Hybrid Model using Three-Stage Algorithm for Simultaneous Load and Price Forecasting

Mehrdad Setayesh Nazar[1], Ashkan Eslami Fard[1], Alireza Heidari[2], Miadreza Shafie-khah[3], and João P. S. Catalão[3,4,5*]

[1] Faculty of Electrical Engineering, Shahid Beheshti University, A.C., Tehran, Iran
[2] The University of New South Wales, Sydney, Australia
[3] C-MAST, University of Beira Interior, 6201-001 Covilhã, Portugal
[4] INESC TEC and Faculty of Engineering of the University of Porto, 4200-465 Porto, Portugal
[5] INESC-ID, Instituto Superior Técnico, University of Lisbon, 1049-001 Lisbon, Portugal

Abstract

Short-term load and price forecasting is an important issue in the optimal operation of restructured electric utilities. This paper presents a new intelligent hybrid three-stage model for simultaneous load and price forecasting. The proposed algorithm uses wavelet and Kalman machines for the first stage load and price forecasting. Each of the load and price data is decomposed into different frequency components, and Kalman machine is used to forecast each frequency components of load and price data. Then a Kohonen Self Organizing Map (SOM) finds similar days of load frequency components and feeds them into the second stage forecasting machine. In addition, mutual information based feature selection is used to find the relevant price data and rank them based on their relevance. The second stage uses Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) for forecasting of load and price frequency components, respectively. The third stage machine uses the second stage outputs and feeds them into its MLP-ANN and ANFIS machines to improve the load and price forecasting accuracy. The proposed three-stage algorithm is applied to Nordpool and mainland Spain power markets. The obtained results are compared with the recent load and price forecast algorithms, and showed that the three-stage algorithm presents a better performance for day-ahead electricity market load and price forecasting.

Keywords: Feature selection; Mutual information; Neural network; Price and load forecasting; wavelet transformation**.**

NOMENCLATURE

Index sets	
t	Discrete time index
l	Length of signal index of wavelet decomposition
j	Decomposed level index of wavelet decomposition
k	Scaling index of wavelet decomposition
Parameters	
η	Correction rate for primary load and price forecast
Δ	Predefined parameter for primary load and price forecast
n	Number of step in primary load and price forecast
a	Spread control of mother wavelet filter
b	Translation parameter of mother wavelet filter

* Corresponding author at Faculty of Engineering of the University of Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal.
E-mail address: catalao@ubi.pt (J.P.S. Catalão).

m	Integer value for mother wavelet filter
n	Integer value for mother wavelet filter
φ	Scaling function of coarse scale coefficients for mother wavelet filter
c	Scaling function of fine scale coefficient for mother wavelet filter
ω	Scaling functions of fine scale coefficient for mother wavelet filter
Ψ	Mother wavelet function
ξ	Forecast horizon for mother wavelet filter
Variables	
$L\ (t)$	Primary load forecast variable
$x(\varpi+1)$	Model state matrix
$A(\varpi)$	State transition matrix
$y(\varpi)$	Measured signal
$C(\varpi)$	Output matrix
$\chi(\varpi)$	System error
$\iota(\varpi)$	Measured error
Q_1	Noise covariance matrix
Q_2	Error covariance matrix
$K(\varpi)$	Kalman gain
$P(\varpi+1)$	Error covariance matrix
$y(\varpi)$	Kalman load forecast
r	Correlation coefficient between two random variables
cov	Covariance
E	Expected value
σ	Standard deviations
Θ	Probability mass function
Ξ	Joint entropy
M , N	Fuzzy membership functions of ANFIS
Υ	Firing strength of the ANFIS rule
$\hat{V}_h$	Forecasted value of prices or load
V	Actual value of prices or load
$\overline{V}$	Average forecasted value

1. Introduction

Load and market price forecasting are important tasks for system operators in restructured power systems [1], [2]. An Independent System Operator (ISO) is responsible for its system security and cost reduction; these tasks are highly dependent on hourly load and market price forecasting according to the fact that the hourly market price is dependent on hourly load [3]. Over the years, extensive works have been performed on the load and price forecasting methods that can be classified into three main categories [4], [5]. The first category is classical statistical methods that use linear analysis [6]. The second category deals with intelligent forecasting algorithms that are used for non-linear forecasting problem. The third category encounters new heuristic ideas in the forecasting paradigms that consist of combined and hybrid models [6]. Other methods can be recognized as a combination of the above categories [7].

The classical statistical methods like Kalman filtering [8], Autoregressive Integrated Moving Average (ARIMA) [9], exponential smoothing [10], state space model [11], and Box-Jenkins models [12] are based on statistical models. The Intelligent techniques include Support Vector Machine (SVM) [13], Support Vector Regression (SVR) [14], fuzzy inference model [15], Knowledge-Based Expert System (KBES) [16], and Artificial Neural Network (ANN) [17], [18].

Hybrid load and price forecasting techniques are the most common methods that show more accurate and acceptable results as compared to custom separate load and price methods [19-21]; thus, in this paper, a hybrid method is proposed. The wavelet decomposition techniques have been used in some hybrid models [22-23] to decompose high and low frequency components of load and price to a set of sub-series. It facilitates the analysis of complex feature of load and price profile, and each part of the sub-series can be predicted easier than that of the original signal. This method is considered in this paper. In [22], Wavelet Transform (WT) and Adaptive-Network-based Fuzzy Inference System (ANFIS) are used. WT decomposes price series into a set of constitutive series, and these series are forecasted using ANFIS. In [23], the wavelet pre-processed time series are used after removing the higher frequency (fast changing) components.

Any market-based load forecasting method cannot work well without considering price as an input. One of the hybrid methods for solving this problem is an iterative model that considers the full dependency of price and load [24-26]; this model is also considered in this paper. A mixed load and price forecasting method is proposed in [24] that consists of a two-level forecast algorithm. The first level uses forecasters for the price and load forecasting. The second level uses two final forecasters that they are equipped with Feature Selection (FS) algorithm. These hybrid methods assume that the Market Clearing Price (MCP) curve has a non-constant variance and average without any pattern [25-26]. Ref. [25] proposes a method that uses the cooperative co-evolutionary approach with adjustable connections in a recursive procedure. In addition, similar days-based methods have been used to investigate the days with similar characteristics including similar week/day indexes or weather parameters during the last two or three years [26]. The major drawbacks of these methods are in the way of finding the similar days and creating a linear function of the past load patterns. For solving this problem, a combination of similar days-based methods and machine learning algorithms is proposed in [27], in which similar days are selected by the felt temperature, and after wavelet decomposition, each frequency pattern is fed into an ANN as a machine-learning algorithm. This method is also considered in this paper.

The ANN-based load-forecasting methods are among the most popular forecasting algorithms, and many researchers have used unsupervised learning ANN (for example, Self-Organizing Map (SOM)) for better performance of Multi-Layer Perceptron (MLP) forecasting algorithms [28-31]. Selecting the best fitting data (as inputs) might be an important issue in load-forecasting methods. The most common inputs for ANN-based methods used in the previous works include weather data [28], historical loads [29], historical prices [30] and week/day index [31]. The historical MCP curve has a different characteristic from load curve, and no similar curve exists among the historical price data; this fact increases the complexity of the price-forecasting problem [31]. Mutual Information (MI) method is one of the FS techniques that can find the most relevant data and rank them according to their relevance to the target, which decreases the redundancy of

data set and is not time-consuming [19]. Elimination of unimportant and redundant data and reducing computational complexity are the main advantages of this method [32-33]; this method is considered in this paper.

Several optimization algorithms have been proposed to optimize the parameters of the hybrid forecasting methods [4, 7, 34-37]. For example, [34] uses fuzzy clustering to find the similar days; the proposed method combines the classical methods into one hybrid method that makes the forecasts based on a combination of recent historical data and similar day data. It consists of three units: a pre-processing unit, which is responsible for detecting that a season has changed and searching for similar days; the second unit is an SVM-based hourly predictor, and the third unit is for optimizing the SVM parameters based on the Particle Swarm Optimization (PSO). In [35], a hybrid model based on a modified firefly algorithm SVR reduces the possibility of trapping in local optima when increasing the convergence criterion. In [36], an algorithm that uses PSO-SVM is proposed, and the obtained results are compared to the classical training methods results. In [37], a hybrid model developed to forecast air conditioning electrical load, and comparisons are made among the applied methods to prove the advantages and applicability of the proposed method.

It was observed that using optimization-based techniques for FS of Simultaneous short–term Price and Load Forecasting (SPLF) might not lead to an acceptable trade-off between accuracy and computational burden. It only increased the complexity of the proposed model and computational efforts without considerable improvement of the algorithm accuracy. Thus, optimization-based FS methods are not used in this paper; rather MI-based FS method has been used.

The authors had many attempts to define a proper general layout for soft computing algorithms and to solve the high error problem of the simultaneous price and load forecasting methods. The results of different competitive soft computing paradigms were compared, and finally, the proposed layout and its soft computing algorithms were selected.

The overall structure of the proposed method is the contribution of this paper. The trade-off between accuracy and computational complexity is one of the main targets of this research.

This paper`s contribution includes:

- introducing a model to find the best datasets of historical loads for ANN-based load forecasting model training;
- using MI method as feature selection to generate the training data sets for ANFIS price forecasting model; and
- improving load-forecasting accuracy by price forecasting amendment.

The rest of the paper is organized as follows: Section 2 describes the problem description and modelling. Section 3 proposes a performance evaluation of the proposed algorithm. Numerical results and analysis are presented in Section 4. Finally, the paper is concluded in Section 5.

2. Problem description and modelling

A pool market typically includes a day-ahead market, several adjustment markets, and balancing markets. In addition, the regulation reserve market is required to ensure the secure system operation.

An ISO must forecast day-ahead load and MCP. The major challenge is that after clearing the day-ahead market, MCP is determined for each hour. It is understood that MCP curve is fluctuating such that it has a characteristic with non-constant variance and average and has no pattern; on the other hand, the daily loads have similar pattern [24, 25]. However, the hourly MCP is dependent on hourly load, and vice versa. Thus, load-forecasting methods that are developed for price taker loads are not adequate tools [30], and SPLF algorithms must be designed.

Based on the above discussion, this paper proposes an iterative method for the price and load forecasting. The block diagram of the proposed method is shown in Fig. 1.

At first, a primary (not necessarily accurate) forecast is performed by an iterative model to generate a next 24-hour load and price profile. This stage uses WT and Kalman forecasting model. Each of the load and price data is decomposed into three frequency components, and Kalman machine is used to forecast each frequency components of load and price data based on the fact that the forecasting of frequency components of load and price is easier than that of the original data. Then a Kohonen SOM finds similar days of load frequency components and feeds them into the second stage load and price forecasting. In addition, MI-based FS is used to find the relevant price data and rank them based on their relevance. The second stage uses MLP-ANN and ANFIS for forecasting of load and price frequency components, respectively. This stage uses Kalman load and price forecasting of frequency components and the Kohonen similar days pattern recognition outputs. The third stage uses the second stage outputs and feeds them into its MLP-ANN and ANFIS machines to improve the load and price forecasting accuracy. The block diagram of the proposed method is shown in Fig. 1.

2.1. Primary forecast model

The primary forecast estimates the trend of next-day load profile. It is assumed that price and load are dependent on each other as they create each other in a recursive mode. Thus, an hourly estimation of price is essential, and a recursive algorithm of price and load forecasting is proposed. Fig. 2 shows the flowchart of this algorithm.

The process is stopped as the difference between two sequential forecasted prices reaches a predefined parameter Δ.

A correction of load L (t) in the n+1 step is calculated as:

$$L_{n+1}(t) = L_n(t) + \eta\Delta. \tag{1}$$

Where, the η parameter is the correction rate and is close to the learning rate of the NN training algorithm. Δ is a predefined parameter.

First, wavelet decomposition is performed on the price and load historical data. Then using two levels of decomposition, three frequency components are obtained. Finally, Kalman forecasting algorithm is used to forecast the load and price frequency components.

2.1.1. Wavelet decomposition

As mentioned earlier, wavelet decomposition is performed on the price and load historical data, and by two levels of decomposition, three frequency components are obtained. WT is used to divide load and price signals to a set of sub-series [38, 39]. This method analyses each part of the sub-series separately because the prediction procedure of sub-series is easier than that of the original signal. Since the hourly load and price series are calculated, Discrete Wavelet Transform (DWT) must be used as:

$$W(m,n) = 2^{-(m/2)} \sum_{t=0}^{l-1} f(t)\varphi(\frac{t-b}{a}), \tag{2}$$

Where, l and t are the length and discrete time indexes of $f(t)$, respectively, and a and b are the spread control and translation parameters of mother wavelet filter and are defined as $a=2^m$ and $b=n2^m$, respectively (m and n are integer values) [3].

The decomposition can be done by iterative upsampling and convolve by the original signal to the desired level. Finally, the main signal $f(t)$ can be reconstructed from the calculated coefficients by the following equation:

$$f(t) = \sum_{k} c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j<j_0} \sum_{k} \omega_{j,k} 2^{\frac{j}{2}} \psi(2^j t - k), \tag{3}$$

Where, j is the decomposed level, k is the scaling index, $\varphi_{j0,k}$ is the scaling function of coarse scale coefficients, and $c_{j0,k,}$ and $\omega_{j,k}$ are the scaling functions of fine scale coefficients. During each level of decomposition, the length of original signal becomes half that is called *down sampling*. To recover the original signal length after reconstruction, a vector of zeros is placed between every two samples, called *upsampling* [23].

There are different kinds of wavelet functions; among them, Daubechies families (dbN) are more common [38]. Any selection among the wavelet functions should make a trade-off between smoothness and wavelength [3]. In the present study, db2 is used as the mother wavelet function, and the price and load signals are decomposed into two levels as depicted in Fig. 3. After finalizing the forecast stages, the decomposed forecasted signal components are combined to reconstruct the unified forecasted load and price by *up sampling* (Fig. 4).

In this paper, wavelet decomposition is performed for load and priced data, and two high frequencies D1 and D2 and a low-frequency A2 are obtained for load and price, separately. High-frequency components may have noise components with values lower than A2. However, the noise values cannot be estimated, and the main signal is reconstructed after deleting the higher frequency components; only low-frequency components are used for reconstruction of the decomposed signals.

Two Kalman forecasting machines are defined in the next stage. The first is called *Kalman`s price forecast* that is used for forecasting the three components of price frequencies, and the second one is *Kalman`s load forecast* that is used for forecasting the three components of load frequencies.

2.1.2. Kalman forecasting model

Based on wavelet decomposition algorithm for D1, D2, and A2 frequencies, three Kalman forecasting processes are used for each of the load and price forecasting procedures, separately.

The Kalman filter is used for the following reasons:

1) The ability of Kalman filter for tracking of signals with time-varying frequencies [40].
2) It does not need learning mechanism.

The basic discrete state-space model of Kalman filter is introduced, which has been implemented in [40] based on weather parameters and historical load as the state vector. The main discrete state equations are as below:

$$x(\varpi+1) = A(\varpi)x(\varpi) + \chi(\varpi), \tag{4}$$

$$y(\varpi) = C(\varpi)x(\varpi) + \iota(\varpi), \tag{5}$$

Where, $x(\varpi+1)$ is model state matrix, $A(\varpi)$ is state transition matrix, $y(\varpi)$ is measured signal at ϖ, $C(\varpi)$ is output matrix, $\chi(\varpi)$ is system error, and $\iota(\varpi)$ is measured error. Noise and error covariance matrices are defined as:

$$E[\chi(\varpi)\chi^T(\varpi)] = Q_1 \text{ and } E[\iota(\varpi)\iota^T(\varpi)] = Q_2, \tag{6}$$

Where, Q_1 and Q_2 are semi-definite and definite matrices, respectively. In this paper, it is assumed that the state vector measurement has no error; so $\chi(\varpi)$ equals zero. Error covariance matrix is defined as Eq. (7):

$$P(\varpi+1) = [A(\varpi) - K(\varpi)C(\varpi)]P(\varpi)[A(\varpi) - K(\varpi)C(\varpi)]^T + K(\varpi)Q_2K^T(\varpi). \tag{7}$$

In Eq. (6), $K(\varpi)$ is Kalman gain and can correct the state vector value as follows:

$$\hat{x}(\varpi+1) = A(\varpi)\hat{x}(\varpi) + K(\varpi)[z(\varpi) - C(\varpi)\hat{x}(\varpi)], \tag{8}$$

In addition, Kalman gain is as follows:

$$K(\varpi) = [A(\varpi)P(\varpi)C^T(\varpi)][C(\varpi)P(\varpi)C^T(\varpi) + Q_2]^{-1}. \tag{9}$$

Using the weighted least squares, a prior estimation of $\hat{x}_0$ and $\hat{P}_0$ can be performed as:

$$\hat{x}(0) = [H^T Q_2^{-1} H]^{-1} H^T Q_2^{-1} Z_0, \tag{10}$$

$$P_0 = [H^T Q_2^{-1} H]^{-1}, \tag{11}$$

Where z_0 is $s \times 1$ matrix and H is $s \times u$ matrix. These matrices are defined as:

$$z_0 = \begin{bmatrix} z(1) \\ z(2) \\ \vdots \\ z(s) \end{bmatrix} \text{ and } H = \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(s) \end{bmatrix}, \tag{12}$$

Where, s is the number of samples defined by the user.

The price and load in the scale of a country will be forecasted. Thus, it is not reasonable to use weather indexes because the weather parameters are not the same in all regions of a country [24]. However, the average temperature is used in the second stage forecast. Hence, the load can be written as a linear model of past loads and prices for a particular time ϖ as:

$$y(\varpi) = load(\varpi) = \sum_{i=1}^{t} \gamma_i . load(\varpi - \varpi_i) + \sum_{i=1}^{j} \varsigma_i . price(\varpi - \vartheta_i), \quad (13)$$

Where, $load(\varpi - \varpi_i)$ and $price(\varpi - \vartheta_i)$ are past load and price parameters, respectively. In addition, γ and ς are derived from adapted transition matrices, and ϖ_i and ϑ_i are the indexes of previous loads and prices, respectively. In this section, a correlation analysis is used to select the most appropriate indexes of the past loads and prices. Thus, the variables that have a higher correlation with the target are retained. The correlation coefficient between two random variables, ι and χ, with expected values (E), μ_ι and μ_χ, and standard deviations σ_ι and σ_χ is defined as [13]:

$$r_{\iota,\chi} = \frac{\mathrm{cov}(\iota, \chi)}{\sigma_\iota \sigma_\chi} = \frac{E[(\iota - \mu_\iota)(\chi - \mu_\chi)]}{\sigma_\iota \sigma_\chi}, \quad (14)$$

Where, *cov* is the covariance.

The mentioned process for each frequency component is performed to find a normalized load profile. The load forecasted by this stage is very similar to the next 24-hour real load. Thus, finding similar days among the historical days can be easily performed.

There are two kinds of data selection in this stage: 1) Data window selection, which determines the number of inputs for Kalman forecasting machine, and 2) Historical day's selection, which determines the number of tracking processes. The historical day's selection has more effects on the process that should be determined by the user according to the structure of the case study.

Then the primary forecast results are fed into the secondary forecast model, and a more accurate price prediction is used for the next hour load forecast.

2.2. SOM-based similar days finding

A Kohonen network reduces the high dimensionality of a set of data to a two-dimensional grid based on the most similar characteristics of a definite pattern. In this paper, an 8×8 hexagonal lattice network is used [41, 42]. The detailed formulation of SOM network will be discussed in the *Numerical Results* section.

2.3. Mutual information

MI has been developed as a feature selection method [43, 44]. The MI method is a statistical and probabilistic method for price forecasting. In [32], the authors present a new MI formulation for price forecasting. Some main concepts of MI are noted here. Entropy is defined as:

$$\Xi(X) = -\sum_{i=1}^{ns} \Theta(X_i) \log_2(\Theta(X_i)), \quad (15)$$

Where, Θ(X) is a Probability Mass Function (PMF), and the discrete variable X has "ns" different possible outcomes. The joint entropy of two variables is defined by:

$$\Xi(X,Y) = -\sum_{i=1}^{ns}\sum_{j=1}^{ms}\Theta(X_i,Y_j)\log_2(\Theta(X_i,Y_j)), \tag{16}$$

Entropy can be calculated as follows:

$$\begin{aligned}\Xi(Y \mid X) &= \sum_{i=1}^{ns}\Theta(X_i)\Xi(Y \mid X = X_i)\\ &= -\sum_{i=1}^{ns}\Theta(X_i)\sum_{j=1}^{ms}\Theta(Y_j \mid X_i)\log_2(\Theta(Y_j \mid X_i))\\ &= -\sum_{i=1}^{ns}\sum_{j=1}^{ms}\Theta(X_i,Y_j)\log_2(\Theta(Y_j \mid X_i)),\end{aligned} \tag{17}$$

According to the chain-rule, it is understood that the entropy of random variable *X* contains some entropy by itself and some common information with the variable *Y*. This can be written as:

$$\Xi(X,Y) = \Xi(X) + \Xi(Y \mid X), \tag{18}$$

Eq. (25) yields the value of common information between *Y* and *X*, known as MI. This concept is very similar to the correlation with the favour of considering probability. Finally, *MI(X, Y)* is as follows.

$$MI(X,Y) = \sum_{i=1}^{ns}\sum_{j=1}^{ms}\Theta(X_i,Y_j)\log_2\left(\frac{\Theta(X_i,Y_j)}{\Theta(X_i)\Theta(Y_j)}\right), \tag{19}$$

The higher MI value between two random variables shows the more relation between them, and zero MI coefficients imply an independent relation. After normalization, the median of each variable set is defined, and any value less than the median is rounded to zero and those higher than the median are rounded to one. More details will be given in the *Numerical Results* section.

2.4. Secondary forecast model

More accurate price and load forecast is the main goal of this forecasting stage. It contains two forecasting machines including ANN and ANFIS; the former predicts the load and uses the first stage load and price as inputs, and the latter improves the price forecast.

Selecting the training data for ANN model is very important. Insufficient or irrelevant data can make a non-perfect relationship between the inputs and outputs and may yield a high error. On the other hand, too many data can cause over-learning.

2.4.1. Multi-layer feedforward perceptron (MLP)

One of the most common forms of ANN is MLP. If the number of inputs and nodes, which form the hidden layer, are selected properly, a good relationship can be established between the inputs and outputs.

ANN inputs play an essential role in its performance. Hence, the input data selection is one of the most important parts of any intelligent model. Since this paper considers a country load forecasting, so the temperature does not have any meaning as an input because the weather changes countrywide, and there is not a constant weather index all over the country. By the way, weather conditions, especially temperature, have a significant effect on load. Therefore, daily

temperature average of a country in the forecasting day and a day before as two ANN inputs are used. Last hours' and days' load and price are also used as inputs for the ANN models, which have a good correlation with the load. For each hour and the frequency component of load, an ANN machine is used separately.

There is no direct approach to find the best and the most optimal structure of an ANN [45-47]. One can find it by trial and error in the way of evaluation of results with different numbers of ANN parameters [5, 48].

In the similar day-based forecasting approach, the daily load patterns are used, which have the most similarity with the load of the forecasting day. Accordingly, a three-layer network with a 20-node hidden layer and 25 inputs is considered. Thus, The MLP ANN trained by backpropagation works properly with fewer neurons in the hidden layer than in the input layer. Many references used the same configuration that the inner layer has fewer neurons than the input layer [48-50].

2.4.2. ANFIS

Since the electricity price for the next 24 hours is not definite, it should be forecasted. However, the price forecasting results have an error, which will have drastic effects on the load forecasting results. Thus, an ANFIS model is employed to improve the price forecasting results. ANFIS has the advantages of both ANN and Fuzzy set theory; however, it causes an unnecessarily large computational burden for large scale of inputs and membership functions [51]. On the other hand, there is not a definite pattern for the price to feed into ANFIS. Accordingly, ANFIS model with just four inputs and MI-based feature selection is employed to find the best-fitted inputs for the second stage price forecasting. Outputs of ANFIS are used to update the price database for the next hour ANN load forecasting; this can improve the load forecasting results. Some papers proposed ANFIS-based models for load forecasting [52], and some others used it or fuzzy systems to refine the load forecasting accuracy [53].

The structure of an ANFIS model with two inputs is shown in Fig. 5.

The functionality of ANFIS is as below:

Each layer has some adaptive nodes as O_i^j , which denotes the output of *j*th node of *i*th layer.

In layer 1, the function of each node is [53]:

$$O_i^1 = \mu M_i(x) \ , \qquad i=1,2, \tag{20}$$

$$O_i^1 = \mu N_i(y) \ , \qquad i=1,2, \tag{21}$$

Where, x and y are the input of *i*th node, and M and N are the linguistic expressions called *membership functions*. Different membership functions can be used according to the case. In this paper, the Gaussian function is used as follows:

$$\mu M_i(x) = e^{-\frac{(x-m1)^2}{2k1^2}}, \tag{22}$$

Where, $m1$ and $k1$ are the parameters to be set during the training.

Layer 2 contains multiplier nodes such that:

$$O_i^2 = \Upsilon_i = \mu M_i(x)\mu N_i(y), \quad i=1,2, \tag{23}$$

Where Υ_i is called firing strength of the ith rule.

Layer 3 normalizes the firing strengths as:

$$O_i^3 = \bar{\Upsilon}_i = \frac{\Upsilon_i}{\sum_j \Upsilon_j}, \quad i=1, 2. \tag{24}$$

The contribution of each rule in the total output is calculated in layer 4 as below:

$$O_i^4 = \bar{\Upsilon}_i z_i = \bar{\Upsilon}_i(\alpha_i x + \beta_i y + \rho_i), \quad i=1,2, \tag{25}$$

Where $\bar{\Upsilon}_i$ is defined based on Eq. (25), and $\{\alpha_i, \beta_i, \rho_i\}$ are consequent parameters.

Layer 5 has one node to sum all of the incoming signals:

$$O^5 = \sum_i \bar{\Upsilon}_i z_i, \quad i=1,2. \tag{26}$$

In this paper, ANFIS is used to forecast the next hour price. The triangular-shaped fuzzy membership functions and Sugeno inference engine are considered.

2.5. Tertiary forecast model

This stage aims to improve the accuracy of the previous price and load forecasting stages. The third stage has ANN and ANFIS machines; the former uses the second stage load and price data as inputs, and the latter uses the improved load forecasting data of the second and third stages as inputs. The parameters of the third stage ANN and ANFIS machines are the same as their corresponding machines in the second stage.

3. Performance evaluation

In most of the papers, Mean Absolute Percentage Error (MAPE) has been used to evaluate the accuracy of forecasting results. The MAPE is defined as [22]:

$$MAPE = \frac{100}{\xi}\sum_{h=1}^{\xi}\frac{|\hat{V}_h - V_h|}{\overline{V}}, \tag{27}$$

Where $\hat{V}_h$ and V_h are, respectively, the forecasted and actual prices or loads at hour h, ξ is the forecast horizon and $\overline{V}$ is the average forecasted value that is expressed as follows:

$$\overline{V} = \frac{1}{\xi}\sum_{h=1}^{\xi} V_h \cdot \tag{28}$$

There is a drawback in *MAPE* criterion; if the mean value of signal *V* is zero, the fraction will be indefinite. The mean value of *D1* and *D2* may be zero; in this case, it is not possible to use *MAPE*, and Mean Absolute Error (*MAE*) is used for this condition instead as below:

$$MAE = \frac{1}{\xi}\sum_{h=1}^{\xi} |\hat{V}_h - V_h| \,. \tag{29}$$

While *MAE* is in the form of *MW* or *$/MWh*, *MAPE* is expressed in percent.

4. Numerical results

The proposed approach was applied to forecast load and price in the electricity market of the DK2 area of Nordpool and the electricity market of the mainland Spain. The algorithm codes were developed in MATLAB, and all of the study cases utilized the neural network, ANFIS and Kohonen SOM toolboxes of MATLAB.

4.1. DK2 area of Nordpool's electricity market load and price forecasting

The data from Nordpool's official website from January 2011 to December 2012 of DK2 area were used to test the proposed price and load forecasting algorithm. The simulation was carried out on a PC (Intel Core 2, 2.93 GHz, 4 GB RAM).

At first, wavelet decomposition was performed. Table 1 shows the level of wavelet decomposition. As seen, the a_2 component, interpreted as the base load, has a large contribution to the original load. The noise components d_1 and d_2 had the effect of less than 5%. Hence, the wavelet decomposition was performed on two levels.

The data window for Kalman was determined according to the correlation between the actual load and the input parameters of forecasting.

Fig. 6 shows a high correlation between a load and its past parameters. For example, when the load in the day *d* and the hour *t* had a perfect linear relation with the hour before, it had not enough correlation with the load of two days before at the same hour. On the other hand, the correlation value between load and price was very little, so it was sorted by the given values. As depicted in Fig. 6, there is no special difference between the correlation of the price of three days before and the same day with the target day load. Table 2 indicates the load and price candidates' correlation.

The number of parameters and candidates was selected according to Table 2, and the correlation values of each price candidate were calculated to select the highest values of correlations.

Then primary load forecasting was performed using the Kalman filter. The frequency outputs' data of wavelet were used for forecasting. Data were delivered to the Kalman filter, and three frequency components of load and price forecast were

found. The starting day was January 1, 2011, and totally, 600 days were considered. Fig. 7 illustrates the outputs of the Kalman filter for June 19, 2012, that was one of the hard cases for accurate load and price forecasting.

In order to recognize similar load patterns and to map them into a two-dimensional grid, the Kalman load and price forecasting results of each frequency were normalized and fed to the SOM network to find similar load among the historical data.

An 8×8 hexagonal lattice Kohonen network was used for each frequency component. The SOM epochs were selected 50, and a selected lattice is shown in Fig. 8.

For example, the similar day load patterns for the D1, D2 and A2 frequency components for June 19, 2012 are:

Similar day load patterns for D1: {126, 152, 161, 167, 175, 179, 511, 517, 524, 528, 529, 530, 531, 532}.

Similar day load patterns for D2: {126, 165, 483, 516, 517, 524, 525, 528, 529, 530, 531}.

Finally, similar day load patterns for A2: {119, 167, 187, 200, 201, 230, 466, 508, 509, 510, 524}.

Figs. 9, 10 and 11 show the clustered load shape by SOM for D1, D2 and A2 frequency components, respectively.

Fig. 12 presents similar days results found by Kohonen for June 19, 2012 load; as shown, there is an acceptable performance of the proposed algorithm in finding the similar days in different components.

Then the MLP ANN was trained based on the loads, prices, and an average temperature of days as inputs. The MLP inputs, hidden layers, and outputs were 25, 20, and 1, respectively. In this paper, *logsig and tansig* transfer functions were considered; *tansig* associates with the lower MAPE in comparison with *logsig* transfer function. The MLP training process was a back-propagation method, and an MLP was used for each hour load forecasting.

Fig. 13 shows the learning procedure of MLP for the low-frequency component of the load for June 19, 2012. It shows that after two epochs, the error reaches to 10^{-8}, and validation tests confirm these results.

Next, load forecasting was performed by the trained MLP for different frequency components. The average temperature of the next day was calculated based on a meteorological forecast. Test data were used to verify the proposed model. Table 3 shows the MAPE values for the test data.

Fig. 14 further compares the real and forecasted loads for different frequency components given by the second stage and shows a very accurate tracking of low-frequency load.

As Figs. 14a, 14b, and 14c depict, D1 and D2 components have less amplitude than the original signal, and their oscillations are faster than the A2 oscillations. In addition, the average amplitudes of D1 and D2 components have negative values.

Then the proposed MI machine was used. The load and price data of the past 50 days were stored in a matrix of size (50×99). The resulting matrix consisted of the following data:

1- The past 24-hours price data (24 elements)

2- The past 24-hours load data (24 elements)

3- The same hour price data for the past 50 days (50 elements)

4- The same hour price data (1 element)

The values of MI were calculated and sorted. Tables 4 and 5 show the values of the calculated MI for 6/19/2012 4:00 PM and 6/19/2012 5:00 PM, respectively.

Next, for every hour and based on Fig. 1, a forecasting procedure was performed by the ANFIS machine. As illustrated in Fig. 6, the correlation between the loads in any hour is very high with the last hours` and days` loads, and very low with their price. Thus, price lost the competition versus load in the ANFIS input selection. Hence, a selection of two load variables and two price variables was proposed to make a trade-off between tracking the trend of price and the best correlation selection. The price forecasting process had better performance, which is depicted in Fig. 15.

Table 6 presents the performance of the proposed method in three stages of price and load forecasting for June 19, 2012. The results showed the improvement in forecasting after using the ANN-ANFIS model to improve the price forecasting.

There were no outstanding differences between the first and second stages` results in $d_1$ and $d_2$ components (Table 6). Obviously, a better performance of the second stage was provided in $a_2$; therefore, the results became remarkably well. This point was also seen in the price forecasting process (Fig. 15 and Table 6). Accordingly, it can be concluded that Kalman`s performance at high frequencies, especially in d_1, was better than that of ANN model for two reasons: 1) High-frequency components had a random pattern, and training data finding was impossible among the historical days, and 2) ANN model`s performance at high frequencies was basically poor; hence, wavelet decomposition was employed.

Final simulation results for this day are depicted in Fig. 16. As shown, there is better tracking of load signal by the third stage forecast and an acceptable prediction for the price. The computation time was about 279 seconds.

Table 7 shows a comparison between one-step-ahead prediction errors for the proposed price forecasting method and four other algorithms [54] for Nordpool electricity market. The same weeks and data as in [54] were selected for comparison of the results.

4.2. Mainland Spain electricity market load and price forecasting

The algorithm was applied to forecast load and price in the electricity market of mainland Spain for two different case studies. For the first case study, the same weeks as in [55] were selected for comparison of the price forecasting results that were four weeks corresponding to four seasons of the year 2002. Table 8 shows a comparison between the proposed price forecasting method and nine other algorithms (AWNN [56], FNN [57], WNN [58], ARIMA [55], Mixed [59], HIS

[60], wavelet-ARIMA [55], MLP-NN [48] and WPA [3]) for the weekly MAPE criteria. For the load and price-forecasting model, a week ahead predictions were computed.

The computation time was about 248 seconds. The proposed algorithm provided a better forecasting accuracy than previous approaches, and improvement of the MAPE index with respect to the nine previous algorithms was 41.4%, 47.34%, 50.08%, 60.24%, 57.41%, 43.18%, 51.17%, 88.77% and 21.89%, respectively.

Then, for the second case study, the SPLF was performed for the electricity market of mainland Spain to predict all the daily loads and prices in summer 2008, autumn 2008, winter 2009 and spring 2009 (365 days), and the same data were selected for comparison of the load and price forecasting results [61].

The Average Seasonal Forecasting Error (ASFE) was used for this case study as below:

$$ASFE = \frac{1}{7}\sum_{j=1}^{7}\frac{1}{I_j}\frac{1}{24}\sum_{i=1}^{I_j}\sum_{t=1}^{24}100\frac{|V_{ij}(t)-\hat{V}_{ij}(t)|}{V_{ij}(t)}, \quad (30)$$

Where I_j is the number of weeks containing the jth day, $V_{ij}(t)$ $\hat{V}_{ij}(t)$ are actual and the forecasted values (load or price), respectively [61].

Table 9 compares the proposed load and price forecasting method with five other algorithms (ARIMA [61], FNP-D [61], FNP-R [61], SFPL-D [61], and SFPL-R [61]) for the average seasonal forecasting error values.

The proposed algorithm provided a better forecasting accuracy than previous approaches, and improvement of the seasonal forecasting error index with respect to the five previous load forecasting algorithms was 25.63%, 36.23%, 29.6%, 23.25% and 17.5%, respectively. In addition, improvement of the seasonal forecasting error index with respect to the five previous price forecasting algorithms was 49%, 52.75%, 52.85%, 46.78%, and 45.38%, respectively. Thus, the results show that the three-stage algorithm presents a better performance for day-ahead electricity market load and price forecasting.

5. Conclusions

In this paper, a new intelligent hybrid model was presented for SPLF based on feature selection. The SPLF was organized based on a three-stage algorithm. At the first stage, a wavelet-Kalman-Kohonen model was introduced to make a primary forecast of load and price, and the results were used to generate the proper data set for MLP-ANN training. At the second stage, an MI method was utilized to train ANFIS for price forecasting. The ANFIS machine improved the first stage results by improving the accuracy of price forecasting. In addition, using wavelet decomposition, ANN performance was improved by cancelling the noise component in the base load. Finally, at the third stage, the accuracy of load and price forecasting was improved using MLP-ANN and ANFIS machines for load and price forecasting,

respectively. Simultaneous prediction of load and price has been ignored by most of the papers, and each one has been assumed known to predict the other one. Unlike the load patterns, price acts like noise, and thus it is difficult to find a proper ANFIS training set of inputs. In this paper, an MI method was used to find the best datasets to train ANFIS for price forecasting. The proposed algorithm was applied to Nordpool and mainland Spain power markets; the results showed that MAPE was about 5% and 4%, respectively. Thus, the proposed algorithm provided a better forecasting accuracy than previous approaches for both electricity markets. At the same time, as future work, some modifications can be performed for better feature selection in order to enhance the efficiency of the proposed method.

Acknowledgements

J.P.S. Catalão acknowledges the support by FEDER funds through COMPETE 2020 and by Portuguese funds through FCT, under Projects SAICT-PAC/0004/2015 - POCI-01-0145-FEDER-016434, POCI-01-0145-FEDER-006961, UID/EEA/50014/2013, UID/CEC/50021/2013, UID/EMS/00151/2013, and 02/SAICT/2017 - POCI-01-0145-FEDER-029803, and also funding from the EU 7th Framework Programme FP7/2007-2013 under GA no. 309048.

References

[1] Julián Moral-Carcedo, Julián Pérez-García, Integrating long-term economic scenarios into peak load forecasting: An application to Spain, Energy, 140 (2017) 682-695.

[2] Fatemeh Chahkoutahi, Mehdi Khashei, A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting, Energy, 140 (2017) 988-1004.

[3] Catalao, J.P.S. , Pousinho, H.M.I. , Mendes, V.M.F. , Hybrid Wavelet-PSO-ANFIS approach for short-term electricity prices forecasting, IEEE Trans. Power Syst. 26 (1) (2011) 137 – 144.

[4] A. Ghasemi, H. Shayeghi, M. Moradzadeh, M. Nooshyar. A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management. Appl. Energy, 177 (2016) 40-59.

[5] Ioannis P. Panapakidis, Athanasios S. Dagoumas. Day-ahead electricity price forecasting via the application of artificial neural network based models. Appl. Energy, 172 (2016) 132-151.

[6] Liye Xiao, Wei Shao, Chen Wang, Kequan Zhang, Haiyan Lu, Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting, Appl. Energy, 180 (2016) 213-233.

[7] Yaoyao H, Rui Liu, Haiyan Li, Shuo Wang, Xiaofen Lu. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl. Energy, 185 (2017) 254-266.

[8] Gupta, P.C., Yamada, K. , Adaptive short-term load forecasting of hourly load using weather information, IEEE Trans. Power Syst. 91 (5) (2007) 2085 – 2094.

[9] Li Wei , Zhang zhen-gang, Based on time sequence of ARIMA model in application of short-term electricity load forecasting, International Conference on Research Challenges in Computer Science (2009) 11 – 14.

[10] Christiaanse, W.R., Short-term load forecasting using general exponential smoothing, IEEE Trans. Power Apparat. Syst.90 (2) (1971) 900 – 911.

[11] Irisarri, G.D., Widergren, S.E., Yehsakul, On-line load forecasting for energy control center application, IEEE Trans. Power Apparat. Syst. 101 (1) (1982) 71-78.

[12] F. Meslier., New advances in short-term load forecasting using Box and Jenkins approach, IEEE/PES Winter Meeting, 1978.

[13] Ping-Feng Pai , Wei-Chiang Hong , Support vector machines with simulated annealing algorithms in electricity load forecasting, Energy Convers. Manage. 46 (17) (2005) 2669–2688.

[14] Jin Xing Che, Jian Zhou Wang, Yu Juan Tang, Optimal training subset in a support vector regression electric load forecasting model, Appl. Soft Comput. 12 (5) (2012) 1523–1531.

[15] Rustum Mamlook, Omar Badran, Emad Abdulhadi, A fuzzy inference model for short-term load forecasting, Energy Policy 37 (4) (2009) 1239–1248.

[16] Ku-Long Ho, Hsu, Yuan-Yih , Chuan-Fu Chen, Tzong-En Lee ,Short term load forecasting of Taiwan power system using a knowledge-based expert system, IEEE Trans. Power Syst. 5 (4) (1990) 1214 – 1221.

[17] Lee, K.Y. , Cha, Y.T., Park, J.H. , Short-term load forecasting using an artificial neural network, IEEE Trans. Power Syst. 7 (1) (1992) 124 - 132.

[18] Sajjad Kouhi, Farshid Keynia, A new cascade NN based method to short-term load forecast in deregulated electricity market, Energy Convers. Manage. 71 (2013) 76-83.

[19] Mansour Sheikhan, Najmeh Mohammadi, Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection, Neural Comput. and Appl. 21 (8) (2012) 1961-1970.

[20] Dongxiao Niu, Bingen Kou, Yunyun Zhang, Zhihong Gu A short-term load forecasting model based on LS-SVM optimized by inertia weight particle swarm optimization algorithm , Advances in Neural Networks 5552 (2009) 242-250.

[21] Goldberg DE, Genetic algorithms in search, optimization, and machine learning, Addison Wesley (New York) (1989).

[22] J.P.S. Catalão, H.M.I. Pousinho, V.M.F. Mendes, Short-term electricity prices forecasting in a competitive market by a hybrid intelligent approach, Energy Convers. Manage. 52 (2) (2011) 1061–1065.

[23] Pandey, A.S., Singh, D., Sinha, S.K. Intelligent hybrid models for short-term load forecasting, IEEE Trans. Power Syst. 25 (3) (2010) 1266 – 1273.

[24] Nima Amjady, Ali Daraeepour, Mixed price and load forecasting of electricity markets by a new iterative prediction method, Electr. Power Syst. Res. 79 (9) (2009) 1329–1336.

[25] A. Karsaz,H. Rajabi Mashhadi, M.M. Mirsalehi, Market clearing price and load forecasting using cooperative co-evolutionary approach, Int. j. of Electr. Power & Energy Syst. 32 (2010) 408–415.

[26] Eugene A. Feinberg, Dora Genethliou, Load forecasting, Applied Mathematics for Restructured Electric Power System, Springer (2005) 269-285.

[27] Ying Chen,Luh, P.B. , Che Guan, Yige Zhao ,Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks, IEEE Trans. Power Syst. 25 (1) (2010) 322 – 330.

[28] T. Rashid, T. Kechadi, A practical approach for electricity load forecasting, Int. J. of Electrical, Electronic Science & Engineering 1 (5) (2007) 722-726.

[29] Zhang Yun, Zhou Quan, Sun Caixin, Lei Shaolan, Liu Yumingm Song Yang ,RFB Neural Network and ANFIS-Based Short-Term Load forecasting Approach in Real-Time Price Environment, IEEE Trans. Power Syst. 23 (3) (2008), 853-858.

[30] Khotanzad, A., Enwang Zhou, Elragal, H. A Neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment, IEEE Trans. Power Syst. 17 (4) (2002) 1273-1282.

[31] Samsher Kadir Sheikh, M. G. Unde ,Short-Term Load Forecasting using ANN technique, Int. J. of Engineering Sciences & Emerging Technologies 1 (2) (2012) 97-107.

[32] Amjady, N., Keynia, F. ,Day-ahead price forecasting of electricity markets by mutual information technique and cascaded Neuro evolutionary algorithm, IEEE Trans. Power Syst. 24 (1) (2009) 306 – 318.

[33] Sebastián Maldonado, Richard Weber, A wrapper method for feature selection using Support Vector Machines, Information Sciences 179 (13) (2009) 2208–2217.

[34] Long Yu, Yihui Zheng, Xin Wang, Lixue Li, Gang Yao, Hongtao Chen, Short-term load forecasting based on fuzzy clustering analysis similar days, unifying electrical engineering and electronics engineering 238 (2014) 191-198.

[35] Abdollah K, Haidar S, Fatemeh M. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst. Appl. 41 (2014) 6047–6056.

[36] Selakov A, Cvijetinovic D, Milovic L, Mellon S, Bekut D. Hybrid PSO–SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank. Appl. Soft Comput 16 (2014) 80–88.

[37] Liao G. Hybrid improved differential evolution and wavelet neural network with load forecasting problem of air conditioning. Electr. Power Energy Syst. 61 (2014) 673–682.

[38] Stéphane Mallat, Wavelet tour of signal processing, Academic Press (2008).

[39] Song Li , Lalit Goel, Peng Wang, An ensemble approach for short-term load forecasting by extreme learning machine. Appl. Energy, 170 (2016) 22-29.

[40] H.M. Al-Hamadi, S.A. Soliman, Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model, Electr. Power Syst. Res. 68 (1) (2004) 47-59.

[41] Teuvo Kohonen, Self-organization and associative memory, Springer Series in Information Sciences 8 (1989).

[42] M. López, S. Valero, C. Senabre, J. Aparicio, A. Gabaldon, Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study, Electr. Power Syst. Res. 91 (2012) 18–27.

[43] G. Brown, A. Pocock, M. Zhao, and M. Luján., Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. JMLR 13 (2012), 27–66.

[44] Farshid Keynia , A new feature selection algorithm and composite neural network for electricity price forecasting, Engineering Applications of Artificial Intelligence 25 (8) (2012) 1687–1697.

[45] L.V. Fausett, Fundamentals of neural networks: Architectures, algorithms and applications, The prentice hall press, 1994.

[46] Michael Arbib, The handbook of brain theory and neural networks, The MIT press, 2003.

[47] M.T. Hagan, H.B. Demuth, M.H. Beale, O. De Jesus, Neural network design, IEEE press series on biomedical engineering, 2014.

[48] J.P.S. Catalão, S.J.P.S. Mariano, V.M.F. Mendes, L.A.F.M. Ferreira, Short-term electricity prices forecasting in a competitive market: A neural network approach, Electr. Power Syst. Res. 77 (10) (2007) 1297–1304.

[49] Anbazhagan, N. Kumarappan, Day-ahead deregulated electricity market price classification using neural network input featured by DCT, Electrical Power and Energy Systems 37 (2012) 103–109.

[50] D. J.C. Sousa, L.P. Neves, H.M. Jorge, Assessing the relevance of load profiling information in electrical load forecasting based on neural network models, Electrical Power and Energy Systems 40 (2012) 85–93.

[51] Vitor N. Coelho and et.al. A self-adaptive evolutionary fuzzy model for load forecasting problems on smart grid environment. Appl. Energy, 169 (2016) 567-584.

[52] Li-Chih Ying, Mei-Chiu Pan, Using adaptive network based fuzzy inference system to forecast regional electricity loads, Energy Convers. Manage. 49 (2) (2008) 205–211.

[53] Ching-Hsue Cheng, Liang-Ying Wei, One Step-ahead ANFIS time series model for forecasting electricity loads, Optimization and Engineering 11 (2) (2011) 303-317.

[54] Najeh Chaabane, A hybrid ARFIMA and neural network model for electricity price prediction, Int. j. of Electr. Power & Energy Syst. 55 (2014) 187–194.

[55] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18(3) (2003) 1014–1020.

[56] N.M. Pindoriya, S.N. Singh, S.K. Singh, An adaptive wavelet neural network based energy price forecasting in electricity markets, IEEE Trans. Power Syst. 23 (3) (2008) 1423–1432.

[57] N. Amjady, Day-ahead price forecasting of electricity markets by a new fuzzy neural network, IEEE Trans. Power Syst. 21 (2) (2006) 887–896.

[58] A. T. Lora, J. M. R. Santos, A. G. Expósito, J. L. M. Ramos, and J. C. R. Santos, Electricity market price forecasting based on weighted nearest neighbors techniques, IEEE Trans. Power Syst. 22 (3) (2007) 1294–1301.

[59] C. Garcia-Martos, J. Rodriguez, M.J. Sanchez, Mixed models for short-run forecasting of electricity prices: application for the Spanish market, IEEE Trans. Power Syst. 22 (2) (2007) 544–551.

[60] N. Amjady and H. Hemmati, Day-ahead price forecasting of electricity markets by a hybrid intelligent system, Eur. Trans. Elect. Power, 19(1) (2009) 89–102.

[61] Vilar Fernández, Juan Manuel; Cao, R. and Aneiros Pérez, German, Forecasting next-day electricity demand and price using nonparametric functional methods. Int. j. of Electr. Power & Energy Syst. 39 (2012) 48–55.

Figures:

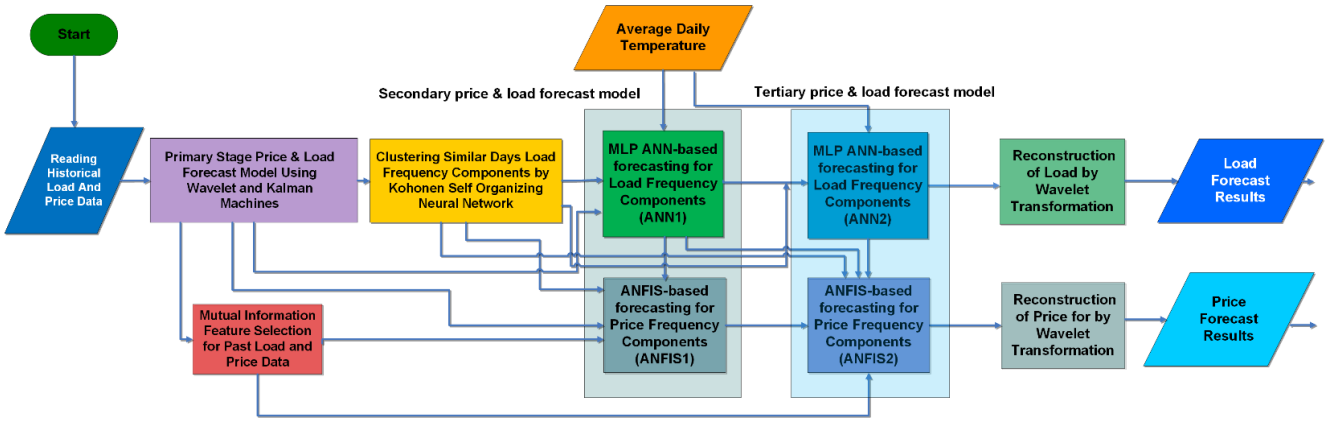

Fig.1. Overall flowchart of proposed forecasting structure.

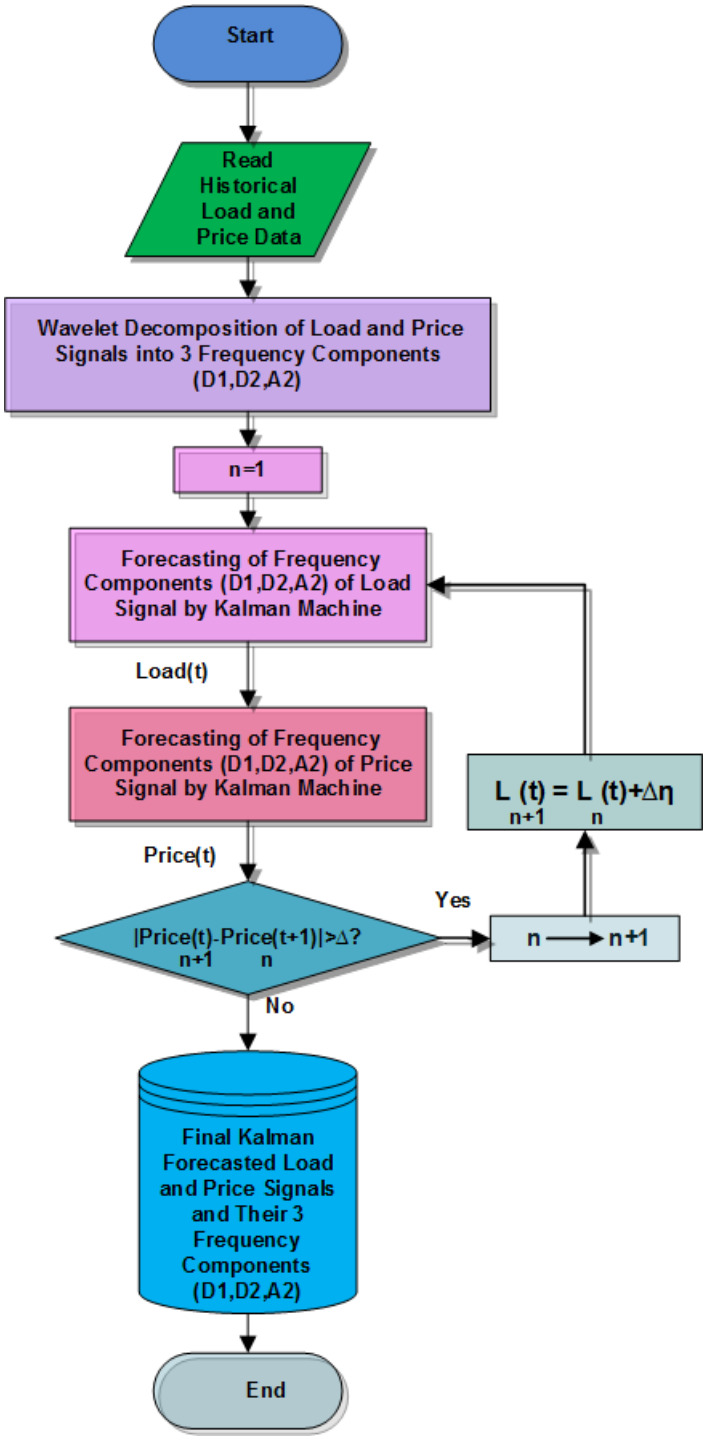

Fig. 2. The proposed first stage iterative load and price forecast model.

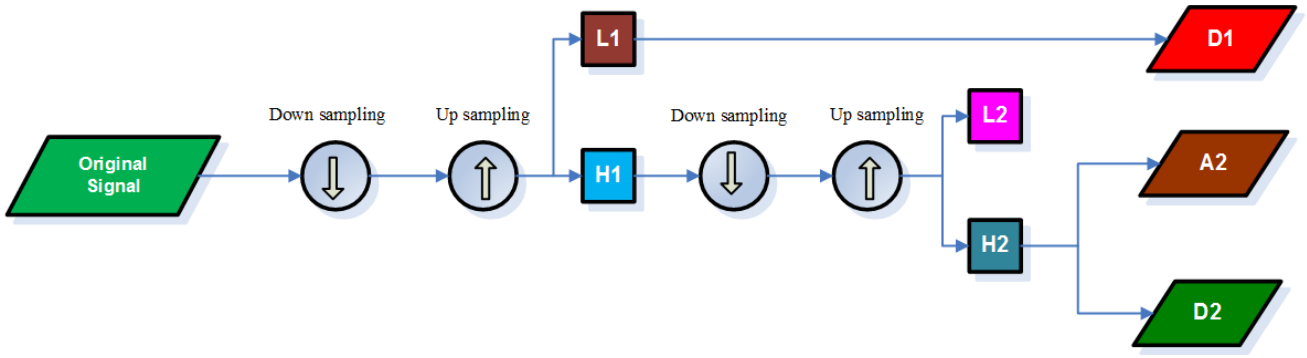

Fig. 3. Two-level wavelet decomposition of signal.

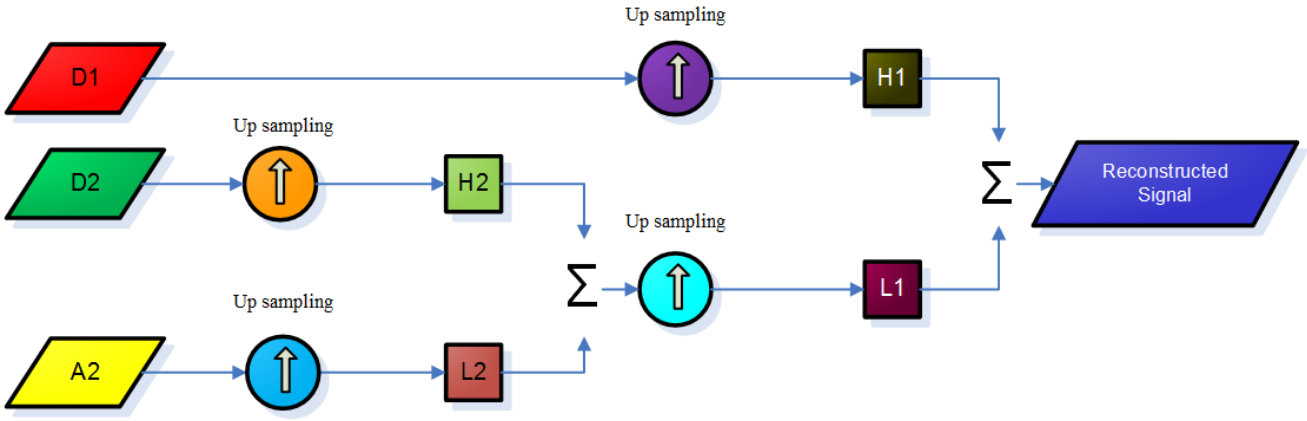

Fig. 4. Wavelet reconstruction of signal.

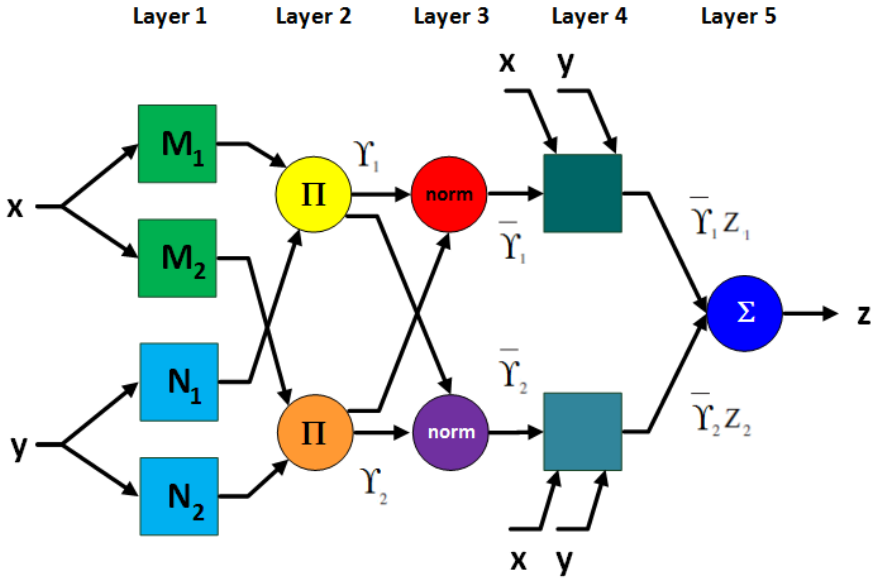

Fig.5. ANFIS structure.

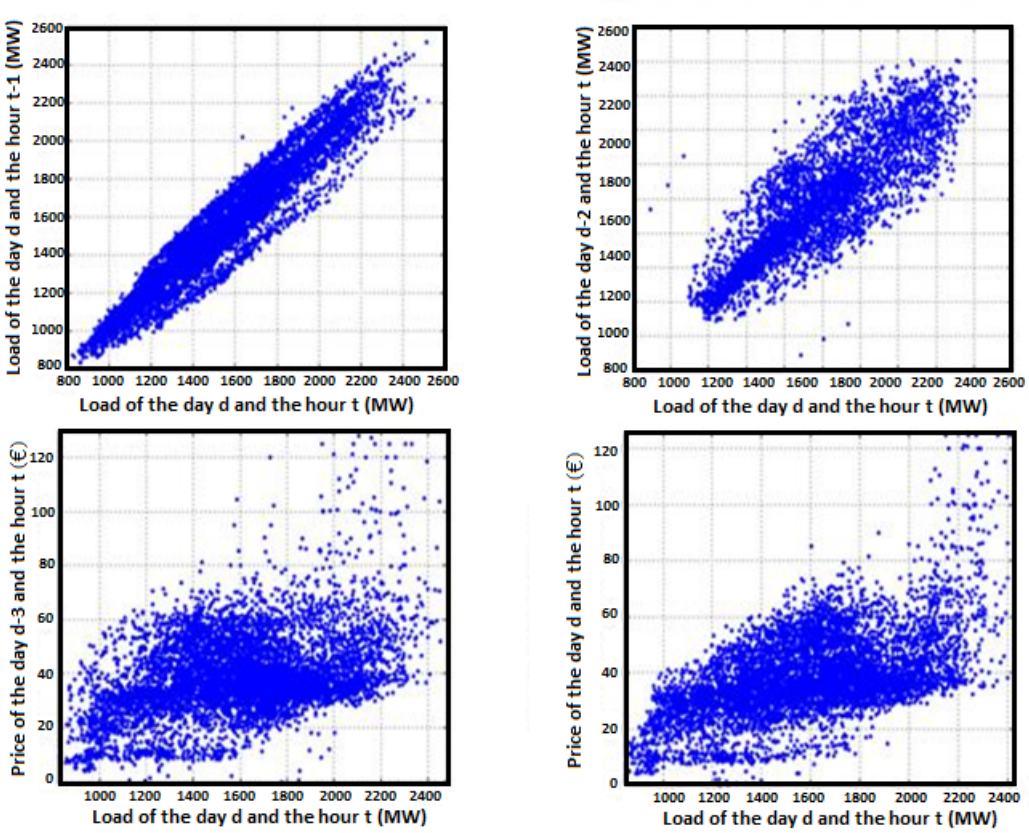

Fig. 6. Dependency between the target and historical loads.

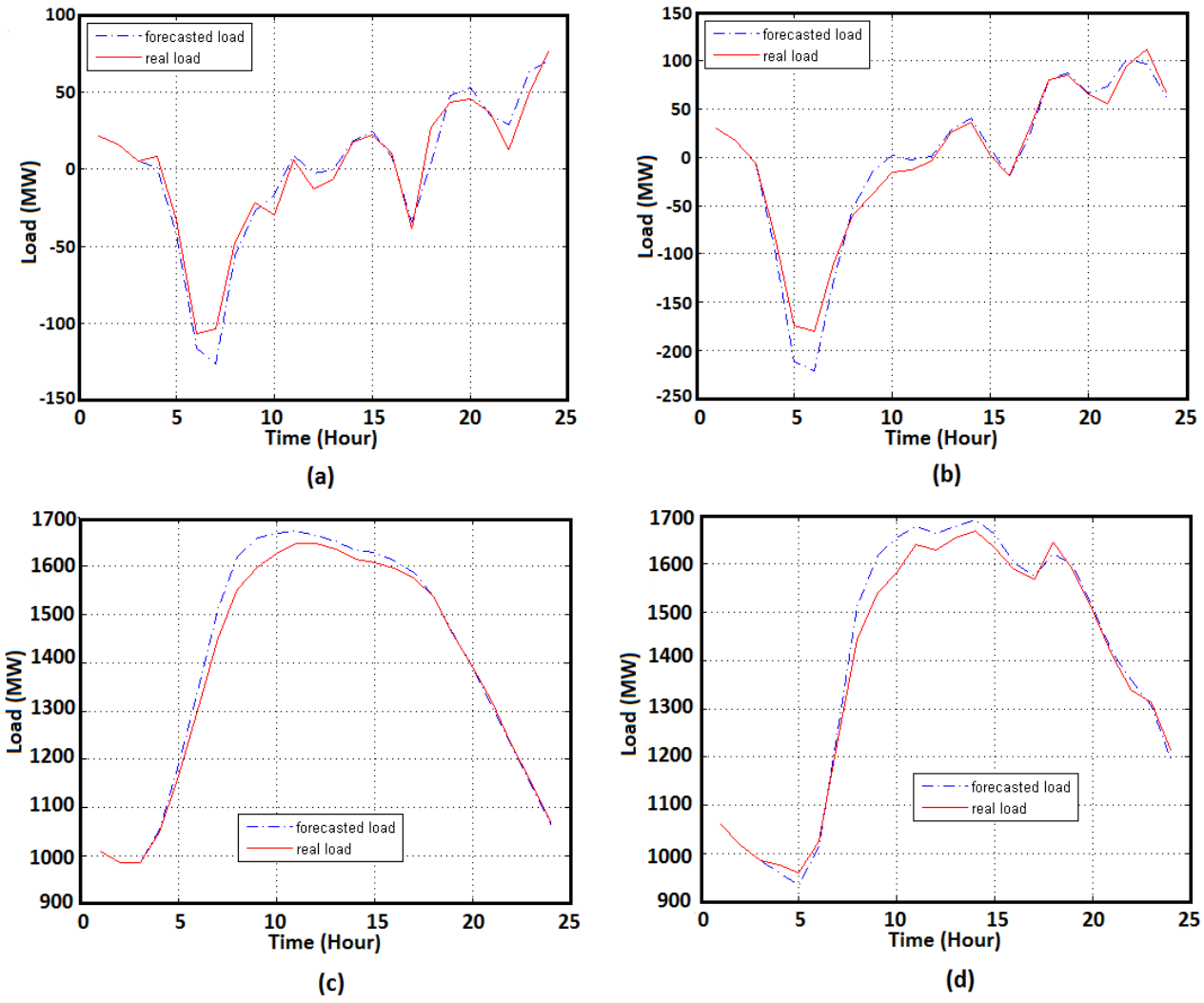

Fig. 7. Primary load forecasting graphs by Kalman for 19 June 2012
(a) A2 frequency component of load, (b) D2 frequency component of load, (c) D1 frequency component of load, and (d) Kalman forecasted load.

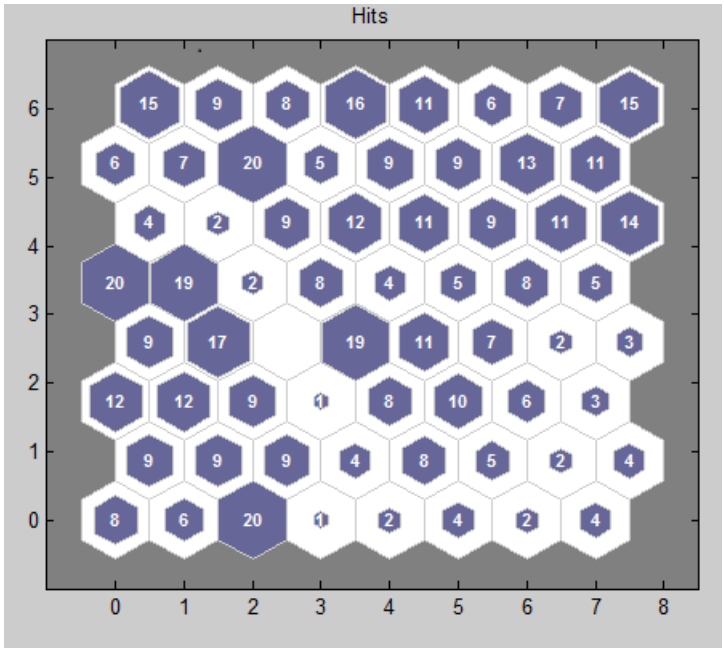

Fig. 8. The 8×8 hexagonal lattice Kohonen network**.**

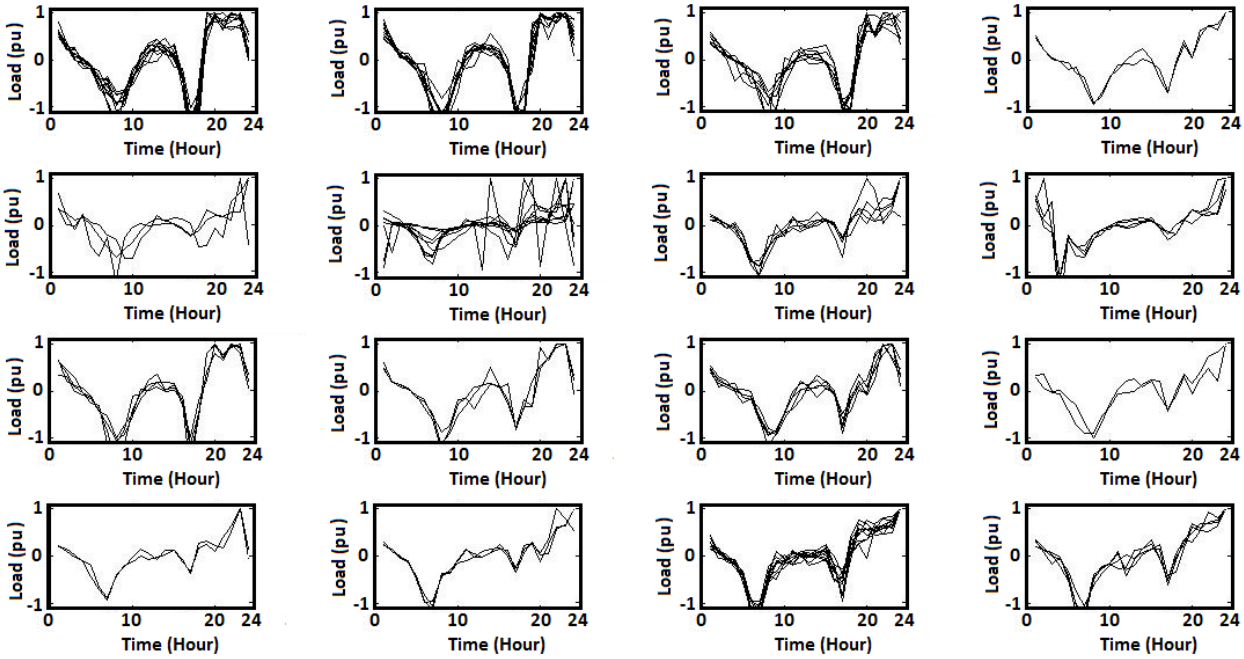

Fig. 9. The clustered load shape for D1 frequency component of load.

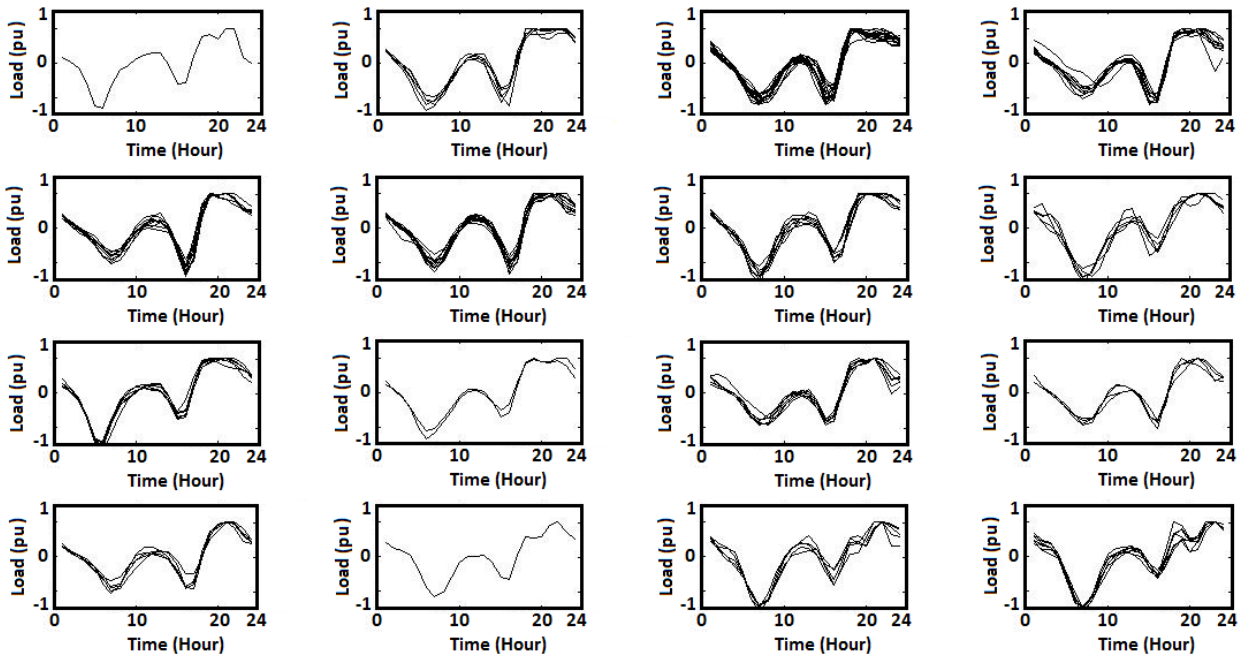

Fig. 10. The clustered load shape for D2 frequency component of load.

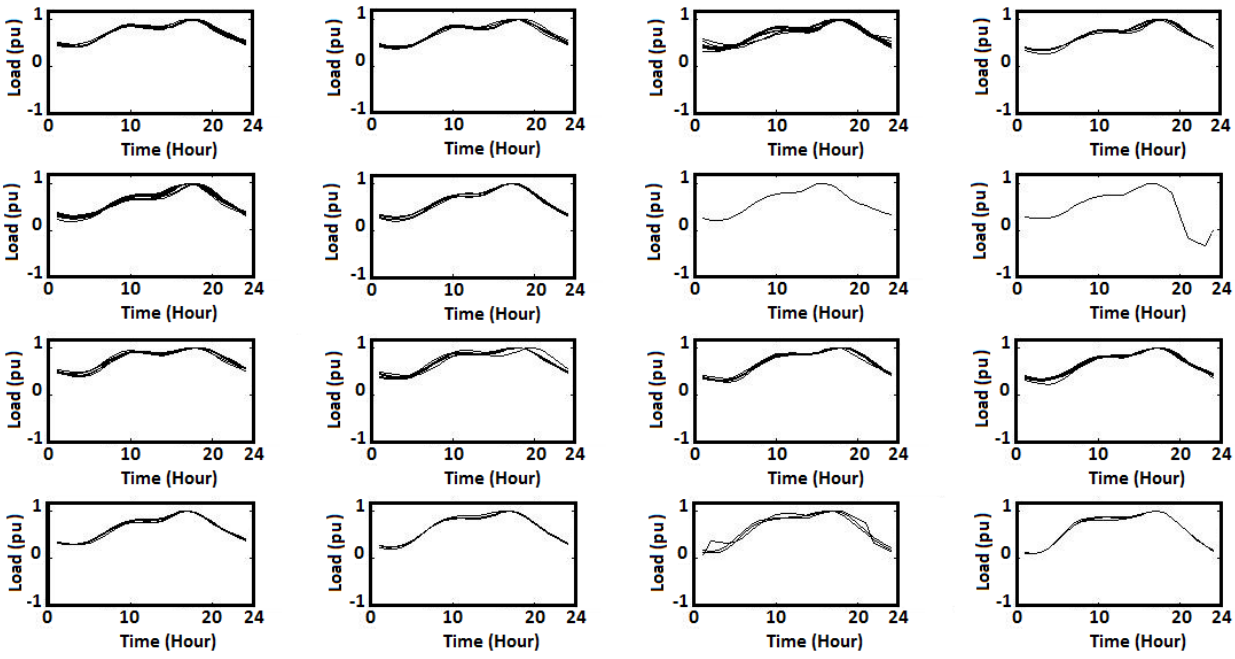

Fig. 11. The clustered load shape for A2 frequency component of load.

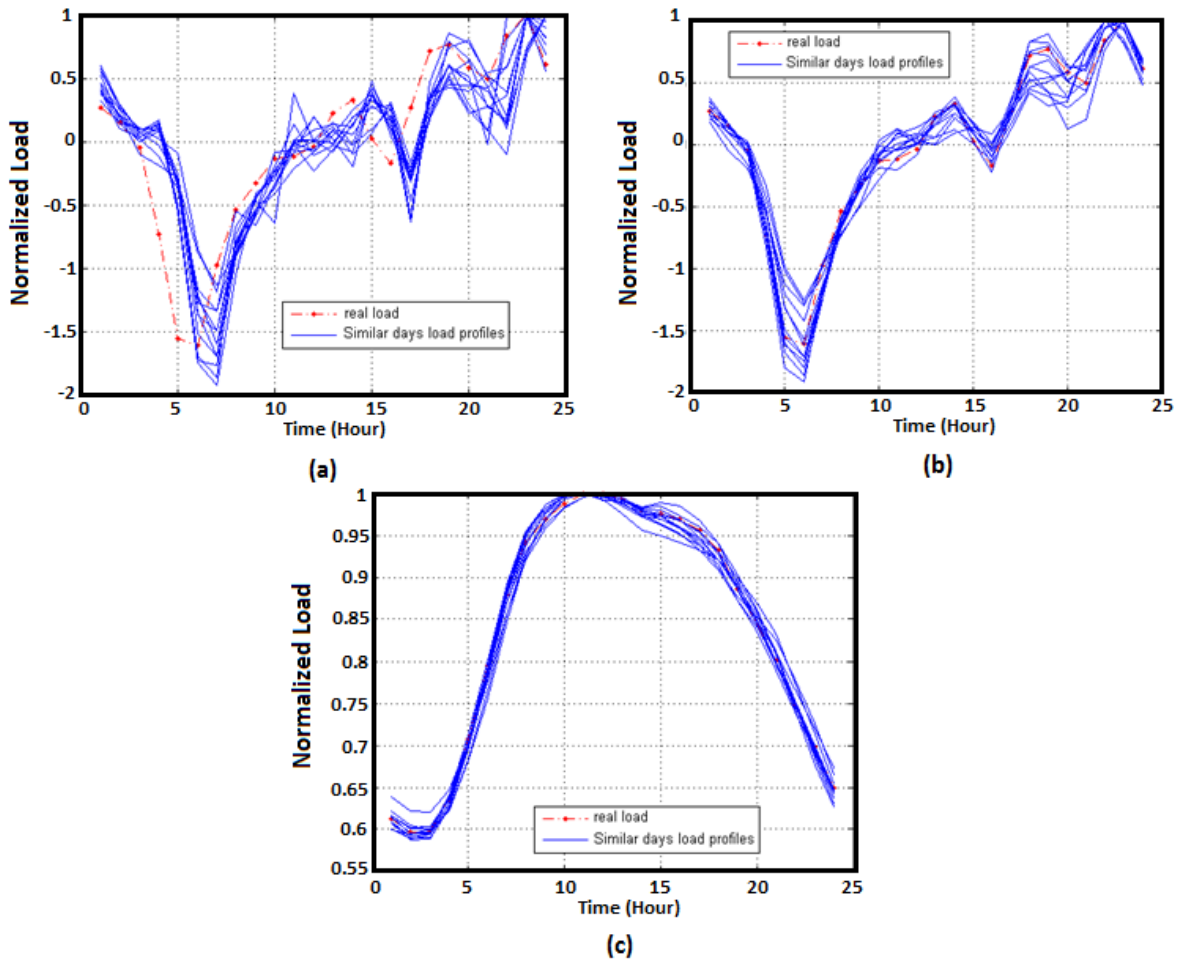

Fig. 12. Similar days finding results by Kohonen for 19 June 2012 load:
(a) D1 frequency component of load, (b) D2 frequency component of load, and (c) A2 frequency component of load.

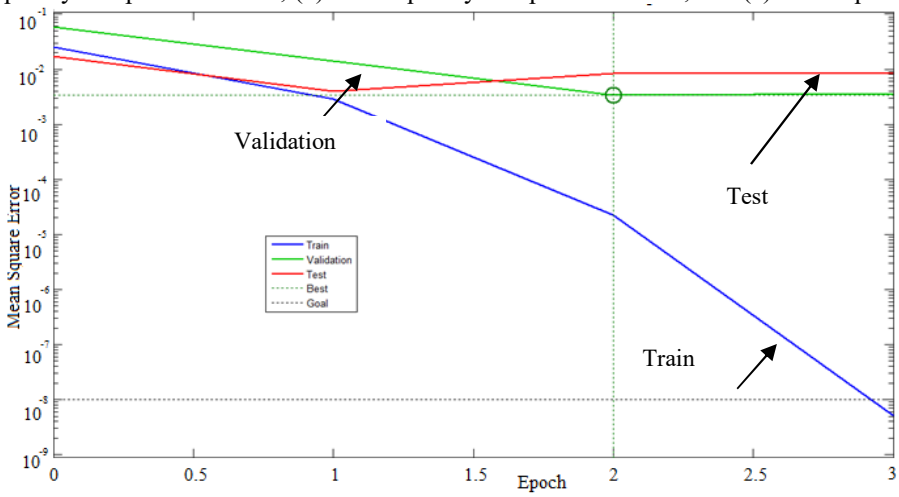

Fig. 13. The learning procedure of MLP ANN for June 19, 2012.

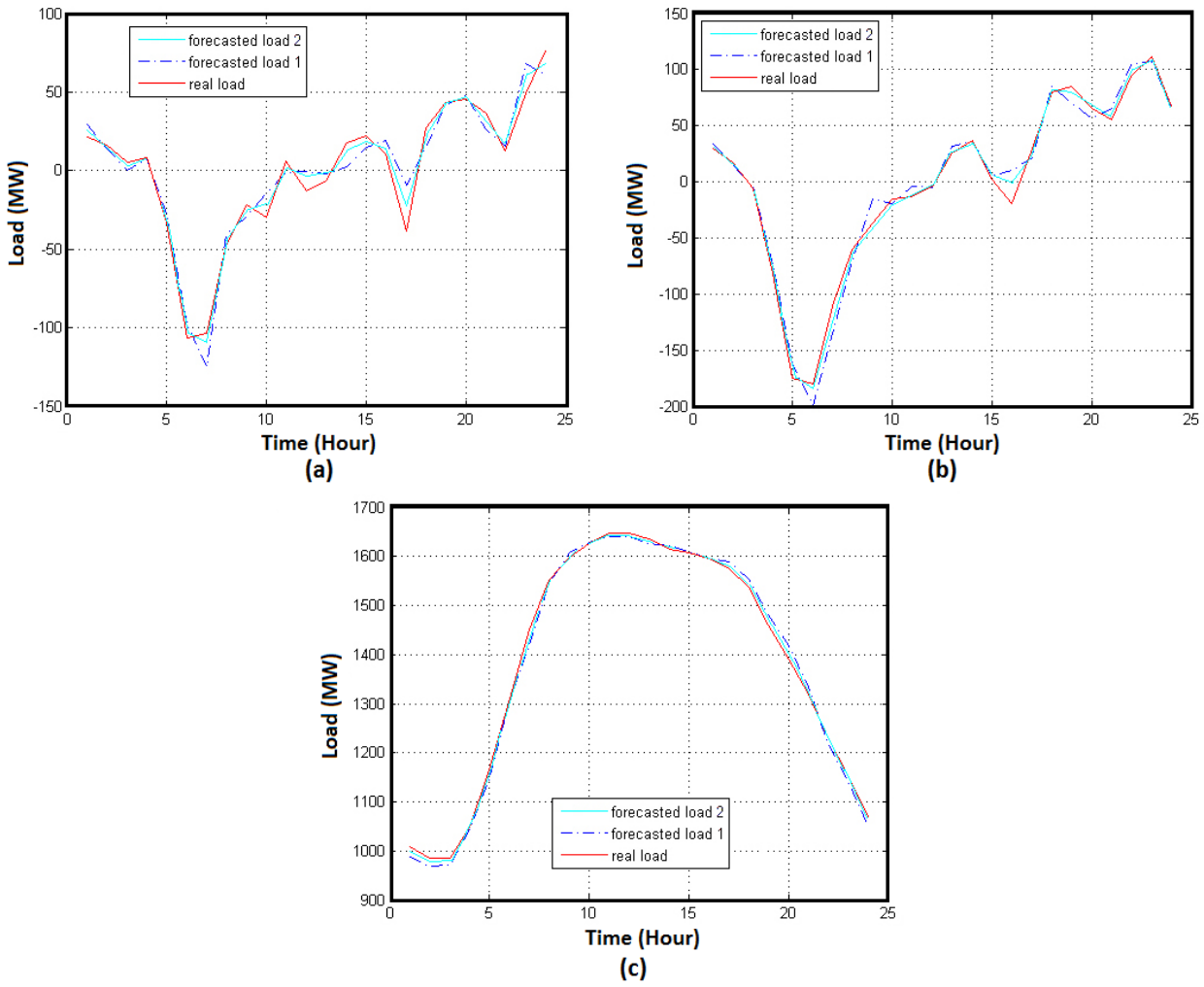

Fig. 14. Second and third stage load forecasting results in different components for 19 June 2012:
(a) D1 frequency component of load, (b) D2 frequency component of load, and (c) A2 frequency component of load.

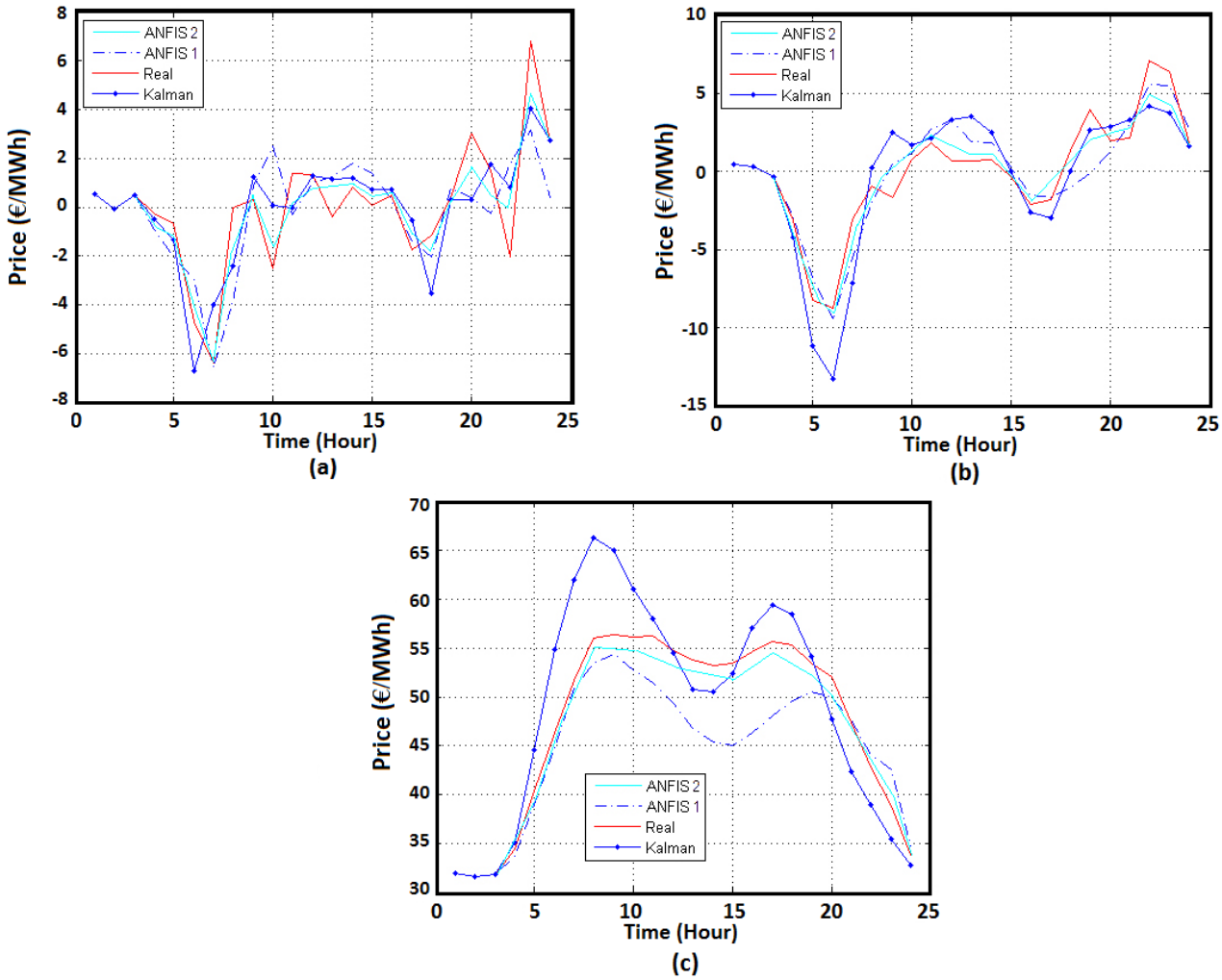

Fig. 15. Price forecasting results in different frequency components given by the second and third stage: (a) D1 frequency component of price, (b) D2 frequency component of price, and (c) A2 frequency component of price.

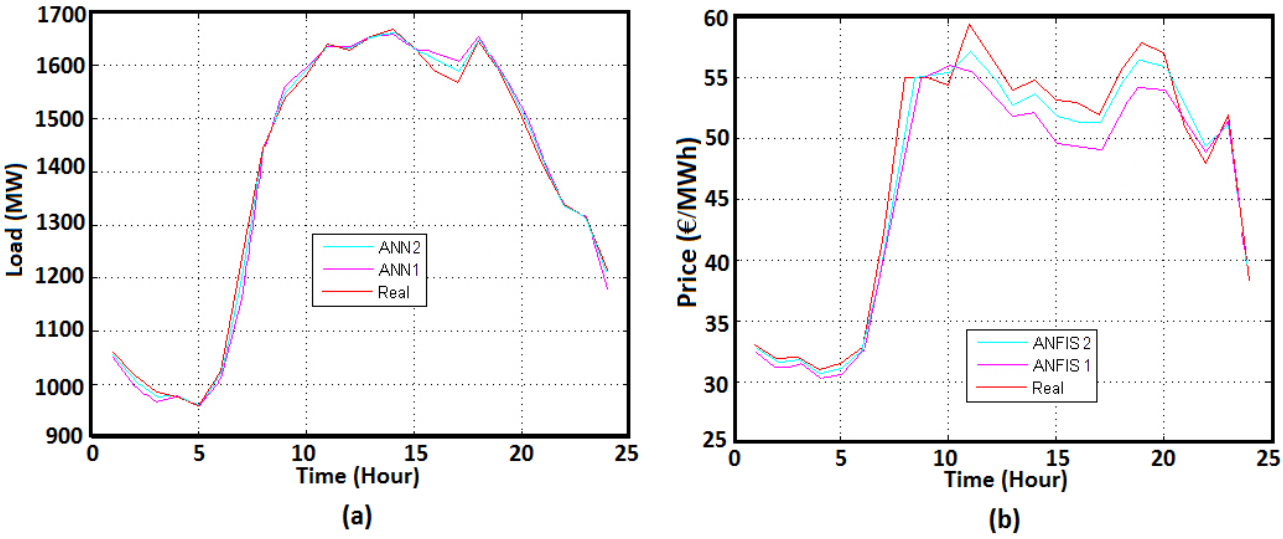

Fig. 16. Final simulation results for 19 June 2012: (a) load, and (b) price.

Tables:

Table 1. Wavelet decomposed frequency components in 2012 DK2 load profiles.

Month	a_2	d_1	d_2
January	0.9524	0.0389	0.0674
February	0.9535	0.0401	0.0667
March	0.9672	0.0414	0.0715
April	0.9832	0.0406	0.0683
May	0.9829	0.0419	0.0655
June	0.9820	0.0381	0.0595
July	0.9870	0.0372	0.0595
August	0.9880	0.0405	0.0702
September	0.9812	0.0432	0.0763
October	0.9595	0.0424	0.0741
November	0.9497	0.0389	0.0703
December	0.9641	0.0384	0.0658
Aver.	0.9709	0.0401	0.0679

Table 2.Load and price candidates' correlation.

Candidate input	Correlation	Candidate input	Correlation
$P_{d,t}$	0.3206	$P_{d-1,t-3}$	0.2947
$P_{d,t-1}$	0.3254	$P_{d-1,t-4}$	0.2989
$P_{d,t-2}$	0.3189	$P_{d-1,t-5}$	0.2997
$P_{d,t-3}$	0.3110	$P_{d-2,t-1}$	0.2523
$P_{d,t-4}$	0.3032	$P_{d-2,t-2}$	0.2528
$P_{d,t-5}$	0.2952	$P_{d-2,t-3}$	0.2528
$P_{d,t-6}$	0.2863	$P_{d-2,t-4}$	0.2525
$P_{d,t-7}$	0.2765	$P_{d-7,t-1}$	0.4925
$P_{d,t-8}$	0.2670	$P_{d-7,t-2}$	0.5028
$P_{d,t-9}$	0.2587	$P_{d-7,t-3}$	0.5133
$P_{d-1,t}$	0.2803	$P_{d-7,t-4}$	0.5222
$P_{d-1,t-1}$	0.2850	$P_{d-8,t-1}$	0.4754
$P_{d-1,t-2}$	0.2898	$P_{d-8,t-2}$	0.4689

Table 3. MAPE of MLP ANN for load forecasting.

Month	Max MAPE	Min MAPE	Month	Max MAPE	Min MAPE
January	2.03	1.12	August	2.59	2.4
February	3.21	2.19	September	1.31	1.51
March	3.32	2.24	October	2.68	1.81
April	2.1	1.62	November	2.29	1.35
May	2.47	1.71	December	2.07	2.05
June	2.20	1.69	Average	2.34	1.75
July	1.85	1.34			

Table 4. The values of the calculated MI for 6/19/2012 4:00 PM

Rank	Candidate	MI	Rank	Candidate	MI	Rank	Candidate	MI
1	P(d,t-1)	0.4706	7	P(d,t-22)	0.0956	13	L(d,t)	0.0573
2	P(d,t-2)	0.4706	8	P(d,t-23)	0.0956	14	L(d,t-1)	0.0573
3	P(d,t-3)	0.4615	9	P(d-12,t)	0.0956	15	L(d,t-2)	0.0573
4	P(d-21,t)	0.2769	10	P(d-46,t)	0.0956	16	L(d,t-3)	0.0573
5	P(d,t-21)	0.1445	11	P(d,t-10)	0.0573	17	L(d,t-23)	0.0573
6	P(d,t-20)	0.0943	12	P(d,t-11)	0.0573	18	L(d,t-24)	0.0573

Table 5. The values of the calculated MI for 6/19/2012 5:00 PM

Rank	Candidate	MI	Rank	Candidate	MI	Rank	Candidate	MI
1	P(d,t-1)	0.2781	7	P(d,t-22)	0.0956	13	L(d,t-6)	0.0573
2	P(d,t-4)	0.1445	8	P(d-25,t)	0.0956	14	L(d,t-24)	0.0573
3	P(d-14,t)	0.1438	9	P(d,t-9)	0.0573	15	P(d-15,t)	0.0573
4	P(d-22,t)	0.1441	10	P(d,t-12)	0.0573	16	P(d-29,t)	0.0573
5	P(d,t-2)	0.0957	11	L(d,t)	0.0573	17	P(d-32,t)	0.0573
6	P(d,t-3)	0.0956	12	L(d,t-5)	0.0573	18	P(d-40,t)	0.0573

Table 6. Comparing MAE of the first, second and third stage results (19 June 2012)

	First Stage		Second Stage		Third Stage	
Frequency Component	MAE (MW)	MAE (€/MWh)	MAE (MW)	MAE (€/MWh)	MAE (MW)	MAE (€/MWh)
A2	18.76	3.48	12.18	3.23	9.74	3.07
D2	10.19	1.59	9.13	1.09	7.26	0.86
D1	7.34	1.13	6.54	0.97	5.81	0.74
Total	22.89	3.28	20.18	3.07	17.59	2.78

Table 7. One-step ahead prediction MAPE for the proposed price forecasting method and four other algorithms for Nordpool electricity market.

	ARFIMA-ANN[52]	ANN[52]	ARFIMA [52]	Zhang's Hybrid model[52]	Proposed
MAPE%	6.47	10.93	13.89	9.23	4.06

Table 8. Weekly MAPE values in terms of percentage (%) for proposed price forecasting method and nine other algorithms for 4 weeks of the Spanish electricity market in year 2002.

Test Week	AWNN[54]	FNN[55]	WNN[56]	ARIMA[53]	Mixed Model[57]	HIS[58]	Wavelet-ARIMA[53]	MLP-NN[42]	WPA[1]	Proposed
Winter	3.43	4.62	5.15	6.32	6.15	6.06	4.78	5.23	3.37	2.82
Spring	4.67	5.30	4.34	6.36	4.46	7.07	5.69	5.36	3.91	3.64
Summer	9.64	9.84	10.89	13.39	14.90	7.47	10.70	11.40	6.50	4.98
Fall	9.29	10.32	11.83	13.78	11.68	7.30	11.27	13.65	6.51	4.11
Average	6.76	7.52	8.05	9.96	9.30	6.97	8.11	8.91	5.07	3.88

Table 9. Average seasonal forecasting error values in terms of percentage (%) for proposed load and price forecasting method and five other algorithms for the Spanish electricity market in year 2008-2009.

	Load forecasting seasonal forecasting errors (%)						Price forecasting seasonal forecasting errors (%)					
	ARIMA	FNP-D	FNP-R	SFPL-D	SFPL-R	Proposed	ARIMA	FNP-D	FNP-R	SFPL-D	SFPL-R	Proposed
Summer	2.55	3.04	3.08	2.78	2.59	2.31	6.78	7.21	7.24	6.33	6.41	3.78
Fall	3.27	4.15	3.82	3.50	3.20	2.59	7.96	8.76	8.58	7.49	7.44	4.38
Winter	4.27	3.86	3.44	3.55	3.43	2.71	11.64	12.16	12.25	11.30	10.95	5.16
Spring	4.14	5.51	4.69	3.96	3.60	2.95	7.98	8.97	9.10	7.82	7.29	3.97
Average	3.55	4.14	3.75	3.44	3.20	2.64	8.59	9.27	9.29	8.23	8.02	4.32