Electricity Market Price Forecasting: Neural Networks versus Weighted-Distance k Nearest Neighbours

Alicia Troncoso Lora[1], José Riquelme Santos[2], Jesús Riquelme Santos[1],
José Luís Martínez Ramos[1] and Antonio Gómez Expósito[1]

[1] Department of Electrical Engineering, University of Sevilla, Sevilla, Spain
[2] Department of Languages and Systems, University of Sevilla, Sevilla, Spain
ali@esi.us.es   riquelme@lsi.us.es   {jesus, camel, age}@us.es

Abstract. In today's deregulated markets, forecasting energy prices is becoming more and more important. In the short term, expected price profiles help market participants to determine their bidding strategies. Consequently, accuracy in forecasting hourly prices is crucial for generation companies (GENCOs) to reduce the risk of over/underestimating the revenue obtained by selling energy. This paper presents and compares two techniques to deal with energy price forecasting time series: an Artificial Neural Network (ANN) and a combined k Nearest Neighbours (kNN) and Genetic algorithm (GA). First, a customized recurrent Multi-layer Perceptron is developed and applied to the 24-hour energy price forecasting problem, and the expected errors are quantified. Second, a k nearest neighbours algorithm is proposed using a Weighted-Euclidean distance. The weights are estimated by using a genetic algorithm. The performance of both methods on electricity market energy price forecasting is compared.

1 Introduction

The new competitive Spanish Electricity Market has been in operation since 1998, and it is mainly based on two separated day-ahead markets [5]:

- The energy market, managed by the Market Operator (MO), where producers and consumers submit production and consumption bids (blocks of hourly energy and the corresponding price in Euros/MWh for each of the 24 hours of the following day). The MO produces a market-clearing price and sets of accepted production and consumption bids for every hour. Additional markets for minor adjustments are also performed on an hourly basis.
- The market for regulation reserves. Once the energy market is cleared, the System Operator establishes the requirements for operating reserves (an hourly margin in MW up and down), that are needed for frequency control, for each of the 24 hours of the following day. The reserve market allocates the margin among the generators that are capable of providing secondary frequency control by using generators' up and down bids which include the offered band (MW) and the price (Euros/MW). A market for additional energy reserves (power that can be provided within 15 minutes for a period of two hours) is also performed.

In this context, forecasting energy prices is extremely important. In the short term, expected price profiles, both in terms of energy and reserve prices, help market participants to determine their bidding strategies. Consequently, accuracy in forecasting hourly energy & reserve prices is crucial for generation companies (GENCOs) to reduce the risk of over/underestimating the revenue obtained by selling energy.

The motivation of this paper is to present and compare two techniques to deal with energy price forecasting time series: an Artificial Neural Network (ANN) and a combined k Nearest Neighbours (kNN) and Genetic algorithm (GA). First, a customized recurrent Multi-layer Perceptron is developed and applied to the 24-hour energy price forecasting problem, and the expected errors are quantified. Second, a k nearest neighbours algorithm is proposed using a Weighted-Euclidean distance. The weights are estimated by using a genetic algorithm.

2 ANN-Based Market Price Forecasting

The ANN approach has been chosen because of its successful performance in the load forecasting problem [1, 2]. Larger errors are expected in this case, however, as the influence of the load level on market clearing prices is only moderate, and other unpredictable factors play an important role in non-perfect oligopolistic markets.

The study reported in this paper is based on the hourly Spanish spot market prices recorded from January 2001 to August 2001. As weekends and holidays constitute separate cases, only data corresponding to working days have been retained and analyzed.

Figure 1a shows the hourly averages and standard deviations of prices for the working days of March 2001, in cents of Euro per kWh.
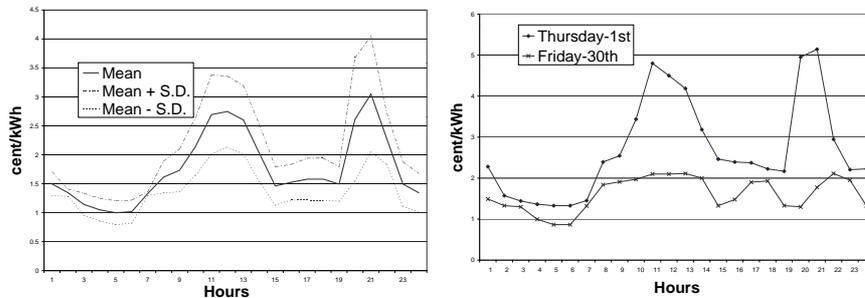

Fig. 1. a) Hourly average of spot market prices for March 2001. **b)** Evolution of energy prices for two days of March 2001.

Average spot prices larger than 2 cent/kWh take place during the morning and evening peak hours (10am-2pm and 8-10pm respectively). Except for a few valley hours, the s.d. of this price exceeds 20% of the mean value, reaching even 40% at 8pm and 9pm.

Figure 1b represents the energy prices for two selected days of March 2001. The significant differences in the prices of the peak hours can not be explained by a change in the demand profile, probably revealing market power mechanisms.

2.1 Structure of the ANN

A brief description of the ANN adopted in this work is provided in this section (the reader interested in ANN background is referred to [3] and [4]).

An ANN is composed of a certain number of perceptrons organized by layers. Each perceptron has several inputs and a single output, whose value is a non-linear function of the inputs. Each perceptron's input is affected by a weighting factor, which must be determined during the training phase. Usually, an ANN is composed of three layers (input, hidden and output), where the outputs of a layer feed the inputs of the next layer.

The two main steps involved in the use of an ANN are:

- Determining its topology, which basically consists of defining the number of perceptrons in the intermediate hidden layer.
- Obtaining the input weighting factors for a given non-linear function (training process).

According to the authors' previous experience, it is decided to feed the ANN with a shifting window of prices comprising 24 hours. This means that the input layer is composed of 24 perceptrons. As far as the number of output perceptrons is concerned, two possibilities have been evaluated.

a) A single output whose value is dictated by the previous 24 hours. Under this scheme, very popular in load forecasting, the window is shifted one hour each time.
b) Twenty four outputs corresponding to the prices of a whole day, whose values are determined by those of the previous day. This implies that the window is shifted 24 hours each time.

Test results have shown better accuracy for scheme b), which is the only one considered in the sequel.

In order to fully define the ANN, it is necessary to determine the number of perceptrons in the intermediate layer and the number of days required for the training process.

Figure 2 shows, for 12, 24 and 36 neurons in the hidden layer, the average forecasting error in the energy price corresponding to the working days of February 2001, versus the number of days used to train the ANN. As can be noted, 20 days are sufficient to train the ANN, the number of neurons not being so important. For the results presented below, 24 neurons in the hidden layer have been used to forecast the spot market energy prices.
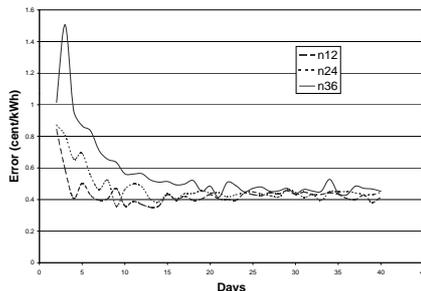
Fig. 2. Average forecasting error in the energy price.

2.2 Results

About two months of the available period (January - February 2001) are used in several experiments to find out and tune the best ANN topology, while the remaining material (March-August 2001) is devoted to check the forecasting errors.

Figure 3a presents the absolute value of the error of the forecasted spot market prices for the two days of March 2001 leading to the largest and smallest average errors.
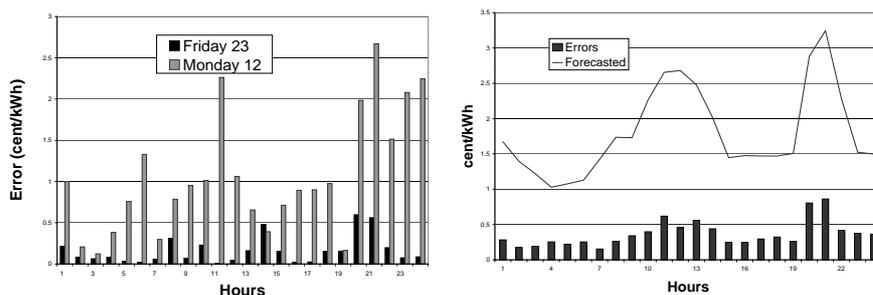

Fig. 3. a) Absolute value of the error of the forecasted energy prices (best and worst prediction). **b)** Hourly average of the forecasted energy prices (March 2001).

	Daily Prices (cent/kWh)			
	Average real price	s. d.	Average absolute error	Maximum error
March-May	2.2588	0.7801	0.3464	2.671
June-August	3.5482	1.0597	0.428	2.0736

Table 1. Forecasting errors for the energy prices.

Figure 3b shows the hourly average of the forecasted energy prices corresponding to March 2001, as well as the resulting prediction errors (obtained by difference with the actual prices of figure 1). Note that the forecasting errors are larger during peak hours.

Finally, table 1 presents the average and s.d. of actual prices, the average absolute value of forecasting errors and the maximum errors for Spring and

Summer seasons. The average error ranges from 12% (Summer) to 15% (Spring) of the hourly average price. As expected, the larger the s.d. of prices the higher the average forecasting error.

3 Nearest Neighbour Based Market Price Forecasting

In this section, a combined kNN [11, 13] and GA [14] for hourly market energy prices forecasting is described. The Genetic Algorithm is used to compute the optimal weights to outweight the most significant hours. The kNN finds the daily market energy prices that are "similar to" the prices of previous days.

The prediction aims at estimating the prices for a certain day from a linear combination of the energy prices of the days that follow the nearest neighbours days.

3.1 K Nearest Neighbours

The prediction of a time series requires the knowledge of some parameters: the time delay, the embedding dimension and the number of nearest neighbours. The optimal determination of these parameters has been studied in the last years [7, 8].

Energy prices are collected every hour and, in consequence, the time delay is an hour.

For short-term forecasting, it has been decided to organize the temporal data with a shifting window of prices comprising 24 hours and consequently the embedding dimension is 24.

As far as the number of steps to predict is concerned, only one possibility has been evaluated: twenty four hours corresponding to the prices of a whole day, whose values are determined by those of the previous day. This possibility implies that the window is shifted 24 hours each time.

Choosing a metric: A time series Y can be considered as a point in a n-dimensional space. Given a sequence query, q, a sequence of Y with the same length as q is searched, z, such that the distance between the sequences is minimum. The choice of the metric to measure similarity between two time series depends mainly on the specific features of the considered series. The most common metric is the square of the Euclidean distance, although other metrics can be used [9]. In this case, the time serie is the electricity market energy prices, and the Weighted-Euclidean distance is preferred because not all the hours of a day have the same influence on the prices of the following day. This distance is defined by

$$d_w^2(q,z) = \sum_{i=1}^{r} w_i \cdot (q_i - z_i)^2 \qquad (1)$$

where $r = \text{length}(q)$ and $w_i \in [0,1]$.

3.2 Prediction

The prediction of stock energy prices for one day $d+1$ is computed through two steps taking into account the weights of all the hours of a day:

1. Calculate the distances between the prices of the day d, P_d, and the preceding points $\{P_{d-1}, P_{d-2}, ...\}$ using (1). Let be $v_1,...,v_k$ the k nearest days to the day d, sorted by closeness.
2. The prediction is:

$$\widehat{P}_{d+1} = \frac{1}{\alpha_1 + ... + \alpha_k} \sum_{l=1}^{l=k} \alpha_l \cdot P_{v_l+1} \qquad (2)$$

where

$$\alpha_l = \frac{d_w(P_d, P_{v_k}) - d_w(P_d, P_{v_l})}{d_w(P_d, P_{v_k}) - d_w(P_d, P_{v_1})} \qquad (3)$$

The two former steps are repeated during the days of the forecasting horizon.

Notice that $0 \leq \alpha_l \leq 1$, i.e., the weight is equal to zero when the considered day is the most distant and one when the considered day is the nearest.

Also, notice that if the k nearest neighbours for a vector P_d are $[P_{v_1}, ..., P_{v_k}]$, where v_i is the ith nearest neighbour, the set of points $[P_{v_1+1}, ..., P_{v_k+1}]$ will usually be the nearest to P_{d+1}.

3.3 Genetic Algorithm

Genetic Algorithms (GA) are computational techniques based on the mechanics of natural selection in which solutions to a problem are coded in a similar way to the genetic structure of biological organisms. Each individual in the population represents a possible solution to the problem, and a fitness factor is assigned to each member of the population, with the best individuals receiving higher fitness factors. A particular GA is characterized by issues such as population size, mutation rates, and selection and new population creation mechanisms.

In the GA used in this paper, each individual is defined by the weight of each hour throughout the whole day, and the fitness function is the inverse of the average square error (ASE) that results from the aforesaid predicting mechanism.

The basic algorithm can be written as follows [12]:

1. An initial population is randomly selected.
2. The fitness factor of each individual is computed.
3. Parents are randomly selected using a tournament selection technique.
4. Parents are crossbred to create new individuals.
5. Several new individuals are mutated.
6. Members with a low fitness factor are replaced with new individuals.
7. If the maximum number of generations has not been reached, go to step 2.

Notice that, in order to compute the fitness factor of each individual, a Nearest Neighbour Algorithm is needed. In consequence, execution times are mainly determined by the kNN module.

Individuals are represented by real row vectors, with columns corresponding to the 24 hours of a day.

The number of days used for the estimation of the parameters $\{w_i\}_{i=1}^r$, and the number of parameters are closely related. If the number of days is too small, the estimated parameters will not be optimal, and if it is too large the estimation of these parameters is computationally very expensive. In consequence, the correct selection of the number of days is quite important to correctly tune the algorithm.

Other implementation issues of the GA are the following:

Initial Population: The initial population is strictly randomly selected.

Fitness function: The fitness function must be able to provide a good measure of the quality of each individual/solution. The kNN+PREDICTION module provides the estimated energy prices according to (2). The fitness function is defined by the inverse of

$$ASE = \frac{1}{n} \sum_{d=1}^{n} (\widehat{P}_d - P_d)^2 \tag{4}$$

where n is the number of days used to calculate the weights. Therefore, in order to calculate the ASE, the k nearest neighbours of the daily prices must be known.

Parent selection and crossover: To produce a new generation, parents are randomly selected using a tournament selection technique that selects the best individuals for reproduction. The probability of a particular individual being selected is in direct proportion to its fitness function, taking into account that the ASE is being minimized.

Regarding the crossover process, the crossover probability has been set to one, i.e., two individuals that have been selected to be parents are always combined to obtain a new individual. Children are obtained by adding the strings that results from random partitions of the row.

Children mutation: Following the crossover process, children are mutated to introduce some new genetic material according to a pre-defined mutation probability. The gene to be mutated is represented by a randomly selected hour, randomly selected individual. The mutation implies adding a random number proportional to the initial value with some previously defined probability.

3.4   Test Results

The kNN+GA has been applied in several experiments to obtain the forecast of Spanish electricity market energy prices. February 2001 has been used to determine the weight of every hour of the day, using the GA described in the former section, and whose main parameters are shown in table 2. The available period of March-August 2001 has been chosen as a test set to check the forecasting errors. The number of neighbours have been considered equal to one in all experiments due to the low influence on the forecasted errors [10].
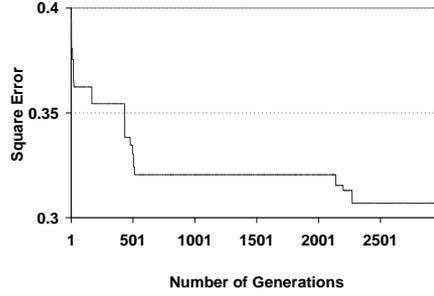
Fig. 4. Evolution of the best individual.

The forecasting errors obtained by the proposed algorithm are compared with the results of an ANN on the same test set described in the first part of this paper.

Figure 4 shows the evolution of the square error of the fittest individual throughout the evolutive process.

Population size	100
Probability of crossover	1
Probability of mutation	0.1
Maximum number of generations	5000

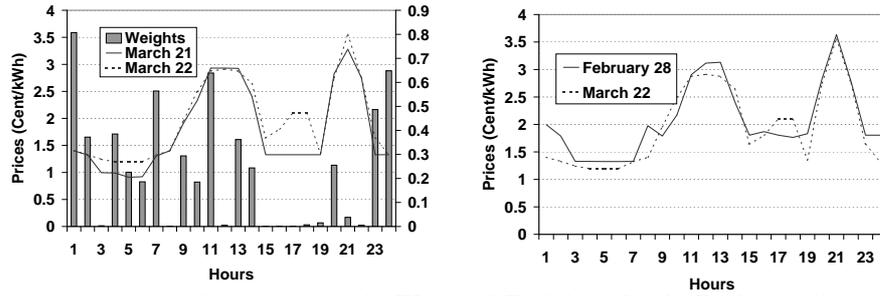**Table 2.** GA's main parameters.

Fig. 5. Nearest neighbour using: **a)** the Weighted-Euclidean, **b)** the Euclidean distance.

Figure 5a represents the relative weights for every hour, along with the energy prices for a selected day, March 22, and its nearest neighbour calculated using the Weighted-Euclidean distance (1). The nearest neighbour corresponds to March 21.

It can be noticed that the hours with larger weights (1am, 7am, 11am, 11pm and 12pm), have almost the same energy prices. However, at 3pm, 4pm, 5pm, 6pm and 7pm, when the differences in prices are greater, the weights are approximately equal to zero.

Figure 5b represents the energy prices of March 22, along with its nearest neighbour calculated using the Euclidean distance. In this case, the nearest neighbour corresponds to February 28. Notice the influence of the chosen distance in the calculation of the nearest neighbour.
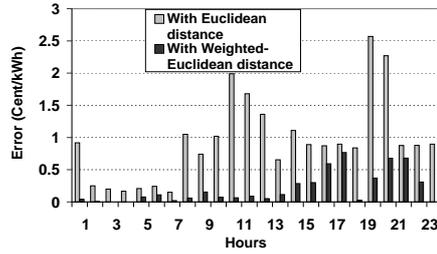
Fig. 6. Absolute value of the error of the forecasted energy prices (March, 23rd).

	Daily Prices (cent/kWh)			
	Average real price	s. d.	Average absolute error	Maximum error
March-May	2.2588	0.2186	0.256	2.13
June-August	3.5482	0.26	0.33	2.36

Table 3. Forecasting errors of energy prices.

Figure 6 presents the absolute value of the error of the forecasted spot market prices for March 23, using the Euclidean distance and the Weighted-Euclidean distance. As expected, the forecasting errors resulting for the selected day are strongly influenced by the calculation of the nearest neighbour.

Finally, table 3 presents the average and s.d. of actual prices, the average absolute value of forecasting errors and the maximum errors for Spring and Summer seasons. Note that, the average error, using an ANN, ranges from 12% (Summer) to 15% (Spring) of the hourly average price, while the average error obtained, by means of a kNN combined with a GA, ranges from 9% (Summer) to 11% (Spring).

Table 4 summarizes the results obtained using both techniques.

4   Conclusions

Optimal bidding strategies are relevant in a competitive market, the computation of a good forecasted price profiles being crucial for the generation companies. In this sense, this paper presents two techniques to deal with forecasting time series: an Artificial Neuronal Network and a combined k Nearest Neighbours and Genetic Algorithm. Both algorithms have been applied to the 24-hour energy price forecasting problems, using real data of the Spanish energy markets, and their performance have been compared. The proposed Weighted-distance kNN reveals much lower forecasting errors for the energy prices that the ANN proposed.

	March-May		June-August	
	kNN+GA	ANN	kNN+GA	ANN
s.d.	0.2186	0.7801	0.26	1.0597
Average absolute error	0.256	0.3464	0.33	0.428
Maximum error	2.13	2.671	2.36	2.0736
Average Relative error (%)	11	15	9	12

Table 4. Comparison of predicted daily prices (Cent/kWh).

Acknowledgments

References

1. R. Lamedia, A. Prudenzi, M. Sforna, M. Caciotta, V. Orsolini Cencellli: A Neural Network Based Technique for Short-Term Forecasting of Anomalous Load Periods. IEEE Trans. on Power System, Vol. 11, pp. 1749-1755. 1996.
2. A. S. Alfuhaid and M. A. El-Sayed: Cascaded Artificial Neural Network for Short-Term Load Forecasting. IEEE Trans. on Power System, Vol. 12, pp. 1524-1529. 1997.
3. M. El-Sharkawi and D. Niebur: Tutorial course on Artificial Neural Networks with Applications to Power Systems. IEEE Catalog Number 96, Tp 112-0, 1996.
4. J. Riquelme, J.L. Martínez, A. Gómez and D. Cros Goma: Possibilities of Artificial Neural Networks in Short-Term Load Forecasting. Proceedings of the IASTED International Conference Power and Energy Systems. pp. 165-170, Marbella, Spain. September 2000.
5. A. Canoyra, C. Illán, A. Landa, J.M. Moreno, J.I. Pérez Arriaga, C. Sallé and C. Solé: The Hierarchical Market Approach to the Economic and Secure Operation of the Spanish Power System. Bulk Power System Dynamic and Control IV, August 24-28, Santorini, Greece.
6. A. J. Conejo, J. Contreras, J.M. Arroyo and S. de la Torre: Optimal Response of an Oligopolistic Generating Company to a Competitive Pool-Based Electric Power Market. To appear in IEEE Trans. on Power System.
7. A. W. Jayawardena, W. K. Li and P. Xu: Neighbourhood selection for local modelling and prediction of hydrological time series. Journal of Hydrology 258, 40-57.
8. James McNames: Local modelling optimization for time serie prediction. Submitted to Physical Review E.
9. D. Berndt and J. Clifford: Using dynamic time warping to find patterns in time series. AAAI-94 Workshop on Knowledge in Databases (KDD-94), Seattle, Washington.
10. R.C. Holte: Very simple classification rules perform well on most commonly used datasets. Machine learning, 11:63-91, 1993.
11. B.V. Dasarathy (Ed): Nearest neighbour (NN) Norms: NN pattern classification techniques. IEEE Computer Society Press, 1991.
12. Z. Michalewicz: Genetic Algorithms + Data Structures = Evolution Programs. Second Edition, Springer-Verlag, 1994.
13. D.W. Aha, D. Kibler and M.K. Albert: Instance-based learning algorithms. Machine Learning 6:37-66, 1991.
14. D. E. Goldberg: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989.