Computational Intelligence in Electrical Engineering, 10 th year, No. 1, 2019
pp. 63-72
Doi: 10.22108/isee.2019.113937.1166
Research Article

Short Term Electricity Price Forecasting by Hybrid Mutual InformationANFIS-PSO Approach

Yaser Raeisi Gahrooei[1], Rahmat-Allah Hooshmand[2]

[1]Isfahan Electric Power Distribution Company, Isfahan, Iran
y.raeisi@eepdc.ir
[2]Department of Electrical Engineering, Faculty of Electrical Engineering,
University of Isfahan, Isfahan, Iran
Hooshmand_r@eng.ui.ac.ir

Abstract :

In a competitive electricity market, an accurate short term price forecasting is essential for all the participants in market as a risk management technique. For both spot markets and long-term contracts, price forecastis necessary to develop bidding strategies or negotiation skills in order to maximize benefit. This paper proposes an efficient tool for short-term electricity price forecasting with a simple model and acceptable computation time by combining several intelligent methods. Using inference, Adaptive Network-based Fuzzy Inference System (ANFIS) is used to determine the nonlinear relation between large quantities of input variables and forecasted price (output variable). To decrease the complexity and improve the accuracy, Mutual Information (MI) technique is used to efficiently select the best set of input variables which have important information concerning forecasted price. Moreover, Particle Swarm Optimization (PSO) algorithm with new strategy in choosing the particles is adopted to tune ANFIS parameters more precisely. To evaluate the accuracy and performance, the proposed hybrid Mutual Information-ANFIS-PSO (MIAP) methodology is implemented on the real world case study of Spanish electricity market. The results show the great potential of this proposed method in fast and accurate short-term price forecasting in comparison with some of the previous price forecasting techniques.

Keywords: ANFIS, Electricity Market, Mutual Information Technique, Short Term Price Forecasting, Swarm Optimization.

1. Introduction

With the introduction of restructuring into the electric power industry, the price of electricity has become one of the most essential subjects in the electricity market. Accurate price forecasts are crucial for producers to maximize their profits by bidding effectively and for bulk electricity customers to maximize their utilities by load scheduling optimally. Therefore, in a competitive electricity market, efficient and robust price forecasting tools are important for all market participants [1]. Price series exhibit a great complexity. In most of competitive electricity markets, the series of prices presents the special features, such as high frequency, no constant mean and variance, daily and weekly seasonality, calendar effect on weekend and public holidays, high volatility and high percentage of unusual prices , hard nonlinear behavior, etc. that make an accurate price forecasting be a challenging task [2].

Due to the importance and complexity of short term price forecasting, it has become an important research area in electrical engineering, and several techniques have been tried out in this task, namely, hard and soft computing techniques. The hard computing techniques, based on time series [3, 4], include Auto Regressive Integrate Moving (ARIMA) [5, 6] and Generalized Auto Regressive Conditional Heteroscedasticity (GARCH) [7] techniques. Regardless of their high computational cost, these approaches are well established to be very accurate; however, due to the use of linear modeling, most of them cannot predict the rapid changes of the price properly [8].

Fortunately, with the development of artificial intelligence,various artificial intelligence forecasting algorithms includingartificial Neural Networks (NN) [2, 8-12], Fuzzy Neural Networks (FNN) [13], Weighted Nearest Neighbors (WNN) [14] and Hybrid

[1] Submission date: 11, 11, 2018
Acceptance date: 22, 02, 2019
Corresponding author: Rahmat-Allah Hooshmand, Department of Electrical Engineering, Faculty of Electrical Engineering, University of Isfahan, Isfahan, Iran

intelligent System (HS) [15] have been developed in price forecasting. The efficiency of soft computing techniques has close relation with the appropriate selection of input variables and the correct tuning of adjustable parameters.To overcome the instability of the NNs' final results, some optimization algorithms, including firefly algorithm [1], whale optimization algorithm [16] and cuckoo search algorithm [17] are employed to optimize the initial weight and threshold of NNs.

To develop hybrid forecasting models, WT [18] and Variational Mode Decomposition (VMD) [19] have been incorporated in the forecasting model for identifying and extracting the main features of the price signal. In WT based models, first WT is applied to decompose the price series, then prediction is made in the wavelet domain using a predictive model like regression model, time series model [20], ANFIS [21], or NN [8, 22-24]; finally the inverse WT is applied to obtain the actual predicted value in time domain. During the process, there may be some loss of information. Moreover, due to the characteristics of some high frequency detailed series, these series cannot be predicted accurately [8].

In ANFIS based approaches, a fuzzy inference system is used for training the ANN [25]. Although it has better performance relative to the ANN, the large amount of training data limit its performance. By considering this fact that an increase in the number of ANFIS input variables significantly increases the number of training data, the selection of appropriate ANFIS input variables is one of the most important aspects of using ANFIS in price forecasting.

In order to make ANFIS a practical approach with high accuracy in forecasting, adopting fast and precise methods in tuning ANFIS parameters is very important. In [25], several techniques are introduced in order to train ANFIS parameters. However, the convergence of parameters by these methods is very slow and depends on the initial values of parameters [26]. In [26], a method based on PSO algorithm is proposed to learn the premise and consequent parameters of ANFIS and is used in [21, 27] for electricity price forecasting. In this method, all parameters are initialized randomly in the first stage and then are being updated using PSO algorithm. In each iteration, one of the parameters is tuned and the others are kept constant. Since ANFIS structure has so many parameters to be trained, this strategy is time consuming and even impossible to reach so acceptable convergence, especially for volatile signals like electricity price.

This paper presents a hybrid forecasting model to improve the accuracy of the short term electricity price forecasting. ANFIS is in the heart of the proposed approach. To limit the number of training ANFIS parameters, MI technique is used to select an appropriate set of input variables. Moreover, PSO algorithm is used to tune all parameters of ANFIS. This paper also adopts a new strategy to form the particles of PSO algorithm by considering premise and consequent parameters in a single matrix.

The accuracy of the hybrid MIAP approach is examined by means of real data from the electricity market of mainland Spain, commonly used as the test case in several price forecasting papers. The proposed method is compared with some of the previous price forecasting techniques to show its effectiveness regarding the accuracy and computation time with simplicity of modeling.

2. Adaptive Network-Based Fuzzy Inference System and PSO Algorithm

In this section, ANFIS structure and PSO algorithm are described in detail.

2.1. Adaptive-network-based Fuzzy Inference System (ANFIS)

ANFIS is a kind of neural network that is based on Takagi–Sugeno fuzzy inference system. Since it integrates both neural networks and fuzzy logic principles, it has the potential to capture the benefits of both in a single frame work. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions. The ANFIS architecture is shown in Fig.1. The circular nodes represent nodes that are fixed whereas the square nodes are nodes that have parameters to be learned.

Layer 1: The output of each node is:

$$O_{1,i} = \mu_{A_i}(x) \quad \text{for } i = 1,2 \tag{1}$$

$$O_{1,i} = \mu_{B_{i-2}}(y) \quad \text{for } i = 3,4 \tag{2}$$

So, the $O_{1,i}(x)$ is essentially the membership grade for x andy. The membership functions could be anything, but for illustration purposes, we will use the bell-shaped function given by:

$$\mu_{A_i}(x) = \exp\left\{-\left(\frac{x-c_i}{a_i}\right)^2\right\} \tag{3}$$

Where a_i, c_i are referred to as ANFIS premise parameters.

Layer 2: Every node in this layer is fixed. This is where the t-norm is used to 'AND' the membership grades - for example the product:

$$O_{2,i} = w_i = \mu_{A_1}(x)\mu_{B_i}(y) \quad i = 1,2 \tag{4}$$

$$O_{2,i} = w_i = \mu_{A_2}(x)\mu_{B_{i-2}}(y) \quad i = 3,4 \tag{5}$$

Layer 3: Contains fixed nodes which calculate the ratio of the firing strengths of the rules:

$$O_{3,i} = \overline{w}_i = \frac{w_i}{\sum_{i=1}^{i=4} w_i} \quad (6)$$

Layer 4: The nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i \left(p_1^i x + p_2^i y + p_3^i\right) \quad (7)$$

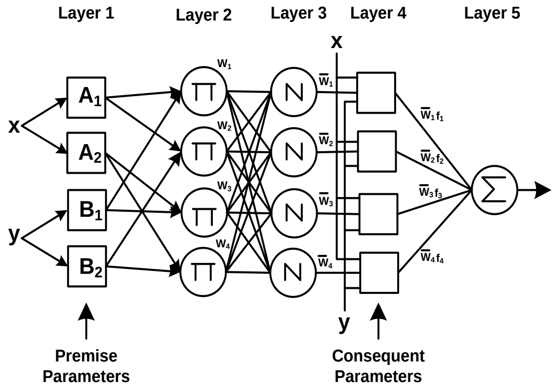

Fig. 1. An ANFIS Architecture for a Two Rule Sugeno System

The parameters in this layer $\left(p_1^i, p_2^i, p_3^i\right)$ are to be determined and are referred to as the consequent parameters.

Layer 5: There is a single node here that computes the overall output *O*:

$$O_{5,i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (8)$$

This, then, is how, typically, the input vector is fed through the network layer by layer.

2.2. PSO Algorithm

PSO is multi-agent search technique that traces its evolution to the emergent motion of a flock of birds searching for food. PSO uses a number of particles that constitute a swarm. The modified velocity and position of search particle can be calculated as shown in the following formulas:

$$v_{j,g}^{t+1} = w.v_{j,g}^t + c_1.rand.(pbest_{j,g} - x_{j,g}^t) + c_2.Rand.(gbest_g - x_{j,g}^t) \quad (9)$$

$$x_{j,g}^{t+1} = x_{j,g}^t + c_1.v_{j,g}^{t+1} \quad (10)$$
$$j = 1,2,\dots,n$$
$$g = 1,2,\dots,m$$

Where n is the number of particles in a group; $m$is number of members in a particle, t is the pointer of iterations (generations); $v_{j,g}^t$ is the velocity of particle j at iteration t and $v_g^{min} \le v_{j,g}^t \le v_g^{max}$; c_1, c_2 are acceleration constant, w is the inertia weight factor; rand, Rand are random number between [0, 1]; $x_{j,g}^t$ is current position of particle j at iteration t; $pbest_j$ is best position of particle j and $gbest_g$ is best position of the particles.

The parameter v_g^{max} determined the resolution with which regions to be searched between the present position and the target position.

The constants c_1, c_2 represent the stochastic acceleration terms that pull each particle toward pbest and gbest positions, respectively ($c_1 = c_2 = 2$). Inertia weight w is the balance factor between global and local exploration and often decrease linearly from about 0.9 to 0.4 versus iteration. In general, the inertia weight is according to the equation (11).

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}}.iter \quad (11)$$

3. Proposed approach

There is a nonlinear relation between the input and output variables in the short term price forecasting. Therefore, the ANFIS method can be employed to determine the nonlinear relation between large quantities of input and output (forecasted price) data by using inference. Although inference is used in the ANFIS approach, this system is not useful with a large quantity of training data. The more input variables are, the more training data will be. To cope with this weakness, the proposed approach limits the number of ANFIS training data by ignoring unimportant input variables. For this purpose, MI technique is used to select the best input variables which have important information concerning output and discard the others. Adapting a precise and fast method in tuning ANFIS parameters is so important to make ANFIS a practical method with high accuracy in price forecasting. The proposed approach uses PSO algorithm in a modified way to train ANFIS parameters that converge fast and increases the forecasting accuracy significantly.

Our strategy is shown in Fig. 2 and can be summarized as follows; first the candidate inputs, include lagged values of prices up to 200 hours ago, and output variable are linearly normalized in the range of [0, 1]. The MI technique selects the best input variables of ANFIS network. By the selected inputs, training and validation samples are constructed based on the data of 49 days ago. Then, PSO algorithm in a modified way tunes the premise and consequent parameters of ANFIS network with high accuracy and less computation time. Finally, next-week prices will be forecasted hourly.

In this section, the structure of ANFIS is mentioned generally and then the role of PSO algorithm in tuning ANFIS parameter and MI technique in selecting input variables are described.

3.1. Training ANFIS Parameters with PSO

Recent works proposed different approaches in order to train ANFIS parameters. Reference [25] introduces Gradient decent based methods and sequential Least Square Error (LSE) method to update the parameters of ANFIS structure. However, since the convergence of ANFIS parameters by these methods is very slow and depends on initial value of parameters, finding the best learning rate is very difficult. Therefore, in [26] a method based on PSO algorithm is proposed to learn the premise and consequent parameters of ANFIS.

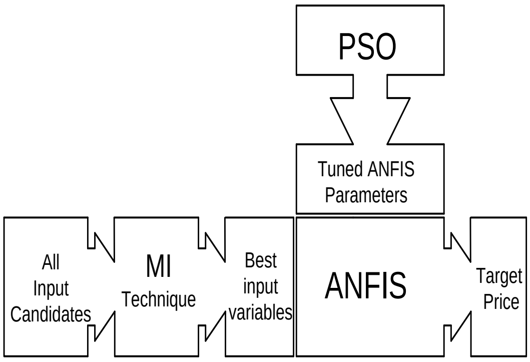

Fig.2. Structure of the Proposed MIAP Strategy

In this method, ANFIS parameters are initialized randomly in the first stage and then are being updated using PSO algorithm. In each iteration, one of the parameters is being trained. In other words, in the first iteration, for example, p_1^i is trained and the other parameters are held fix, then in the second iteration, p_2^i is trained and the other parameters are held fix. After training all parameters again, the first parameter trains and so on. It means that every change in a parameter at each iteration changes the output value and consequently the error function. Since ANFIS has so many parameters, it is so time consuming to reach acceptable convergence if parameters are trained by this strategy. On the other hand, each of ANFIS parameter has mutual interaction with ANFIS output and they aren't independent. This paper considers premise and consequent parameters in a single matrix P. In each iteration, PSO algorithm trains all the entries of matrix (parameters) simultaneously instead of training them individually.

To explain how to construct matrix P (Eq. 11), suppose we have n bell shaped membership functions for each input. There are two parameters at each membership function (its mean and variance) and then, there are 2n parameters. If we have m input variables for ANFIS, there are $2n \times m$ parameters in premise part of it and is shown in upper part of matrix P. For instance, in Fig. 1 by $m = 2$ input variables and $n = 2$ membership function there are $2 \times 2 \times 2 = 8$ premise parameters. In consequent part (layer4), there are n^m nodes and there are $m + 1$ parameters for each node (i.e. for two inputs, the output of ith node is $p_1^i x + p_2^i y + p_3^i$). So there are $(m + 1)n^m$ parameters in consequent part and is shown in downer part of matrix P. For instance, in Fig. 1 there are $(2 + 1)2^2 = 12$ consequent parameters.

Now, we can define a particle in PSO algorithm as the following equation:

$$P = \left[\begin{array}{ccccc} c_1^1 & c_1^2 & \cdots & c_1^m & 0 \\ a_1^1 & a_1^2 & \cdots & a_1^m & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ c_i^1 & c_i^2 & \cdots & c_i^m & 0 \\ a_i^1 & a_i^2 & \cdots & a_i^m & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ c_n^1 & c_n^2 & \cdots & c_n^m & 0 \\ a_n^1 & a_n^2 & \cdots & a_n^m & 0 \\ \hline p_1^1 & p_2^1 & \cdots & p_m^1 & p_{m+1}^1 \\ \vdots & \vdots & & \vdots & \vdots \\ p_1^{n^m} & p_2^{n^m} & \cdots & p_m^{n^m} & p_{m+1}^{n^m} \end{array}\right] \quad (12)$$

In PSO algorithm, N initial populations that contain N different matrices in the form of P are generated randomly. Selecting initial population is a key point. Some particles can be selected by random generating mean and the variance values of membership functions and consequent parameters separately and testing every possible permutation to get minimum error function. After generating initial population, the ability of each particle is evaluated. If its value is better than current pbest or gbest, that particle will be considered as pbestorgbest. For next iteration, the position and velocity of each particle are updated based on Eq. 9 and Eq. 10. This procedure is continued for all iterations. The last gbest includes trained ANFIS parameters.

3.2. Selecting Proper ANFIS Input Variables by Mutual Technique

ANFIS is a classifying system that maps input features onto output classes. Some of features have important information concerning output and the others contain little information. To avoid the curse of complexity and dimensionality, we try to select features that contain more information about the outputs. Mutual information is a technique that we use to preprocess our data and discard unimportant features.

The entropy $H(X)$ of a discrete random variable with values $X_1, X_2, \cdots, X_n$ and probabilities $P(X_1)$, $P(X_2), \cdots, P(X_n)$, respectively, is defined as follows [28]:

$$H(X) = -\sum_{i=1}^{n} P(X_i)\log_2\big(P(X_i)\big) \quad (13)$$

For two discrete random variables X and Y with their joint pdf $p(X, Y)$, the joint entropy of X and Y is defined as:

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i, Y_j)\log_2\left(P(X_i, Y_j)\right) \quad (14)$$

When certain variables are known and others are not, the remaining uncertainty is measured by the conditional entropy:

$$H\left(\frac{Y}{X}\right) = \sum_{i=1}^{n} P(X_i) H\left(\frac{Y}{X} = X_i\right)$$
$$= -\sum_{i=1}^{n} P(X_i) \sum_{j=1}^{m} P\left(\frac{Y_j}{X_i}\right) \log_2\left(P\left(\frac{Y_j}{X_i}\right)\right) \quad (15)$$
$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i, Y_j)\log_2\left(P(Y_j/X_i)\right)$$

The joint entropy and the conditional entropy have the following relation:

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y) \quad (16)$$

This is known as 'chain-rule', implies that the total entropy of random variables X and Y is the entropy of X plus the remaining entropy ofY for a given X as shown in Fig. 3.

The information found commonly in two random variables is of importance in our work and this is defined as the mutual information between two variables.

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \quad (17)$$

If the mutual information value between two random variables is large, it means that these are two correlated variables. If the mutual information becomes zero, the two random variables are totally unrelated or the two variables are independent.

The mutual information and the entropy have the following relation, as shown in Fig. 3:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \quad (18)$$

We use the ability of mutual information technique in extracting important set of input variables from all input candidates in order to forecast the price. For price forecasting of an hour, we use the input set of lagged values of price up to 200 hours ago. In this the paper previous 49 days is considered as the training periods. The mutual technique uses logarithm in base 2. So all inputs and output data must maps to binomial distributions. For this purpose, all inputs and output data are linearly normalized in the range of [0,1]. If each normalized variable is bigger than its median, it will map to 1 and otherwise, it will map to 0. Now, 200+1 binomial distributions for inputs and output of mutual technique are generated.

According to Eq. 18, the mutual information $I(X, Y)$ between input X and output Y is calculated. For this purpose, we define the joint probability $P(X, Y)$ similar to [2] as follows:

$$U = Y + 2X \quad (19)$$

where U can have four digits {0,1,2,3}. The number of U (out of all N training samples) that are 0 is defined as U_0 and similarly for U_1, U_2 and U_3.

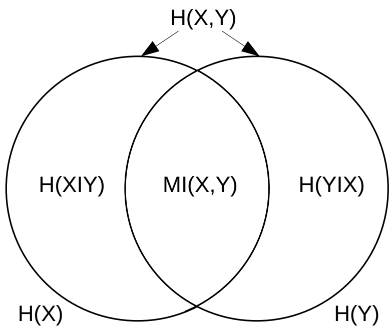

Fig. 3. Representation Mutual Information and different Entropies [28]

Table 1. Four states of the joint probability $P(X, Y)$ and two state probabilities $P(X)$ and $P(Y)$

Y / X	Y=0	Y=1
X=0	$P(X=0, Y=0) = U_0/N$	$P(X=0, Y=1) = U_1/N$
X=1	$P(X=1, Y=0) = U_2/N$	$P(X=1, Y=1) = U_3/N$

According to above definitions, four states of the joint probability $P(X, Y)$ and two state probabilities $P(X)$ and $P(Y)$ are computed as shown in Table 1. Based on the individual and joint probabilities and replacing in Eq. (18), the mutual information between the candidate input X and target variable Y is computed. This procedure is repeated for all inputs candidates and output and the mutual information between each input candidates and output is calculated and then are ranked. Input candidates with high mutual information are selected as important input variables for price forecasting by ANFIS algorithm and the others are discarded.

As a case study on Spanish Market, the prices of sample testing days February 18, May 20, August 19 and November 18 at different seasons of 2002 and their previous 49 days are considered.In Spanish electricity market, there are 24 hourly observations each day. The data are divided into two subsets: the training sample and the testing sample. The training sample is used tobuild the forecasting model, while the testing data is applied for the robustness check of the model.

By applying the mutual information technique, as described above, important lagged prices as input variables based on the value of their Normalized Mutual Information (NMI) that have the price of $h(P_h)$ are ranked and shown in Table 2.

4. Forecasting Accuracy Assessment

Several criteria defined in [20] have been used to measure the forecast and the accuracy of the proposed MIAP approach. Mean Absolute Percentage Error (MAPE) takes the absolute error of each forecast ($|P_h^{act.} - P_h^{fore.}|$), and divides it by the value ofthe actual price atthat hour ($P_h^{act.}$), and then averages these percentage errors in that forecasted hours (N). In order

Table 2. Ten high ranks input variables for testing days of Spanish Market (2002)

Rank	Feb. 18		May. 20		Aug. 19		Nov. 18	
	Selected Feature	N.MI Value	Selected Feature	N.MI Value	Selected Feature	N.MI Value	Selected Feature	N.MI Value
1	P_{h-1}	1.000	P_{h-1}	1.000	P_{h-168}	1.00	P_{h-1}	1.000
2	P_{h-168}	0.953	P_{h-2}	0.801	P_{h-1}	0.94	P_{h-168}	0.897
3	P_{h-24}	0.804	P_{h-168}	0.722	P_{h-169}	0.88	P_{h-169}	0.778
4	P_{h-169}	0.803	P_{h-169}	0.653	P_{h-167}	0.88	P_{h-167}	0.769
5	P_{h-167}	0.801	P_{h-167}	0.652	P_{h-166}	0.74	P_{h-2}	0.734
6	P_{h-23}	0.711	P_{h-3}	0.651	P_{h-170}	0.73	P_{h-144}	0.661
7	P_{h-144}	0.698	P_{h-166}	0.597	P_{h-2}	0.71	P_{h-192}	0.658
8	P_{h-2}	0.683	P_{h-144}	0.585	P_{h-192}	0.69	P_{h-170}	0.650
9	P_{h-25}	0.676	P_{h-170}	0.569	P_{h-144}	0.66	P_{h-166}	0.626
10	P_{h-72}	0.665	P_{h-143}	0.552	P_{h-193}	0.64	P_{h-24}	0.603

to avoid the adverse effect of hourly prices close to zero, the MAPE is divided by the mean actual price of that period -day or week- $\overline{P_h^{act.}}$ instead of the price in that hour. This work forecasts the price of a week ahead of electricity market and then uses weekly MAPE criterion (N=168) which can be computed as follows:

$$MAPE\% = \frac{1}{168}\sum_{h=1}^{N=168}\frac{|P_h^{act.} - P_h^{fore.}|}{\overline{P_h^{act.}}} \times 100 \qquad (20)$$

Error variance criterion is the other index used to measure the uncertainty of the proposed model. Small error variance gives more precise price prediction [20]. Weekly error variance can be defined as:

$$\sigma_{e,week}^2 = \frac{1}{168}\sum_{h=1}^{N=168}\left(\frac{|P_h^{act.} - P_h^{fore.}|}{\overline{P_h^{act.}}} - e_{week}\right)^2 \times 100 \qquad (21)$$

Where:

$$e_{week} = \frac{1}{168}\sum_{h=1}^{N=168}\frac{|P_h^{act.} - P_h^{fore.}|}{\overline{P_h^{act.}}} \qquad (22)$$

At the following, the weekly MAPE and weekly error variance are used to compare the accuracy of proposed MIAP approach with previously published approaches.

The proposed hybrid MIAP approach is applied to forecast next-week in Spanish electricity market for the year of 2002 which is publicly available on its website [29]. It should be noted that the mainland Spain electricity market, a duopoly, is having a dominant player, which changes the price strategic bidding and thus it is hard to predict the electricity prices accurately for the next day [20]. Besides, we compare the approach with at least ten previously published price forecasting with the other methods previously reported in [5], [8], [9], [2], [13-15, 20, 21],[27] and [30], the same test weeks are considered and are shown in Table 3 and no exogenous input variables are selected.

The forecasted price obtained with MIAP during the spring test week is shown in Fig. 4 along with actual price. It can be seen from Fig. 4 that forecasted prices of the week accurately match with the actual prices even the mild spikes and valleys on Sunday that is a unique characteristic for hybrid MIAP approach. For the winter, summer, and fall test weeks, forecasted prices are shown in Figs. 5 to 7, respectively.

Table 3. Hourly price data for forecasting methods

Markets	Seasons	Historical hourly price data	Test weeks
Spanish (2002)	Winter	Jan. 1-Feb. 17	Feb. 18- Feb. 24
	Spring	April 2- May 19	May 20-May 26
	Summer	July 2- Aug. 18	Aug. 19- Au g. 25
	Fall	Oct. 1- Nov.17	Nov. 18- Nov.24

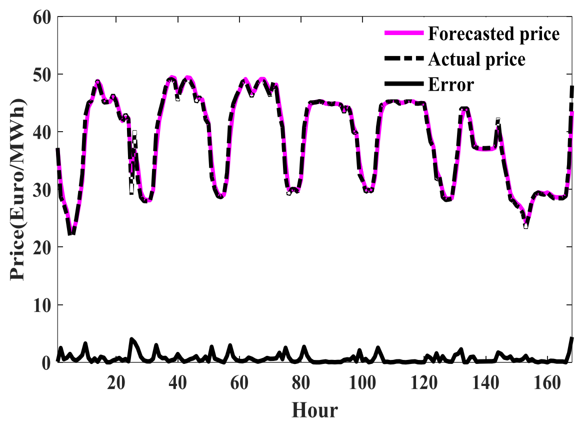

Fig. 4. Forecast Results of the Spring Test Week (2002)

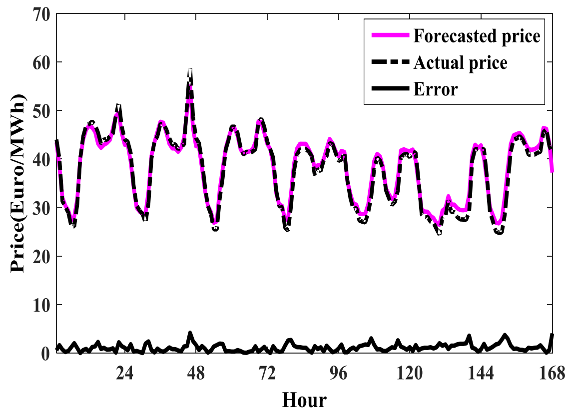

Fig. 5. Forecast Results of the Winter Test Week (2002)

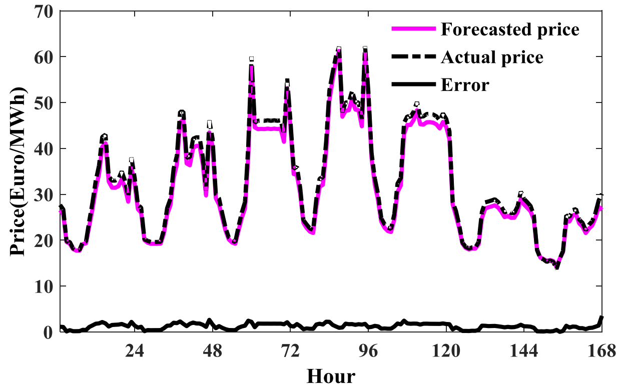

Fig. 6. Forecast Results of the Summer Test Week (2002)

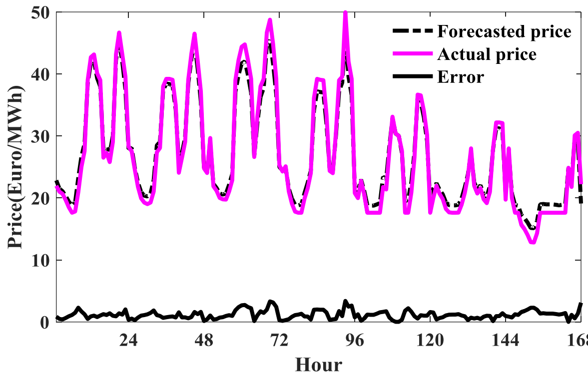

Fig. 7. Forecast Results of the Fall Test Week (2002)

Table 4 compares hybrid MIPA approach and twelve other approaches regarding the MAPE criterion. In the second and third rows, results of ARIMA model and ARIMA plus wavelet transform as an external decomposer/composer have been quoted from [5, 20].The fourth row shows the results of the new fuzzy neural network quoted from [13] in which instead of decomposing price signal, it divides smoothly input space in the form of fuzzified classification. The fifth row of Table 4 represents the obtained results from an MLP neural network with LM learning algorithm, quoted from [9].In the sixth row, results of mixedmodel in which separate ARIMA model is implemented for each hour of the day are represented (quoted from [30]).The idea of separate modeling of the 24 hourly time series of electricity price is also used in [2] with neural network as forecast block. This reference uses mutual information algorithm as the feature selection technique. Results of Cascaded Neuro-Evolutionary Algorithm (CNEA) method are shown in the seventh row. The eighth row of Table 4, represents the obtained results of Adaptive Wavelet Neural Network (AWNN) quoted from [8] in which wavelet neural network is used instead of classical feed-forward neural network. Results of Hybrid Intelligent System (HIS) method [15] are shown in the ninth row. Weighted nearest neighbors (WNNs) method has been presented for electricity price forecast technique in [14]. Results of implementing this method on forecasting test weeks are shown in the tenth row. In the eleventh row, results of Wavelet-PSO-ANFIS approach are quoted from [21]. Wavelet transform is as an external decomposer/ compose. The hybrid algorithm [23] include various methodologies like Wavelet Transform (WT), Fuzzy Adaptive Particle Swarm Optimization (FA-PSO) and Feed Forward Neural Networks (FFNN)).It should be noted that although the idea of using hybrid PSO-ANFIS (HPA) algorithmin electricity price forecasting has been presented in [21, 27], our paper uses a different strategy, explained in Section 3.1, in using PSO algorithm for the fine-tuning of adjustable ANFIS parameters so that this strategy improves forecasting accuracy. Besides, the other part of our proposed technique includes mutual information based technique in selecting the best input features for price forecasting and disregarding the others. The results of our proposed MIAP algorithm are shown in the last row. The MAPE for the Spanish market has an average value of 57.3%. Improvement in the average MAPE of MIPA with respect to the previous eleven approaches (ARIMA, Mixed-model, NN, W-ARIMA, WNN, FNN, HIS, AWNN, CNEA, WPA, Wavelet-FA-PSO-FFNN, HPA) is 68.8%, 66.6%, 65.2%, 61.7%, 61.4%, 58.7%, 55.5%, 54.08%, 41.7%, 38.8%, 37.5%, 41%, respectively. It indicates that the average of the weekly MAPE values of the proposed approach is considerably lower than other techniques. It means that the proposed hybrid MIAP outperforms in the forecasting the electricity prices compared to other approaches.

Table 4. Comparative MAPE results for four weeks of Table 3

Technique	Winter	Spring	Summer	Fall	Average
ARIMA [5]	6.32	6.36	13.39	13.78	9.96
Wavelet-ARIMA [20]	4.78	5.69	10.70	11.27	8.11
FNN [13]	4.62	5.30	9.84	10.32	7.52
NN [9]	5.23	5.36	11.40	13.65	8.91
Mixed-model[30]	6.15	4.46	14.9	11.68	9.30
CNEA[2]	4.88	4.65	5.79	5.96	5.32

AWNN [8]	3.43	4.67	9.64	9.29	6.75
HIS[15]	6.06	7.07	7.47	7.30	6.97
WNN[14]	5.15	4.34	10.89	11.83	8.05
WPA[21]	3.37	3.91	6.50	6.51	5.07
Wavelet-FA-PSO-FFNN [23]	4.04	3.43	5.32	6.51	4.8
HPA[27]	3.65	4.19	6.76	6.53	5.28
Proposed MIAP	2.95	1.49	3.7	4.27	3.10

In Table 5, a comparison between the variance of the prediction error of the other nine approaches (ARIMA, NN, WARIMA, FNN, HIS, AWNN, CNEA, WPA) and the proposed method is shown. Smaller weekly error variance by the proposed MIAP indicates its less uncertainty in the forecasting.

Table 5. Comparative weekly forecasting error variance for four weeks of Table 3

Technique	Winter	Spring	Summer	Fall	Average
ARIMA [5]	0.0034	0.0020	0.0158	0.0157	0.0092
NN [9]	0.0017	0.0018	0.0109	0.0136	0.0070
Wavelet-ARIMA [20]	0.0019	0.0025	0.0108	0.0103	0.0064
FNN [13]	0.0018	0.0019	0.0092	0.0088	0.0054
HIS [15]	0.0012	0.0031	0.0074	0.0075	0.0048
AWNN [8]	0.0009	0.0017	0.0074	0.0049	0.0037
CNEA [2]	0.0036	0.0027	0.0043	0.0039	0.0036
WPA [21]	0.0008	0.0013	0.0056	0.0033	0.0027
Proposed MIAP	0.0006	0.0001	0.0002	0.0009	0.0004

Time series models such as ARIMA are linear predicator. Due to nonlinear behavior of electricity price signal as a function of its input, these forecasting techniques cannot predict rapid changes of the electricity price. As is shown in Table 4, the accuracy of ARIMA based techniques [5, 20] in predicting the same test weeks is less than artificial intelligence based techniques which can model the complex nonlinear relation between input variables and target price.

Some previous techniques uses wavelet transform as a decomposer/composer. The accuracy of our proposed approach is better and it seems that due to high frequency components of price signals, Wavelet transform may lose some information. This fact can impact on the accuracy of wavelet based approaches.

As obtained results show, ANFIS based techniques [21], and proposed MIAP approaches which uses a fuzzy inference system for training the ANN has better performance relative to NN in electricity price forecasting. Generally, the efficiency of artificial intelligence based techniques has close relation with proper selection of input variables and fine-tuning of adjustable parameters of their architecture. Input candidate can be lagged values of the price signal and exogenous variables such as load and generation data provided that they are available in the corresponding electricity market. In order to avoid the curse of dimensionality and to limit the number of adjustable parameters which is increased with number of input variables, important input variables that contribute forecasting accuracy should be selected among this potentially huge number of input variables. For this purpose, MI technique is used in the proposed MIAP approach. The proposed strategy in correct and fast tuning ANFIS parameters by PSO algorithm, explained in Section 3.1, plays important role in the accuracy of proposed price forecasting approach.

Our proposed MIAP approach was executed for different weeks in different seasons of the year 2002 for the Spanish electricity market. The obtained weekly MAPEs and error variances for these weeks are close to the results of Tables 4 and 5. In order to evaluate the daily price forecasting of the proposed MIAP approach, it is examined in price forecasting of the days of the other testing week of the Spanish electricity Market (August 25-31, 2000).

The accuracy of the proposed method is compared with the CENA, ARIMA and mixed-model and is shown in Table 6. Simulation results show that the daily mean errors at all days of the testing week are considerably lower than CENA, ARIMA and mixed-model. Improvement in the average mean error of the proposed approach with respect to the three mentioned approaches is 62%, 73%, 68%, respectively.

Fig. 8 shows a graphical view regarding the forecast accuracy ofthe proposed method in the priceforecasting of the testing week. The weekly MAPE for thetest week has an average value of 2.21%.

Table 6. Daily mean error for August 25-31. 2000 of the Spanish Electricity market

Technique	Mixed-Model [30]	ARIMA[5]	CNEA[2]	Proposed MIAP
Day1	4.8%	4.34%	4.65%	2.05%
Day2	7.3%	7.99%	6.62%	2.28%
Day3	5.4%	4.57%	5.22%	2.92%
Day4	4.6%	10.81%	4.59%	1.63%
Da5	5.1%	6.12%	5.56%	1.12%
Day6	14.9%	17.34%	8.52%	3.05%
Day7	7.2%	6.05%	6.11%	2.41%
Average	7.04%	8.17%	5.90%	2.21%

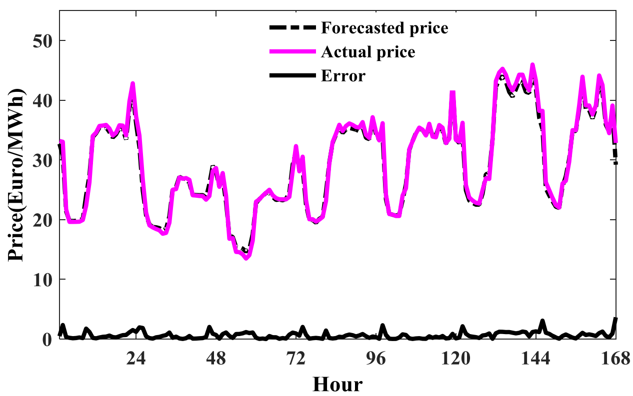

Fig.8. Forecast Results of Summer Test Week (Aug. 2000)

The simulation analyses are implemented in MATLAB. The average computational time is 54 seconds by using a Pentium P4 2.3-GHz personal computer with 6 GB of RAM memory. Considering the forecasting time horizon, this set up time is acceptable within a short-term decision making framework. This characteristic makes MIAP as a practical approach with high accuracy.

5. Conclusion

Accurate electricity price forecasting is of great significance for the entire electrical power system. In doing so, this paper proposes a new hybrid method by combining the ANFIS, PSO, and MI algorithms. The proposed forecasting method is examined by using data from the Spanish market. The results confirm the considerable performance of the proposed MIAP approach in forecasting short-term electricity prices. The great performance of the proposed hybrid algorithm can be attributed to the following three causes. Firstly, the ANFIS algorithm has nonlinear mapping capabilities, which can capture the nonlinear component of electricity pricesmore easily. Secondly, PSO algorithm can well tune ANFIS parameters, in which choosing inappropriate adjusting variables will cause either under or over fitting. Thirdly, MI technique can appropriately select and limit the input variables of ANFIS structure. Extension of this method that includes more concentrating on predicting the rapid changes of price signal will be considered in the future work.

6. References

[1] D. Wang, H. Luo, O. Grunder, Y. Lin, and H. Guo, "Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm," *Applied Energy,* vol. 190, pp. 390-407, 2017/03/15/ 2017.

[2] N. Amjady and F. Keynia, "Day-Ahead Price Forecasting of Electricity Markets by Mutual Information Technique and Cascaded Neuro-Evolutionary Algorithm," *IEEE Transactions on Power Systems,* vol. 24, no. 1, pp. 306-318, 2009.

[3] R. Weron and A. Misiorek, "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models," *International Journal of Forecasting,* vol. 24, no. 4, pp. 744-763, 2008/10/01/ 2008.

[4] A. J. Conejo, J. Contreras, R. Espínola, and M. A. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," *International Journal of Forecasting,* vol. 21, no. 3, pp. 435-462, 2005/07/01/ 2005.

[5] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Transactions on Power Systems,* vol. 18, no. 3, pp. 1014-1020, 2003.

[6] J. P. González, A. M. S. Roque, and E. A. Pérez, "Forecasting Functional Time Series with a New Hilbertian ARMAX Model: Application to Electricity Price Forecasting," *IEEE Transactions on Power Systems,* vol. 33, no. 1, pp. 545-556, 2018.

[7] G. P. Girish, "Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models," *Energy Strategy Reviews,* vol. 11-12, pp. 52-57, 2016/06/01/ 2016.

[8] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets," *IEEE Transactions on Power Systems,* vol. 23, no. 3, pp. 1423-1432, 2008.

[9] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Electric Power Systems Research,* vol. 77, no. 10, pp. 1297-1304, 2007/08/01/ 2007.

[10] K. Meng, Z. Y. Dong, and K. P. Wong, "Self-adaptive radial basis function neural network for short-term electricity price forecasting," *IET Generation, Transmission & Distribution,* vol. 3, no. 4, pp. 325-335, 2009.

[11] P. Areekul, T. Senjyu, H. Toyama, and A. Yona, "A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market," *IEEE Transactions on Power Systems,* vol. 25, no. 1, pp. 524-530, 02// 2010.

[12] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping," *IEEE Transactions on Power Systems,* vol. 27, no. 4, pp. 2055-2062, 2012.

[13] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Transactions on Power Systems,* vol. 21, no. 2, pp. 887-896, 2006.

[14] A. T. Lora, J. M. R. Santos, A. G. Exposito, J. L. M. Ramos, and J. C. R. Santos, "Electricity Market Price Forecasting Based on Weighted Nearest Neighbors Techniques," *IEEE Transactions on Power Systems,* vol. 22, no. 3, pp. 1294-1301, 2007.

[15] N. Amjady and M. Hemmati, "Day-ahead price forecasting of electricity markets by a hybrid intelligent system," *European Transactions on Electrical Power,* vol. 19, no. 1, pp. 89-102, 2009.

[16] P. Du, J. Wang, W. Yang, and T. Niu, "Multi-step ahead forecasting in electrical power system using a hybrid forecasting system," *Renewable Energy,* vol. 122, pp. 533-550, 2018/07/01/ 2018.

[17] L. Xiao, W. Shao, M. Yu, J. Ma, and C. Jin, "Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system

forecasting," *Applied Energy,* vol. 198, pp. 203-222, 2017/07/15/ 2017.

[18] H. Shayeghi, A. Ghasemi, M. Moradzadeh, and M. Nooshyar, "Day-ahead electricity price forecasting using WPT, GMI and modified LSSVM-based S-OLABC algorithm," *Soft Computing,* vol. 21, no. 2, pp. 525-541, 2017/01/01 2017.

[19] S. Lahmiri, "Comparing Variational and Empirical Mode Decomposition in Forecasting Day-Ahead Energy Prices," *IEEE Systems Journal,* vol. 11, no. 3, pp. 1907-1910, 2017.

[20] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems,* vol. 20, no. 2, pp. 1035-1042, 2005.

[21] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1-1.

[22] Y. H. S. L. Z. Z. X. Y. C. S. J. Yao, "Prediction of System Marginal Prices by Wavelet Transform and Neural Networks," *Electric Machines & Power Systems,* vol. 28, no. 10, pp. 983-993, 2000/10/01 2000. M. Khooban, "Probabilistic Forecasting of Hourly Electricity Price by Generalization of ELM for Usage in Improved Wavelet Neural Network," *IEEE Transactions on Industrial Informatics,* vol. 13, no. 1, pp. 71-79, 2017.

[25] J. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 23, no. 3, pp. 665-685, 1993.

[26] M. A. Shoorehdeli, M. Teshnehlab, A. K. Sedigh, and M. A. Khanesar, "Identification using ANFIS with intelligent hybrid stable learning algorithm approaches and stability analysis of training methods," *Applied Soft Computing,* vol. 9, no. 2, pp. 833-850, 2009/03/01/ 2009.

[27] H. M. I. Pousinho, V. M. F. Mendes, and J. P. S. Catalão, "Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach," *International Journal of Electrical Power & Energy Systems,* vol. 39, no. 1, pp. 29-35, 2012/07/01/ 2012.

[28] N. Kwak and C. Chong-Ho, "Input feature selection by mutual information based on Parzen window," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, no. 12, pp. 1667-1671, 2002.

[29] *Market Operator of the Electricity Market of Mainland Spain, MEL, 2010. [Online]. Available: http://www.omel.es/.*

[30] C. Garcia-Martos, J. Rodriguez, and M. J. Sanchez, "Mixed Models for Short-Run Forecasting of Electricity Prices: Application for the Spanish Market," *IEEE Transactions on Power Systems,* vol. 22, no. 2, pp. 544-552, 2007.