates: a) the variation of electricity price with the irradiation; and b) Load Consumption and Deviation from the Average Value of Irradiation.

From Fig.1 a) it is noticeable that there is a negative correlation (inversely proportional evolution) for the data of both variables in the analysed temporal window. This is the data from July 1, 2015 until June 30, 2017, although the data used for the day-ahead forecast model was only July 1, 2016 through June 30, 2017, as they would be relatively short-term forecasts. This option was chosen since no older data would

be required given the apparent correlation between these variables. Although there is some correlation between electricity price and load consumption, and electricity price and irradiation, the correlation between load and irradiation did not have an apparent value. This led to the creation of a new variable: the deviation between irradiation and its own average value. This happens because it is on the seasons with peaks of radiation (be them positive or negative), that the value of load consumption is higher, see Fig. 1 b).
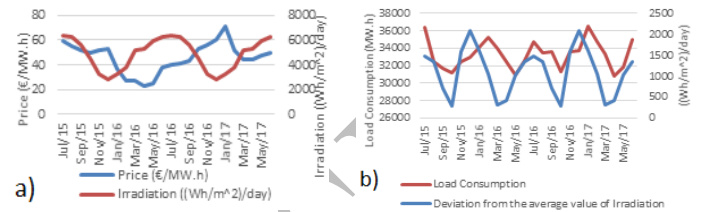

Fig. 1. Variation between: a) Electricity price and Irradiation; b) Load Consumption and Deviation from the Average Value of Irradiation

Using the Spearman Correlation method is possible to analyse the value of correlation between variables, with precision. The analysis is performed in relation to the price and the load consumption. Since in this data set there is a great diversity of variable types, they are separated according to their characteristics. Data refers to production and consumption for the Iberian Peninsula. The notation used as "(D-1)", for example, indicates the value of the referred variable, 24 hours before the time for which it is intended to predict (day-1). Likewise, "(D-2)" refers to 48 hours before, and "(W-1)" to 168h before (a week before). In case the reference is "D", it indicates that it is the value for the same period for which it is intended to predict. Initially, in many of the variables tested here, the reference used is "D" as a way to make a previous analysis if it would be worthwhile to test for other references, since the value in this situation will be, in the great majority of variables, higher than previous references, be it 24, 48 or 168 hours before. Table I presents the Spearman coefficient results obtained for the data related to energy production by technology, and consumption.

Here some groups of energy technologies that could have some influence on the price were tested, as the value of load consumption and these energy technologies individually. Once the set with a higher correlation (MIX3) was obtained, it served to create a new variable to be tested: ratio between MIX3 / Load Consumption. This proved to be, along with PRO Thermal Energy, the most influential to the price evolution, as such, both were added to serve as inputs to the forecast models.

Initially, entries with the chronological reference "(D-1)", would be used as inputs for the models. However, after verifying that consumption and production data for Portugal were only available around 3 pm the following day, these values were abandoned and gave place to "(D-2)" or "(W-1)", depending on the value of the correlation proves more preponderant in one case or the other. This value is usually higher in the variables one week before (168h), given the importance of the weekly seasonality mentioned above.

The meteorological data collected for the Iberian Peninsula was also evaluated, and this data were generally more

influential in price development. Together with this data, and because they are also meteorological data, although from another source and monthly values instead of day based as in the other cases, solar radiation values will be included. The correlations of the spoken data can be verified in Table II.

TABLE I.   PRICE CORRELATIONS WITH LOAD AND PRODUCTION BY TECHNOLOGY

Correlation coefficients relative to Price, using Spearman Coefficient		
Tested Variable	Chronological reference	Correlation value in percentage (ρ)
Load Consumption	D	50,30%
	D-1	38,94%
	D-2	28,64%
	W-1	45,11%
Hydraulic	D-1	31,29%
	D-2	19,00%
	W-1	33,43%
Thermal PRO	D-1	51,88%
	D-2	36,02%
	W-1	51,78%
Eolic	D	-23,30%
MIX1 – Eolic + Photovoltaic	D	-22,37%
MIX2 – Eolic + PV + Thermal PRE	D	-20,82%
MIX3 – Eolic + PV + Thermal PRE + Nuclear	D	-24,75%
MIX4 – Hydr. + Eolic + PV + Thermal PRE	D	6,83%
Ratio MIX3/Load Consumption	D-1	-39,73%
	D-2	-22,34%
	W-1	-36,07%

TABLE II.   PRICE CORRELATION WITH METEOROLOGICAL DATA

Correlation coefficients relative to Price, using Spearman Coefficient		
Tested Variable	Chronological reference	Correlation value in percentage (ρ)
Irradiation	D	-83,22%
Average Temperature	D	-57,28%
Wind	D	-34,74%
Precipitation	D-1	-20,96%

Analyzing the values presented in this table, we can see the importance of both solar radiation and temperature in price evolution. Finally, we will analyze the influence that recent historical price values will have on this same variable. The values of these Spearman coefficients are shown in Table III.

TABLE III.   PRICE CORRELATION WITH ITS OWN CHRONOLOGICAL DATA

Correlation coefficients relative to Price, using Spearman Coefficient		
Tested Variable	Chronological reference	Correlation value in percentage (ρ)
Price	D-1	76,94%
	W-1	73,79%
	W-2	71,16%

These values, as expected, have a great correlation with the price that is intended to forecast, so these will be the main inputs to be included in the models. As previously described, a load forecast model will be developed to assist in obtaining accurate price values. For this, the correlation of some data with the value of the load was also evaluated, being the values obtained with respect to the correlations of these data, presented in Table IV.

As it was concluded verifying Table IV, it is demonstrated here the greater percentage of correlation, in relation to the evolution of the load, of the deviation of the average values of solar radiation and of average temperature ("D.Aver.Irrad." and "D.Aver.Temp." , respectively), than of the verified values of one and other variables. This is explained by the connection between these, and by the fact that when more extreme temperatures and irradiance values occur, there is also a need for greater consumption, for heating / cooling, or even lighting.

TABLE IV.   LOAD CONSUMPTION CORRELATION WITH METEOROLOGICAL DATA

Correlation coefficients relative to Price, using Spearman Coefficient		
Tested Variable	Chronological reference	Correlation value in percentage (ρ)
Irradiation	D	-4,20%
Deviation of the average value of Irradiation	D	72,73%
Average Temperature	D	-4,00%
Deviation of the average value of Temperature	D	51,64%
Wind	D	15,78%
Precipitation	W-1	-9,33%

Once again, historical variables will play a major role in predicting the load, giving an idea of the most recent range of values in the Iberian Peninsula. The weekly seasonality of the load is again demonstrated when checking the correlation values in Table V.

TABLE V.   LOAD CONSUMPTION CORRELATION WITH ITS OWN CHRONOLOGICAL DATA

Correlation coefficients relative to Price, using Spearman Coefficient		
Tested Variable	Chronological reference	Correlation value in percentage (ρ)
Load Consumption	D-1	81,74%
	D-2	64,90%
	W-1	90,66%
	W-2	86,55%

It is noticeable, as in the analogous case of price, that there is a large correlation between the most recent available load consumption data values and those which are expected to be predicted. These variables will then be, once again, the main data to be used in the forecast models to be applied. In addition to the inputs tested in the tables presented in this section, chronological variables will also be used, due to the variability of the load and the price depending on the intensity of the business activities throughout the day, week and, to some extent, the year. With this in mind, the variables of this type to be used will be Hour (H), Day (D), Day of Week (DoW), Month (M), National Holidays (NH) and Seasonality (Summer, Winter or intermediate seasons).

B. Forecasting model

A test ANN is used, based on the nftool and nntool methods of Matlab. The nftool is used to implement, create and train the network, and the nntool is used to simulate the results. Although with a small difference, the method with a smaller error, being the selected method was method 2 (3.52% versus 3.77% of method 1) which used only nntool to create, train and simulate the networks. All simulations and networks were created according to this same method, with the algorithm feedforward backpropagation, in which the information flow occurs unidirectionally, progressing from the input neurons to the output neurons. According to the architecture selected for all the networks developed in this paper, these are multi-layer perceptrons, and in this type of networks where there is a hidden layer that contains a variable number of neurons, this being a fact that makes this type of networks widely used. Two-layer networks were used, and being that there is no universal rule for the selection of the number of neurons N in the hidden layer the rule expressed in (2) has been used, where n is the number of network inputs.

$$N = 3n + 2 \qquad (2)$$

In the other parameters to be defined in the networks, Levenberg-Marquardt backpropagation [10] was used as a training function, and as learning and performance adaptation functions the chosen were, learngdm (gradient descent with momentum weight and bias learning function) and mse (mean-squared error), respectively. In layer 1 (hidden layer) the function tansig was used as transfer function, and in layer 2 (output layer) the function purelin.

To implement all the case studies of this paper, it was used a computer with an Intel (R) Core i7-4710HQ CPU @ 2.50 GHz 2.50 GHz, 8 GB of RAM and a Windows 10 Home 64-Bit operating system.

III. CASE STUDY

A. Specifications

In order to evaluate the developed artificial neural networks, the data collected and treated for its construction were divided. The data to be used to the day-ahead forecast model, were the data corresponding to the time window from July 1, 2016 until June 30, 2017. These data were distributed in such a way regarding that their division was about 85% for training data, and 15% for test data. In this way, the training data would serve to train the neural networks in order to "teach" it how it would behave, given the presented data. This learning process would then be simulated, with data corresponding to 15%, representing the test data. The breakdown resulted in the random choice of 8 weeks (about 15.34% of the total sample of 365 days) to be tested, providing the input data to the network, but not the expected output data. The 8 weeks of the test set chosen were as follows:

- August 23, 2016 until August 29, 2016
- September 19, 2016 until September 25, 2016
- January 6, 2017 until January 12, 2017
- May 13, 2017 until May 19, 2017
- November 17, 2016 until November 23, 2016

- January 19, 2017 until January 25, 2017
- February 12, 2017 until February 18, 2017
- November 30, 2016 until December 6, 2016

These designated training and test sets were used in the created models. The Day-Ahead Forecasting Model has a great quantity of recent information available, which means we have many variables to choose from. This also makes the error tolerance small to a model like this, compared to others with greater time horizons, like a weekly or monthly model.

In order to guarantee all the precautions to be taken to ensure that the forecasts made in this daily model are as accurate as possible, additional tests were carried out. Therefore, before proceeding to the training and simulation of the networks of the different models developed, tests were carried out regarding the optimum parameterization to be used and the number of trainings to be performed in the simulations of this model. For these tests, as in the previous one of the Matlab® tool to be used, the same test network was used. These aspects were evaluated according to the subsequent analysis of the MAPE error result.

The first test that was performed for this model was the parameterization to be used in the training of the networks, in order to maximize the accuracy of the models to be created. For this purpose, five models with different parameterization values were designed, varying the values of maximum number of validation checks without performance improvement (*max_fail*) and minimum magnitude gradient (a value that will become very small as the training reaches the minimum) (*min_grad*). These values are used to finalize the training of neural networks, being able to become critical parameters, by too premature stop of the training.

Each model was tested five times in order to have more than one output to test the results more effectively. Five networks were then developed for each model, to make up the desired number of outputs, and five training sessions were carried out in each one, before the simulation of the results, which can be seen in Table 6.

TABLE VI. TABLE 1 PARAMETRIZATION NETWORKS

Parametrization Networks	Min_grad	Max_fail	MAPE
Model of parametrization 1	1e-07	6	3,70%
Model of parametrization 2	1e-08	30	3,73%
Model of parametrization 3	**1e-08**	**100**	**3,49%**
Model of parametrization 4	1e-07	30	3,63%
Model of parametrization 5	1e-06	100	3,60%

In order to minimize the error obtained in the simulations of the price forecast, we then created the models to be developed for this purpose, and the subsequent analysis of the MAPE error for each one of them. For the development of these models we used the variables selected based on the studies carried out in previously on this paper. Thus, in a constructive way, new inputs were added to the previous models. Then, changes were made to the previous models in order to develop others that had more attention to variables with a positive or negative correlation, respectively, the Daily Price Models (DPM) 04 and 05.

One of the aspects to take into account also in the choice of the inputs to use, from a practical point of view, was its availability so that the forecast could be made from one day to the next, without the need to wait for a certain hour later on that day, to make this information available. With this in mind, it was decided not to use the Load Consumption data of the day prior to the forecast, in the case of the forecast model of this variable, discussed later on this paper. Since this data, although for Spain were available almost immediately, for Portugal this data was only available for download on the platform for this purpose, more than 12 hours later. This led to opt instead for the Load entry (D-2) (value of the load in the Iberian Peninsula, 48 hours before the value to be predicted) instead of Load (D-1) (the same value, 24 hours before the forecasted value).

The distribution of the variables by the models created for the forecast of price for the Day-Ahead Market was as follows:

- DPM 01: Price(D-1) + H + D + Price(W-1)
- DPM 02: previous variables + DoW + M + NH
- DPM 03: previous variables + Irradiation + Av.T.
- DPM 04: previous variables + Price(W-2) + Thermal PRO(W-1) – M
- DPM 05: DPM 03 + MIX3/Load ratio(W-1) + Wind – M

In order to test the influence of the value of load consumption value on the forecast to be made of the electric energy price, 5 other models were made, similar to the previous ones, but with the inclusion of the real value of the verified load consumption:

- DPM 06: Price(D-1) + H + D + Price(W-1) + Load Consumption
- DPM 07: previous variables + DoW + M + NH
- DPM 08: previous variables + Irradiation + Av.T.
- DPM 09: previous variables + Price(S-2) + Thermal PRO(W-1) – M
- DPM 10: DPM 08 + MIX3/Load ratio(W-1) + Wind – M

A. Price forecasting results

Next, we proceed to the training and simulation of the neural networks created through these models, being that five networks were developed for each one of the models, so that 5 different outputs were obtained in each of the cases, being evaluated the error for each network and averaging next. This comparison of errors between the different models, with and without load, can be verified in Table 7.

TABLE VII. ELECTRICITY PRICE FORECAST FOR DAY-AHEAD MARKET

Models without Load Consumption values		Models with Load Consumption values	
DPM 01	9,77%	DPM 06	9,59%
DPM 02	**9,49%**	DPM 07	9,97%
DPM 03	11,28%	DPM 08	10,66%
DPM 04	10,73%	DPM 09	10,78%
DPM 05	9,56%	**DPM 10**	**8,34%**

As it can be verified, the model that obtained a smaller value of MAPE was the DPM 02, in the group of those who did not include the value of load consumption, and the DPM 10, in the group of those that included that same value. Since

the error value in DPM 10 was still lower than in DPM 02, then there is the need to develop and test a load consumption prediction model to follow the development of this one. It should be noted that a model is not always more precise because it has more variables in its constitution, as an example, the DPM 02, which is one of the developed models that has the least number of variables. This may be due to the fact that some of the variables used have a direct correlation with the electricity price and others have an indirect correlation, affecting the final forecast result when used together.

B. Price and load forecasting results

For the load prediction model, two new variables studied for this purpose were created, as mentioned in previously, which revealed a higher correlation than the base variables. These were the deviation of the average value of temperature and the deviation of the average value of solar radiation (D.Temp. and D.Irrad., respectively), and are explained by the increase in load consumption, on periods in which both the temperature and the solar radiation , are more extreme, for example in their positive and negative peaks, reached by norm in summer and winter, in that order.

Thus, 5 models were created with the following characteristics, in order to predict the load consumption on MIBEL (denominated DLM - Daily Load Model):

- DLM 01: H + D + Load Consumption(W-1)
- DLM 02: previous variables + DoW + M + NH
- DLM 03: previous variables + Load Consumption(W-2) + Load Consumption(D-2)
- DLM 04: previous variables + D.Irrad.+ D.Temp.
- DLM 05: DLM 02 + D.Irrad. + Load Consumption(W-2)

The construction of these models, followed the logic of the models developed for the price forecast, only to mention that in DLM 05, the variables of the model DLM 02 were used, and only the entries of the later models that had greater values of correlation were added, Load Consumption(W-2) in the case of DLM 03, and D.Irrad. in the case of DLM 04.

In these models, the procedures that were previously tested for the method to be used, parameterization were again followed. The results obtained in this simulation can be found in Table 8.

TABLE VIII. CREATED MODELS TO LOAD CONSUMPTION FORECAST

Model	MAPE
DLM 01	5,26%
DLM 02	5,29%
DLM 03	**4,03%**
DLM 04	4,10%
DLM 05	6,74%

The model that obtained a smaller MAPE was DLM 03, with 4,03%. The average values for load consumption obtained for each hour in the simulations of this model, were then used as input, to test the model DPM 10, in order to verify if the error obtained using the predicted value of load consumption would compensate. This was done to compare with the use of the best model without the value of the load consumption in its inputs (DPM 02).

Again, the simulation of this new network (DPM 10 with load consumption forecasting) was carried out taking into account the criteria previously determined, and with 5 outputs. The mean error obtained in these simulations resulted in a MAPE of about 9.32%. As expected, this error value is higher than the value obtained for DPM 10 with the real load consumption values. This is justified because the load forecast itself has an associated error. However, this value remains below the MAPE value of the best model without use of the load consumption values (MAPE_ (DPM 02) = 9.49%), making them the models selected for forecasting the price in the day-ahead market.

Finally, a procedure was applicated to handle the error. To achieve this, it were used dispersion measures, as a measure of error management, in order to identify and minimize it. After calculating the mean and standard deviation of the outputs for each period, in the selected final models, a range was constructed by adding and subtracting the resulting standard deviation value to the mean value obtained for each moment (hour). Then, the values of outputs that did not integrate the created interval would not be included in the final calculation of the mean, whose value was then considered for the final predicted value for the moment in question. The objective of this process was to reduce the dispersion of the predicted values, ignoring the values that deviated too much from the average, and in this way, to reduce the amplitude of the values, since it is preferable to work with values with a smaller variance. Applying this process, it was possible to obtain a reduction of the error of the final model, which decreased from 9.32% to 8.24%.

Finally, the Spearman coefficient of the final results obtained with the real price values was calculated, and a graphic demonstration of this correlation was also used. The value obtained was 90.54%, and Figure 4 shows the graph that correlates these two variables. The correlation of the data calculated in this way did not obtain a result as promising as when calculated by the Spearman coefficient, however it has a very acceptable value (about 0.84). Then, the outputs were compared with the actual price values, for the entire test set, through Figure 5**Fig**. In this, it is also possible to check, in gray, the area representative of the standard deviation of the analyzed data. The precision of the selected model can be ascertained, being that the model has greater difficulties in tracking the actual price when peaks exist.

IV.   CONCLUSIONS

Electricity market prices forecast is an active domain of research due to the in-creasing need for adequate forecasting models that are able to consider the different factors that influence electricity market prices.This paper presents a methodology for electricity market prices forecasting with artificial neural networks and data analysis, using the spearman's rank correlation co-efficient. Results using real data from the Iberian electricity market show that the proposed approach is able to achieve very low forecasting errors, through the proper identification of the most relevant variables to be considered in the training process of the artificial neural network. As future work, different models will be tried, as alternative to artificial neural networks. In particular, hybrid neuro-fuzzy inference systems and ensemble learning approaches are considered, due to their promising nature in dealing with the association between multiple input variables.
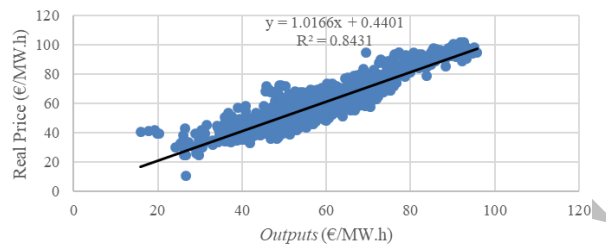

Fig. 2 Correlation between the outputs of the model and electricity prices
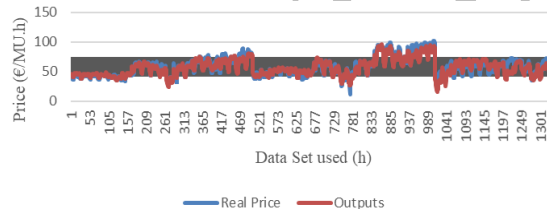

Fig. 3 Comparison between real electricity price and generated outputs

REFERENCES

[1] European Commission, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on ensuring the cross-border portability of online content services in the internal market," Brussels, 2015.

[2] F. Rahimi and A. Ipakchi, "Using a Transactive Energy Framework: Providing Grid Services from Smart Buildings," IEEE Electrif. Mag., vol. 4, no. 4, pp. 23–29, Dec. 2016.

[3] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," Renew. Sustain. Energy Rev., vol. 81, pp. 1548–1568, 2018.

[4] "MIBEL - Mercado Ibérico de la Electricidad," 2018. [Online]. Available: http://www.mibel.com.

[5] W.-Y. Zhang, Z.-W. Wei, B.-H. Wang, and X.-P. Han, "Measuring mixing patterns in complex networks by Spearman rank correlation coefficient," Phys. A Stat. Mech. Appl., vol. 451, pp. 440–450, 2016.

[6] M. M. Baig, M. M. Awais, and E.-S. M. El-Alfy, "AdaBoost-based artificial neural network learning," Neurocomputing, vol. 248, pp. 120–126, 2017.

[7] M. G. De Giorgi, A. Ficarella, and M. Tarantino, "Error analysis of short term wind power prediction models," Appl. Energy, vol. 88, no. 4, pp. 1298–1311, 2011.

[8] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," Appl. Energy, vol. 87, no. 7, pp. 2313–2320, 2010.

[9] Y. Mu, X. Liu, and L. Wang, "A Pearson's correlation coefficient based decision tree and its parallel implementation," Inf. Sci. (Ny)., vol. 435, pp. 40–58, 2018.

[10] S. Mammadli, "Financial time series prediction using artificial neural network based on Levenberg-Marquardt algorithm," Procedia Comput. Sci., vol. 120, pp. 602–607, 2017.