Abstract—A novel hybrid approach, combining particle swarm optimization and adaptive-network based fuzzy inference system, is proposed in this paper for short-term electricity prices prediction in a competitive market. Results from a case study based on the electricity market of mainland Spain are presented. Finally, conclusions are duly drawn.

Index Terms—Electricity market, fuzzy logic, neural networks, price forecasting, swarm optimization.

I. INTRODUCTION

All over the world, the electricity industry is converging toward a competitive framework. In this competitive framework, short-term electricity prices prediction is required by producers and consumers to derive their bidding strategies to the electricity market. Deregulation brings electricity prices uncertainty, placing higher requirements on forecasting.

Therefore, price forecasting tools are essential for all market participants for their survival under deregulated environment [1].

In most competitive electricity markets the series of prices presents the following features: high frequency, non-constant mean and variance, daily and weekly seasonality, calendar effect on weekend and public holidays, high volatility and high percentage of unusual prices [2].

Price forecast is a key issue in competitive electricity markets, and several techniques have been tried out in this task. In general, hard and soft computing techniques could be used to predict electricity prices.

The hard computing techniques include auto regressive integrated moving average (ARIMA) [3], wavelet-ARIMA [4], and mixed model [5] approaches.

Usually, an exact model of the system is required, and the solution is found using algorithms that consider the physical phenomena that govern the process. Although these approaches can be very accurate, they require a lot of information, and the computational cost is very high.

The soft computing techniques include neural networks (NN) [6], fuzzy neural networks (FNN) [7], weighted nearest neighbors (WNN) [8], adaptive wavelet neural network (AWNN) [9], hybrid intelligent system (HIS) [10], and cascaded neuro-evolutionary algorithm (CNEA) [11] approaches. A combination of neural networks with wavelet transform (NNWT) has also been recently proposed [12].

Usually, an input-output mapping is learned from historical examples, thus there is no need to model the system. Hence, these approaches can be much more efficient computationally and as accurate as the first ones, if the correct inputs are considered [13].

In this paper, a novel hybrid approach is proposed for short-term electricity prices prediction. The proposed approach is based on the combination of particle swarm optimization (PSO) and adaptive-network based fuzzy inference system (ANFIS).

Our HPA (hybrid PSO-ANFIS) approach is examined on the electricity market of mainland Spain, commonly used as the test case in several price forecasting papers [3]-[12]. It has been concluded that the Spanish market has a hard nonlinear behavior and time variant functional relationship [4], [7]. So, this market is a real world case study with sufficient complexity.

This paper is organized as follows. Section II presents the proposed approach to predict electricity prices. Section III provides the different criterions used to evaluate the forecasting accuracy. Section IV provides the results from a case study based on the electricity market of mainland Spain. Section V outlines the conclusions.

II. PROPOSED APPROACH

The proposed approach is based on a combination of PSO and ANFIS. The PSO is used to improve the performance of ANFIS, tuning the membership functions required to achieve a lower error.

A. Particle Swarm Optimization

Particle swarm optimization is a heuristic approach first proposed by Kennedy and Eberhart in 1995 [14] as an evolutionary computational method developed for dealing with the optimization of continuous and discontinuous function decision making. The PSO algorithm is based on the biological and sociological behavior of animals such as schools of fish and flocks of birds searching for their food [15].

PSO is a population-based search method where each potential solution is represented as a particle in a population (called swarm).
Particles change their position in a multidimensional search space until equilibrium or optimal state has been reached or until computation limitations are exceeded.

Empirical evidence has been accumulated to show that the algorithm is a useful tool for optimization [16]. PSO has been applied to many optimization problems, for instance [17].

Consider an optimization problem of $D$ variables. A swarm of $N$ particles is initialized in which each particle is assigned a random position in the $D$-dimensional hyperspace such that each particle’s position corresponds to a candidate solution for the optimization problem.

Let $x$ denote a particle’s position (coordinate) and $v$ denote the particle’s flight velocity over a solution space. Each individual $x$ in the swarm is scored using a scoring function that obtains a fitness value representing how good it solves the problem.

The best previous position of a particle is $P_{best}$. The index of the best particle among all particles in the swarm is $G_{best}$. Each particle records its own personal best position ($P_{best}$), and knows the best positions found by all particles in the swarm ($G_{best}$). Then, all particles that fly over the $D$-dimensional solution space are subject to updated rules for new positions, until the global optimal position is found.

Velocity and position of a particle are updated by the following stochastic and deterministic update rules

$$v_i(t) = \omega v_i(t-1) + \rho_1 (x_{P_{best}} - x_i(t)) + \rho_2 (x_{G_{best}} - x_i(t)) \quad (1)$$

$$x_i(t) = x_i(t-1) + v_i(t), \quad t = t + 1 \quad (2)$$

where $\omega$ is an inertia weight, $\rho_1$ and $\rho_2$ are random variables. The random variables are defined as $\rho_1 = r_1 c_1$ and $\rho_2 = r_2 c_2$, with $r_1, r_2 \sim U(0,1)$, and $C_1$ and $C_2$ are positive acceleration constants.

Fig. 1 illustrates a search mechanism of a PSO technique using the velocity update rule (1) and the position update rule (2).

![Fig. 1. Updating the position mechanism of PSO.](image_url)

Acceleration constants $C_1$ and $C_2$ represent the weights of the stochastic acceleration terms that push a particle toward $P_{best}$ and $G_{best}$, respectively. Small values allow a particle to roam far from target regions. Conversely, large values result in the abrupt movement of particles toward target regions. In this work, constants $C_1$ and $C_2$ are both set at 2.0, following the typical practice in [18].

Suitable correction of inertia $\omega$ in (2) provides a balance between global and local explorations, thereby reducing the number of iterations when finding a sufficiently optimal solution. An inertia correction function called “inertia weight approach (IWA)” is utilized in this work [18]. During the IWA, the inertia weight $\omega$ is modified according to the following equation

$$\omega = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{Itr_{max}} Itr \quad (3)$$

where $\omega_{max}$ and $\omega_{min}$ are the initial and final inertia weights, $Itr_{max}$ is the maximum number of iteration, and $Itr$ is the current number of iteration.

### B. ANFIS

NN are simple, but powerful and flexible tools for forecasting, provided that there are enough data for training, an adequate selection of the input-output samples, an appropriated number of hidden units and enough computational resources available. Also, NN have the well-known advantages of being able to approximate any nonlinear function and being able to solve problems where the input-output relationship is neither well defined nor easily computable, because NN are data-driven. Multi-layered feedforward NN are specially suited for forecasting, implementing nonlinearities using sigmoid functions for the hidden layer and linear functions for the output layer [6].

Just like NN, a fuzzy logic system is a nonlinear mapping of an input vector into a scalar output, but it can handle numerical values and linguistic knowledge. In general, a fuzzy logic system contains four components: fuzzifier, rules, inference engine, and defuzzifier. The fuzzifier converts a crisp input variable into a fuzzy representation, where membership functions give the degree of belonging of the variable to a given attribute. Fuzzy rules are of the type “if–then”, and can be derived from numerical data or from expert linguistic. Mamdani and Sugeno inference engines are two of the main types of inference mechanisms.

The Mamdani engine combines fuzzy rules into a mapping from fuzzy input sets to fuzzy output sets, while the Takagi–Sugeno type relates fuzzy inputs and crisp outputs. The defuzzifier converts a fuzzy set into a crisp number using the centroid of area, bisector of area, mean of maxima, or maximum criteria.

NN have the advantage over the fuzzy logic models that knowledge is automatically acquired during the learning process. However, this knowledge cannot be extracted from the trained network behaving as a black box. Fuzzy systems, on the other hand, can be understood through their rules, but these rules are difficult to define when the system has too many variables and their relations are complex [13].
A combination of NN and fuzzy systems has the advantages of each of them. In a neuro-fuzzy system, neural networks extract automatically fuzzy rules from numerical data and, through the learning process, the membership functions are adaptively adjusted.

ANFIS is a class of adaptive multi-layer feedforward networks, applied to nonlinear forecasting where past samples are used to forecast the sample ahead. ANFIS incorporates the self-learning ability of NN with the linguistic expression function of fuzzy inference [19].

The ANFIS architecture is shown in Fig. 2. The ANFIS network is composed of five layers. Each layer contains several nodes described by the node function. The node function is described next. Let \( O^j_i \) denote the output of the \( i \)th node in layer \( j \).

In layer 1, every node \( i \) is an adaptive node with node function

\[
O^1_i = \mu A_i (x), \quad i = 1, 2
\]

or

\[
O^1_i = \mu B_{i-2} (y), \quad i = 3, 4
\]

where \( x \) (or \( y \)) is the input to the \( i \)th node and \( A_i \) (or \( B_{i-2} \)) is a linguistic label associated with this node.

Thus, \( O^1_i \) is the membership grade of a fuzzy set \( A \) (= \( A_1, A_2, B_1, \) or \( B_2 \)) and it specifies the degree to which the given input \( x \) (or \( y \)) satisfies the quantifier \( A \). The membership functions for \( A \) and \( B \) are usually described by generalized bell functions, e.g.

\[
\mu A_i (x) = \frac{1}{1 + \left| \frac{x - q_i}{p_i} \right|^r_i}
\]

where \( \{p_i, q_i, r_i\} \) is the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions on linguistic label \( A_i \).

In fact, any continuous and piecewise differentiable functions, such as triangular-shaped membership functions, are also qualified candidates for node functions in this layer [20]. Parameters in this layer are referred to as premise parameters.

In layer 2, each node \( \prod \) multiplies incoming signals and sends the product out

\[
O^2_i = w_i = \mu A_i (x) \mu B_j (y), \quad i = 1, 2
\]

Hence, each node output represents the firing strength of a rule.

In layer 3, each node \( N \) computes the ratio of the ith rule’s firing strength to the sum of all rules’ firing strengths

\[
O^3_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2
\]

The outputs of this layer are called normalized firing strengths.

In layer 4, each node computes the contribution of the ith rule to the overall output

\[
O^4_i = w_i z_i = \overline{w_i} (a_i x + b_i y + c_i), \quad i = 1, 2
\]

where \( \overline{w_i} \) is the output of layer 3 and \( \{a_i, b_i, c_i\} \) is the parameter set. Parameters of this layer are referred to as consequent parameters.

In layer 5, the single node \( \sum \) computes the final output as the summation of all incoming signals

\[
O^5 = \sum_i w_i z_i = \sum_i \frac{w_i}{w_j} z_i
\]

Thus, an adaptive network is functionally equivalent to a Sugeno-type fuzzy inference system.

In this paper, ANFIS employs PSO method to adjust the parameters of the membership functions, as in [21]. The PSO techniques have the advantage of being less computationally expensive for a given size of network topology. The membership functions considered in this study are triangular-shaped.

### III. Forecasting Accuracy Evaluation

To evaluate the accuracy of the HPA approach in forecasting electricity prices, different criterions are used. This accuracy is computed in function of the actual prices that occurred. The mean absolute percentage error (MAPE) criterion, the sum squared error (SSE) criterion, and the standard deviation of error (SDE) criterion, are defined as follows.

The MAPE criterion is given by

\[
MAPE = \frac{100}{N} \sum_{h=1}^{N} \frac{\hat{p}_h - p_h}{\overline{p}}
\]

where \( \hat{p}_h \) and \( p_h \) are respectively the forecasted and actual electricity prices at hour \( h \), \( \overline{p} \) is the average price of the forecasting period and \( N \) is the number of forecasted hours.
Electricity price can rise to tens or even hundreds of times of its normal value at particular hours, and it may drop to zero at other hours. Hence, the average price is used in (11) to avoid the adverse effect of prices close to zero [22].

The SSE criterion is given by

$$SSE = \sum_{h=1}^{N} (\hat{p}_h - p_h)^2$$

The SDE criterion is given by

$$SDE = \sqrt{\frac{1}{N} \sum_{h=1}^{N} (e_h - \bar{e})^2}$$

$$e_h = \hat{p}_h - p_h$$

$$\bar{e} = \frac{1}{N} \sum_{h=1}^{N} e_h$$

where $e_h$ is the forecast error at hour $h$ and $\bar{e}$ is the average error of the forecasting period.

IV. NUMERICAL RESULTS

The HPA approach is applied to predict next-week prices in the electricity market of mainland Spain. Price forecasting is computed using historical data of year 2002 for the Spanish market. For the sake of simplicity and clear comparison, no exogenous variables are considered. Also, for the sake of a fair comparison, the same test weeks as in [3]-[12] are selected, which correspond to the four seasons of year 2002. To build the forecasting model, the hourly historical price data of the 42 days previous to the day of the week whose prices are to be forecasted have been considered.

Numerical results with the HPA approach are shown in Figs. 3 to 6 respectively for the winter, spring, summer and fall weeks. Each figure shows the actual prices, solid line, together with the forecasted prices, dash-dot line.

Table 1 presents the values for the criterions to evaluate the accuracy of the proposed approach in predicting electricity prices. The first column indicates the week, the second column presents the MAPE, the third column presents the square root of the SSE, and the fourth column presents the SDE.

<table>
<thead>
<tr>
<th>Week</th>
<th>MAPE</th>
<th>$\sqrt{SSE}$</th>
<th>SDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>3.65</td>
<td>27.39</td>
<td>1.38</td>
</tr>
<tr>
<td>Spring</td>
<td>4.19</td>
<td>31.38</td>
<td>1.54</td>
</tr>
<tr>
<td>Summer</td>
<td>6.76</td>
<td>51.95</td>
<td>2.93</td>
</tr>
<tr>
<td>Fall</td>
<td>6.53</td>
<td>36.86</td>
<td>1.85</td>
</tr>
</tbody>
</table>

The HPA approach presents good forecasting accuracy. The MAPE has an average value of 5.28%. For comparison purposes, the average MAPE values for ARIMA [3], NN [6] and FNN [7] approaches were 9.96%, 8.91%, and 7.52%, respectively. Moreover, the average computation time is less than 5 seconds, using MATLAB on a PC with 1 GB of RAM and a 2.0-GHz-based processor.

An average value of 5.32% for the MAPE has been recently reported using a CNEA [11] approach. However, this approach has a major drawback: the computation time of about 40 minutes. Hence, the proposed HPA approach presents a good trade-off between forecasting accuracy and computation time.
VI. REFERENCES


V. CONCLUSIONS

A novel hybrid approach is proposed in this paper for short-term electricity prices prediction. The proposed approach is based on the combination of particle swarm optimization and adaptive-network-based fuzzy inference system. The MAPE has an average value of 5.28%, while the average computation time is less than 5 seconds. Hence, the proposed approach presents good forecasting accuracy with a negligible computation time.
VII. BIOGRAPHIES

H. M. I. Pousinho (S’09) received the M.Sc. degree from the University of Beira Interior, Covilha, Portugal, in 2009.

He is currently a Ph.D. student at the University of Beira Interior, in collaboration with the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal. His research interests include hydro scheduling, unit commitment, and price forecasting.

V. M. F. Mendes received the M.Sc. and Ph.D. degrees from the Instituto Superior Técnico, Lisbon, Portugal, in 1987 and 1994, respectively.

He is currently a Coordinator Professor with Aggregation at the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal. His research interests include hydrothermal scheduling, optimization theory and its applications, and renewable energies. He is the author or co-author of more than 110 scientific papers presented at international conferences or published in reviewed journals.

J. P. S. Catalão (M’04) received the M.Sc. degree from the Instituto Superior Técnico, Lisbon, Portugal, in 2003 and the Ph.D. degree from the University of Beira Interior, Covilha, Portugal, in 2007.

He is currently an Assistant Professor at the University of Beira Interior. His research interests include hydro scheduling, unit commitment, price forecasting, wind energy systems, and electricity markets. He is the author or co-author of more than 70 scientific papers presented at international conferences or published in reviewed journals.

Dr. Catalão is an Associate Editor for the International Journal of Power and Energy Systems, and a Member of the Editorial Board of Electric Power Components & Systems. Also, he is a reviewer for IEEE TRANSACTIONS ON POWER SYSTEMS, and other IEEE and International Journals.