Paper accepted for presentation at 2009 IEEE Bucharest Power Tech Conference, June 28th - July 2nd, Bucharest, Romania

Forecasting Electricity Prices in Spot Markets - One Week Horizon Approach

A. F. Duarte, J. N. Fidalgo, *Member, IEEE,* and J. T. Saraiva, *Member, IEEE*

***Abstract*—This paper describes the methodology developed to build estimates of electricity prices having the horizon of one week. This approach uses artificial neural networks and includes a particular treatment of weekends and national holidays as a way to improve the quality of the results. The developed approach was tested using data obtained from the Spanish market operator for the time period of 2006 to 2008. The obtained value of MAPE - Mean Absolute Percentage Error - was 12,62% for workdays and 10,73% for holidays and weekends. The obtained results show that this study has interest to the market agents in question, since realistic forecasting was achieved.**

***Index Terms*—Electricity Sector Restructuring, Electricity Markets, Artificial Neural Networks, Forecasting Electricity Prices.**

I. Introduction

THE restructuring process of the electricity sector and the emergence of electricity markets in most countries of continental Europe began during the 90s. On electricity markets, transactions are negotiated in advance before the physical delivery - a day, a few hours or even minutes - based on the forecast of consumption.

The adjustment of the imbalances (inevitable) that arise between the negotiated values and the values registered in generation and consumption, are worked by procedures that may or may not have a competitive nature.

While searching for a competitive environment, the electricity price is no longer set by its own pricing methods, and starts to be established by market mechanisms [1][2].

The electricity spot market, also known as Pool, is a structure that relates generation to consumption. It was against this background that came an interest in this work, which is the forecast of electricity prices resulting from the Spanish Pool.

The need to forecast a week in advance is justified when used in the construction of market strategies for power plants, which means making proposals to be submitted to the pool.

In the past few years, many papers report the use of ANNs in short-term electricity demand forecasting [3], [4], [5], [6]. Most recently, some papers have repported the use of ANNs to forecast next-day electricity prices [7], [8]. Yet, some market strategies ask for a longer horizon. In such case, forecasting a week in advance is expected to add extra value.

The paper is structured as follows: Section 2 provides an overview of the ANN forecasting methodology; Section 3 describes the adopted architecture and the data series used to train and test our methodology.

A. F. Duarte is with Efacec Engenharia S. A., Porto, Portugal (e-mail: andre.freire.duarte@fe.up.pt).

J. N. Fidalgo is with the Department of Electrical Engineering at Faculdade de Engenharia da Universidade do Porto, Portugal (e-mail: jfidalgo@fe.up.pt).

J. T. Saraiva is with the Department of Electrical Engineering at Faculdade de Engenharia da Universidade do Porto, Portugal (email: jsaraiva@fe.up.pt).

Section 4 discusses the obtained results compared to actual data. The paper is rounded off by Section 5, which draws relevant conclusions from the study.

II. Short-Term Forecasting Using Artificial Neural Networks

A. General Aspects

The Artificial Neural Networks - ANNs -, as a methodology applied to the short-term forecasting, have experienced considerable development since 1980, with overall results that can be considered satisfactory.

ANNs are mathematical models based on the functioning of the human brain, and are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems and therefore are called neurons.

These neurons have inputs, an internal processing unit and outputs. The internal processing unit transforms inputs into outputs.

The literature explains certain characteristics of the ANNs that make them particularly useful in predicting time series [7]. It will be noted that the most relevant characteristic is their ability to approach almost any function [9], [10]. This means that the ANNs has the facility to approximate any function to a data set, which turns especially important when dealing with complex functions.

Another relevant feature of ANNs methodology is its speed. Solutions can be generated very quickly for most problems, although a number of conditions must be fulfilled [7]:

- Network configuration should not be too large;
- The training examples set should not be too long. The smaller the set, the more times every example will go through the network and the faster the solution will be obtained;
- The composition of training examples must be homogeneous. The more similar the examples, the faster the learning process.

Given these conditions, one way to accelerate the process is to make a preliminary selection of the data sets used in the training examples.

B. Training the ANN

Regarding the learning of ANN, it was chosen a training function based on the *Back-Propagation*algorithm. The adopted function updates the weights of the connections between neurons in accordance with the method of optimization of *Levenberg-Marquardt*. The use of this method is typically recommended for supervised learning however it does require more memory compared to some other algorithms [11].

978-1-4244-2235-7/09/$25.00 ©2009 IEEE

III. Training and Testing Data

In this study, the ANN output (target) is, obviously, the electricity price. In what ANN inputs is concerned, they have to be selected according to their relevance in the electricity price forecasting.

Within the available set of variables, potentially related to the electricity prices, some are commonly known to have a significant effect on these prices.

However, there are others where its relevance can only be demonstrated heuristically. After a phase of ANN input selection, the data was divided into two groups: training and test.

The training data corresponds to the period from January 1, 2006 to December 31, 2007, thereby completing a two-year period. Regarding to the test set, the data relates to the first two months of 2008, i.e.January and February 2008.

For the test set, it was decided to use a period subsequent to the training group because this is the condition to check in actual operation**.**

A. Workdays and Month of the Year

The day of the week affects on a large scale the price of electricity resulting from a competitive market. This happens because there are proper consumptions in each day of the week that lead to specific price evolutions along the day.

Typically the driest months originate the increase of generation by thermal power plants, which, by their nature, have higher marginal costs.

This contributes to the increase of the electricity price in the day-ahead markets.

Similarly, in the months of highest rainfall, there will be a larger use of hydro power plants, so that market prices tend to fall daily.

B. Brent Index and Natural Gas Prices

A considerable part of electricity generation comes from thermal groups using oil, its derivatives and natural gas as primary sources [12].

This clearly suggests that the prices of these primary sources in international commodity markets are used as input variables given to their influence on the price of electricity.

The price of these commodities results from a negotiation process in international markets and it is influenced by several factors as reserves of primary resources, the decisions of producers (in several cases grouped in well known organizations), speculation, the variation of the value of international currencies, namely the US dollar.

C. Electricity Price in Previous Days

As previously mentioned, due to the weekly cycle of the demand, there are typical values of the electricity demand in each week day that display a clear auto-correlation and that differ from the demand patterns in some other days.

As a result, the demand in day d will be strongly auto-correlated with the demand in day $d-7$, that is, with the demand in the same day of the previous week.

As indicated in Table I, the auto-correlation coefficient obtained for the day $d-1$ is the largest one. However, since the developed forecasting exercise had an horizon of one week, it was not possible to use this information regarding day $d-1$.

In any case, the auto-correlation coefficients also display large values for days $d-7$ and $d-14$. This indicates that the price of the electricity in each hour of day d is strongly correlated with the price in the same hour in the previous weeks.

Given this information, we performed several experiences including these variables as inputs of the ANN. These experiences indicated that the use of the hourly prices that occurred 7 days before the day in which the forecast is to be made were beneficial for the results.

However, including the hourly prices that occurred 14 days before would not improve the quality of the results.

As a consequence, these prices were not used as input variables of the ANN. According to the values in Table I, the auto-correlation coefficient displays a larger value for day $d-1$.

As mentioned before, the hourly prices in day $d-1$ were not used as input variables since the developed forecasting exercise had an horizon of 1 week.

Table I
Auto-correlation Coefficients for Several Analyzed Scenarios

	7AM	*12AM*	*9PM*		*7AM*	*12AM*	*9PM*
$d-1$	**0.847**	**0.709**	**0.860**	$d-9$	0.659	0.517	0.750
$d-2$	0.758	0.606	0.814	$d-10$	0.636	0.488	0.736
$d-3$	0.731	0.568	0.790	$d-11$	0.634	0.462	0.730
$d-4$	0.713	0.551	0.785	$d-12$	0.653	0.476	0.736
$d-5$	0.719	0.570	0.782	$d-13$	0.694	0.542	0.739
$d-6$	0.740	0.614	0.786	$d-14$	**0.726**	**0.643**	**0.757**
$d-7$	**0.779**	**0.753**	**0.817**	$d-15$	0.672	0.473	0.710
$d-8$	0.713	0.578	0.777				

D. Architecture of the network

In this work, the selection of the most adequate architecture of the ANN was based on a trial and error process in which we considered the following aspects:

1) Complex architectures are usually not required in forecasting applications given that the temporal series that are associated with them are not described by complex functional relations. In series of prices of electricity, it is typically possible to detect a trend that is related with the demand habits. It is also important to refer that complex ANN architectures easily internalize data with noise, which is clearly undesirable;
2) The comparison between the errors of the training set and of the test set corresponds to a good estimator of the performance of the ANN. If the observed errors are large, one can conclude that the ANN has an insufficient capacity to model and emulate the data. If that is the case, it will be necessary to increase the number of units of the hidden layer of the ANN;
3) On the other hand, when the error observed in the test set is considerably larger than the error obtained for the training set, one can conclude that the ANN captured the atypical data but it was unable of generalizing them when it is exposed to the data in the test set. In this case, the number of units in the hidden layer should be

reduced.

The application of these general indications guided the construction of the ANN architecture displayed in Figure 1. The final architecture was selected after several tests we made, both regarding the input data as well as the number of units in the hidden layer.

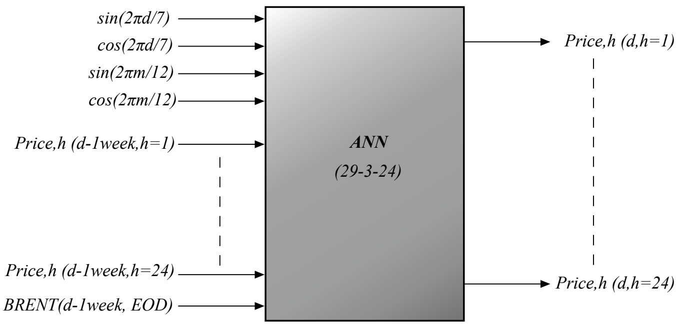

Figure 1. ANN architecture that produced the best results.

IV. Results

During the entire design process, several ANNs were trained seeking the one with the best performance. The performance indicator used on the training process was the Mean Square Error (MSE). The choice of the number of neurons from the hidden layer was conclusive for the selection of the best configuration.

It was shown that the ANN had the best performance with three neurons on the hidden layer. The next procedure corresponded to the analysis of its ability to keep up with the typical evolution of the electricity prices throughout the day and for the different workdays.

A. Workdays

Observing the evolution of electricity prices throughout the day, we can notice a pattern explained by the load diagram representative of workdays. The influence of trading markets and industry is determinant nowadays. During the day break, the price remains low, reaching its minimum value at 5 am.

Since then, the price starts rising until it reaches its peak at 11 am. The highest peak occurs between 8 pm and 9 pm. During this period, the domestic demand is significantly influent because the referred peak represents its maximum activity.

As wishful, the ANN manages to interpret that pattern with great success. Figure 2 illustrates what was last said.

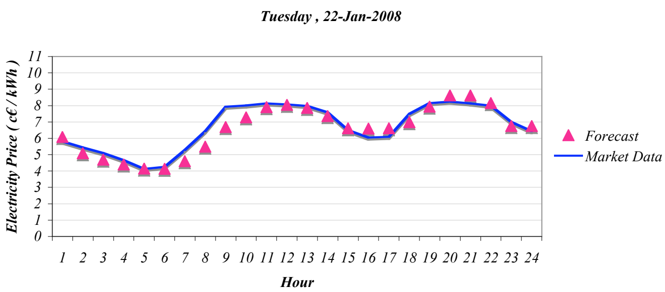

Figure 2. Evolution of the electricity price on Tuesday, 22-Jan-2008.

B. Weekends

On what concerns weekends, the price evolution throughout the day is completely different, what would be expected, seeing that the consumpion pattern is totaly different on those days. At dawn, between 1 am and 2 am, a slight increase of the price, in comparison with workdays, is noticed.

The ANN interprets this event successfully. After that period, a latter peak happens around 12 am, although substantially inferior to what takes place in workdays. The highest summit on weekend days happens around 9 pm, which is reflected on the electricity price.

Once again, the ANN performs satisfactorily for weekends. Figures 3 and 4illustrate the comparison between the registered value on the Spanish market and the value obtained by the adopted forecast.

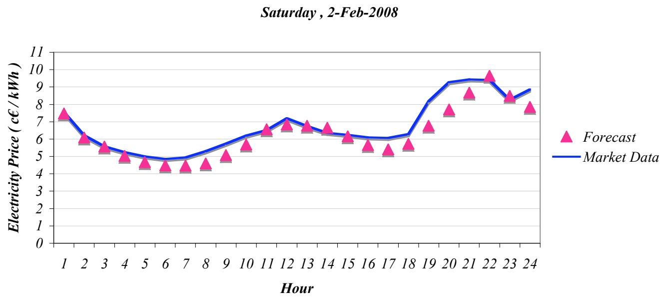

Figure 3. Evolution of the electricity price on Saturday, 2-Feb-2008.

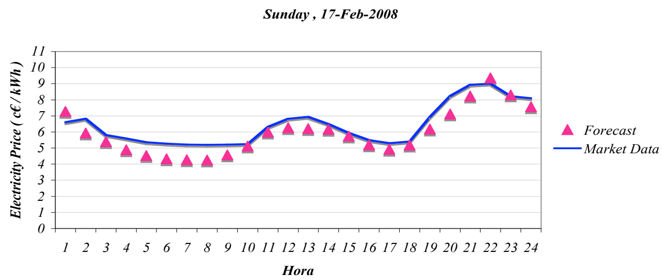

Figure 4. Evolution of the electricity price on Sunday, 17-Feb-2008.

C. National holidays

National holidays are considered as non-typical events on electricity pricing sets. Although their rate of occurrence is relatively low, in relation with regular days, the effects caused on the performance are substantial.

Two types of effects caused by national holidays can be considered [6]:

1) On one hand, the forecasting for those days tends to be larger than it should be, because the use of electricity is comparatively lower on national holidays (which is reflected on the price of electricity);
2) On the other hand, the forecast that depends on series that correspond to national holidays have a tendency to be lower than they should.

An approach to the treatment of these events will be shown on the next phase. During the process of investigation, different methods were found to cope/ deal with national holidays, but the most frequent is one that is based on the inclusion of a variant of extraordinary entrance that defines if the day of the forecasting is a national holiday or not.

In this method, for instance, it can be given the value 1 to national holidays and the value 0 to regular days as ANN's entrance. This method was adopted on a first phase, however no significant improvements were demonstrated when applied to the test set.

Based on this dissatisfaction, a different treatment was adopted, which will be explained next.

On a first stage, having as a support the holidays[1] calendar referring to the years 2006 and 2007, a comparison was made (on the training set) between all holidays and the previous/further days.

So, taking a holiday d_h, the daily mean value of the electricity price was compared between the day d_{h-1}, d_h and d_{h+1}. This analysis showed that the daily mean value of electricity prices on holidays is approximately 9% lower than its nearby. Attending that our data set refers to the period going from the 1st of January to the 29th of February 2008, the existing holidays are the first and the sixth of January 2008.

In order to study the specific behavior on those two cases, a more profound analysis was developed. We applied the neighbour-comparison principle, which was already used, but this time that analysis was extended to all periods of hour. We have concluded that both days have very particular behaviours:

1) The 1st of January shows a very elevated price on the initial hours which symbolizes the high consumption of electricity that occurs during the early hours of New Year. The period of dawn, where a great variation can be noticed, is located between 1 am and 7 am. As it can be seen on Table II.

Table II
RELATIVE PERCENTUAL RAISE FOR 1 JAN 2008

Hour	Relative percentual raise
1h	60%
2h	80%
3h	80%
4h	80%
5h	60%
6h	50%
7h	20%

2) The 6th of January also represents a festivity day, although very different from New Year. On this day a slight increase of the price is noticed during the period of 7 pm until midnight. This can be seen on Table III.

[1]Note that they were just considered holidays in Spain. Regional holidays were not considered because they do not reflect the behavior of the whole country but simply the demand of a region or a city.

Table III
RELATIVE PERCENTUAL RAISE FOR 6 JAN 2008

Hour	Relative percentual raise
19h	30%
20h	30%
21h	30%
22h	20%
23h	20%
24h	20%

Figures 5 and 6 illustrate the positive contribution that the individual treatment on holidays causes in forecasting. It is certainly a longer process than the inclusion of a variable that identifies the existence of a holiday. However, in this case, it revealed itself to be much satisfactory.

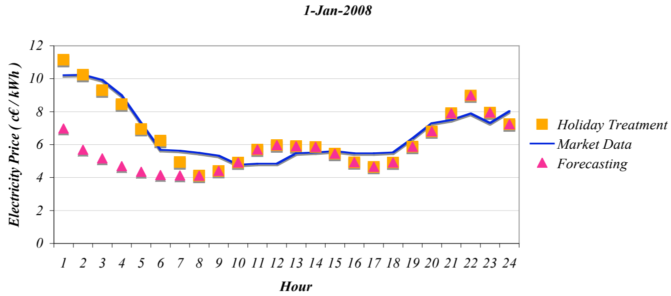

Figure 5. Holiday treatment contribution to forecasting, 1-Jan-2008.

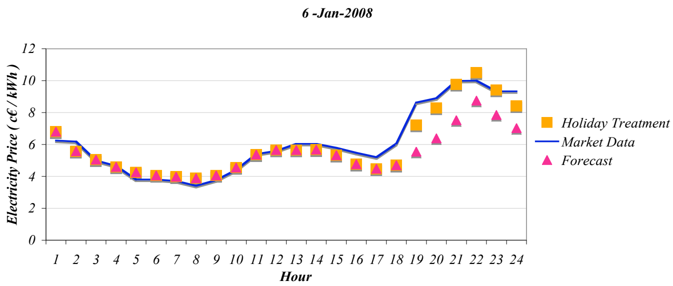

Figure 6. Holiday treatment contribution to forecasting, 6-Jan-2008.

After numerous experiments, we reached the architecture that revealed the best performance for the forecasting exercise. Through the selection process it was possible to conclude that the choice of the number of neurons of the hidden layer is determinant.

Using a larger number of neurons on the hidden layer, a general dissatisfaction on the results was experienced, which resulted from the fact that the network was exposed to the 'over-adaptation' phenomenon. That number was reduced systematically until the one with the best performance on the ANN was identified.

The chosen variables went through tests that finally droven us to the best architecture. The adopted holidays treatment, although hardworking, proved itself to be effective to the tested set.

The obtained value of MAPE - Mean Absolute Percentage Error - (performance indicator much used in forecasting exercises) was 12,62% for workdays and 10,73% for holidays and weekends (with the use of holidays treatment). The MAPE global value resulted in 12,25%. On Table IV the hourly-based values from this indicator are represented.

Table IV
HOURLY-BASED MAPE OBTAINED WITH THE BEST ANN ARCHITECTURE AND IMPROVED BY THE NATIONAL HOLIDAYS TREATMENT

Hour	$MAPE$	Hour	$MAPE$	Hour	$MAPE$
$1h$	12.332 %	$9h$	15.052 %	$17h$	13.561 %
$2h$	9.669 %	$10h$	12.849 %	$18h$	12.896 %
$3h$	9.552 %	$11h$	11.085 %	$19h$	12.223 %
$4h$	12.252 %	$12h$	10.518 %	$20h$	9.523 %
$5h$	14.200 %	$13h$	10.758 %	$21h$	8.159 %
$6h$	14.302 %	$14h$	11.591 %	$22h$	9.804 %
$7h$	14.085 %	$15h$	12.722 %	$23h$	11.957 %
$8h$	17.394 %	$16h$	13.862 %	$24h$	13.570 %

V. CONCLUSIONS

The achieved values were compared with a very similar study regarding the same market [7], proving the achievement of satisfactory results, considering the difference in temporal horizon between both of them.

The obtained results reveal that this study will certainly be of interest to market agents, since realistic forecasting was achieved.

As following directions for future work, we suggest a continued consolidation of the training data in order to test the consistence and durability of the developed model.

VI. REFERENCES

[1] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proceedings of the IEEE*, vol. 88, pp. Pages 163–169, February 2000.

[2] N. Fabra and J. Toro, "Price wars and collusion in the spanish electricity market," *International Journal of Industrial Organization*, vol. 23, no. 3-4, pp. 155–181, 2005.

[3] H. Bacha and W. Meyer, "A neural network architecture for load forecast," *Neural Networks, 1992. IJCNN., International Joint Conference on*, vol. Volume 2, no. ISBN 0-7803-0559-0, pp. 442–447, 1992.

[4] K. Lee, W. Cha, and J. Park, "Short-term load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. Volume 7, February 1992.

[5] C. Chen, Y. Tzeng, and J. Hawang, "The application of artificial neural networks to substation load forecasting," *Electric Power Systems Research*, vol. Volume 38, pp. 153–160, 1996.

[6] J. Fidalgo and M. A. Matos, "Forecasting portugal global load with artificial neural networks," *Lecture Notes in Computer Science*, vol. Volume 4669/2007, no. ISBN 978-3-540-74695-9, pp. 728–737, 2007.

[7] R. Pino, J. Parreno, A. Gomez, and P. Priore, "Forecasting next-day price of electricity in the spanish energy market using artificial neural networks," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 1, pp. 53–62, 2008.

[8] J. Catalao, S. Mariano, V. Mendes, and L. Ferreira, "Application of neural networks on next-day electricity prices forecasting," *Universities Power Engineering Conference, 2006. UPEC '06. Proceedings of the 41st International*, vol. Volume 3, pp. 1072–1076, September 2006.

[9] G. Cybenko, "Approximations by superpositions of sigmoidal functions," *Math, Control, Signals Systems*, vol. Volume 2, pp. 303–314, 1989.

[10] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feed-forward networks are universal approximators," *Neural Networks*, vol. Volume 2, pp. 359–366, 1989.

[11] The MathWorks, Inc., *MATLAB Neural Network Toolbox User Guide*, 7.6.0 (r2008a) ed., February 2008.

[12] E. F. Dominguez and J. X. Bernat, "Restructuring and generation of electrical energy in the iberian peninsula," *Energy Policy*, vol. 35, no. 10, 2007.

VII. BIOGRAPHIES

André F. Duarte was born in Porto, Portugal in 1983.

In 2008 he got his MSc in Electrical and Computer Engineering from the Faculdade de Engenharia da Universidade do Porto, FEUP. In 2008 he joined Efacec Engenharia S.A. – a leading Portuguese engineering group – where he presently collaborates in several projects related with the development of EMS and DTS systems. His research interests include wholesale power markets, energy policy, energy economics and forecasting in electricity markets.

José N. Fidalgo was born in Porto, Portugal in 1961.

He is an Assistant Professor in the Dept. of EE of the Fac. of Eng. of Porto and a Senior Researcher at Power Systems Unit of Institute of Engineering Systems and Computers (INESC Porto). He obtained a PhD degree in Electrical Engineering and Computers from the same University in 1995. His interests include pattern recognition, computational intelligence and neural networks and its applications to power systems.

João T. Saraiva was born in Porto, Portugal in 1962.

In 1987, 1993 and 2002 he got his MSc, PhD, and Agregado degrees in Electrical and Computer Engineering from the Faculdade de Engenharia da Universidade do Porto, FEUP, where he is currently Professor.

In 1985 he joined INESC Porto – a private research institute – where he was head researcher or collaborated in several projects related with the development of DMS systems, quality in power systems, and tariffs due for the use of transmission and distribution networks. Several of these projects were developed under consultancy contracts with the Portuguese Electricity Regulatory Agency.