Applied Energy 172 (2016) 132–151

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Day-ahead electricity price forecasting via the application of artificial neural network based models

Ioannis P. Panapakidis [a,b,*], Athanasios S. Dagoumas [b]

[a] Department of Electrical Engineering, Technological Education Institute of Thessaly, 41110 Larisa, Greece
[b] School of Economics, Business & International Studies, University of Piraeus, 80 Karaoli & Dimitriou Str., 18534 Piraeus, Greece

HIGHLIGHTS

- The paper focuses in short-term price load forecasting.
- Several Day-ahead forecasting models are proposed and tested.
- The clustering tool is combined with neural network.
- We focus on no pre-processed data.

ARTICLE INFO

Article history:
Received 2 December 2015
Received in revised form 3 March 2016
Accepted 18 March 2016
Available online 31 March 2016

Keywords:
Artificial neural networks
Electricity price forecasting
Day-ahead market
Time-series clustering

ABSTRACT

Traditionally, short-term electricity price forecasting has been essential for utilities and generation companies. However, the deregulation of electricity markets created a competitive environment and the introduction of new market participants, such as the retailers and aggregators, whose economic viability and profitability highly depends on the spot market price patterns. The aim of this study is to examine artificial neural network (ANN) based models for Day-ahead price forecasting. Specifically, the models refer to the sole application of ANNs or to hybrid models, where the ANN is combined with clustering algorithm. The training data are clustered in homogenous groups and for each cluster, a dedicated forecaster is employed. The proposed models are characterized by comprehensive operation and by high level of flexibility; different inputs can be taken under consideration and different ANN topologies can be examined. The models are tested on a data set that consists of atypical price patterns and many outliers. This approach makes the price forecasting problem a more challenging task, providing evidence that the proposed models can be considered as useful and robust forecasting tools to the actual needs of market participants, including the traditional generation companies and self-producers, but also the retailers/suppliers and aggregators.

1. Introduction

1.1. Motivation

Modern power systems planning includes a variety of resources to cover the increasing demand subject to the various techno-economic and environmental constrains [1]. Load forecasting is of fundamental importance for power systems operation and this fact is reflected by the plethora of related researches. Many methodologies that differ in the data preprocessing, model selection, calibration and testing phases, have been presented [2]. The load forecasting literature is expanded. It includes single models or more sophisticated models that combine various computational intelligence algorithms [3].

On the other hand, the literature on price forecasting is less numerous [4]. This is due to the fact that most markets were structured as monopolies until recently; wholesale competition was absent or limited. While electricity markets become competitive, price forecasting is gathering research momentum. Price forecasting is a relatively more difficult task due to the endogenous characteristics of price time series [5]. Since the determination of the hourly market clearing price (MCP) is held within a dynamic and competitive environment, MCP is characterized by volatility [6]. MCP's chronological evolution is influenced by a set of diverse parameters like demand, fuel prices such as coal and natural gas, merit order of generation plants, hydropower capacity, market

* Corresponding author at: Department of Electrical Engineering, Technological Education Institute of Thessaly, 41110 Larisa, Greece.
E-mail address: ipanap@ee.auth.gr (I.P. Panapakidis).

http://dx.doi.org/10.1016/j.apenergy.2016.03.089
0306-2619/

participants strategies, network congestion and others [7]. Thus, special care should be placed on the inputs selection, model's parameters and calibration, model's assessment and generally, on the experimental set-up that will result in a robust forecaster.

The electricity price forecasting problem influences various processes expanded on different time frames in modern power systems operation. It is closely related with other contemporary scientific and engineering problems such as the optimal power generation units scheduling, fuel consumption, energy resources exploitation, GHG emissions, power systems simulation and electricity demand modeling. Therefore, it is inter-connected with other problems and tasks. Due to this attribute, it is a research topic targeting multi-disciplinary audiences within the power systems community. Also, while price forecasting is related to the energy market transactions and operations, various market participants (utilities, grid operators, retailers, aggregators and others) are interested in obtaining short-term predictions.

The aim of this work is to explore the potential of ANN based models on price forecasting. We examine models built solely on ANN and hybrid models that combine unsupervised machine learning via the clustering tool and supervised machine learning via the ANN. A reliable forecaster should be easy to implement and characterized by high level of parameterization. The analysis of the study is centered toward these goals.

In the price forecasting literature, many models have been proposed based on computational intelligence. The models refer to sole algorithms or more complex hybrid approaches. The present paper examines a set of single and hybrid models on the case of using no preprocessed data. The single models refer to ANNs and more specifically, to Multi-Layered Perceptrons (MLPs) and the hybrid combine MLPs and clustering. In the literature, many robust models such as Fuzzy Neural Network (FNN), Support Vector Machines (SVMs) and Radial Basis Function (RBF) networks. Our preference on the MLPs over the others is justified by the following reasons:

(a) Our focus is mainly on the current data. We adapt the most common approach of the literature (i.e., the MLPs) to examine their performance on special data sets. Also, by using single MLPs, it is easy to examine various configurations that differ in the number and type of inputs. The comparison with more complex models such as the FNN is left for future research.
(b) A basic advantage of the MLPs is their training flexibility. The user can select between a large set of modified back-propagation algorithms in contrast to SVMs and RBFs networks. Different training algorithms apply to various problems, providing the user a flexible and adaptable modeling tool.
(c) The MLPs require less training time contrary to SVMs and RBFs. For example, in the case of the RBFs networks, the number of RBF units equals to the number of patterns. This fact may increase the problem's complexity when dealing with large data sets.

Prior to entering the data into the forecasting models, some researches utilize the wavelet transform to decompose the original signal into low and high-frequency subseries (wavelet domain) [8]. This approach leads to better predictions in some cases. In the present paper, our main focus is to test the accuracy of some models on the raw data coming directly from the metering system. The wavelet analysis appears to be prominent and will be the regarded into a future study by the authors.

1.2. Solution approach

Due to their potential of simulating data with complex and non-linear relationships, ANNs are preferable in cases where a model that describes the data is absent [9]. ANNs are data driven models that are trained with a limited number of data and are to provide a generalization of their operation. A forecaster built on ANN receives as inputs the parameters that influence the quantity under examination, i.e. the MCP. For instance, the inputs include past MCP values, exogenous variables like temperature, fuel prices, day type identification codes and others. The majority of the price forecasting related literature focuses in specific electricity markets with relatively smooth data. Our approach differs from the related literature on the attributes of the data set. The models are tested on a raw data sample that contain null values and have missing entries. At least theoretically, this approach increases the difficulties that will prevent an analysis to formulate a robust forecasting model. To further analyze the problem of working with raw data, we explore the potential of utilizing the clustering tool for the purpose of increasing the forecasting accuracy of a feed-forward neural network trained by the Levenberg–Marquard algorithm [10].

The developed models are applied on the Southern (SUD) Italy electricity market [11–13]. The available data set covers the period between 01/02/2012 and 30/04/2015. Among them, the period between 01/02/2012 and 31/12/2014 is used as the training set and the rest is used as test set. The role of the training set is the determination of the optimal ANN configuration, i.e. the optimal selection of the type of neurons activation function, number of hidden layers, number of neurons in the hidden layer(s) and maximum number of training epochs. One training epoch corresponds to one forward pass and one backward pass of all the training examples. The test set is used for the models comparison.

1.3. Literature survey and contributions

From a market's participant perspective, the estimation of the MCP in short-term horizon aids on the adoption of a proper strategy in wholesale market exchanges, i.e. the establishment of bilateral contracts or the generation units scheduling. The importance of the estimation of MCP is evident in profit maximization problems [14,15]. The MCP is treated as a stochastic variable and a set of scenarios are constructed to estimate its future variation.

Reviews of the state-of-the-art on the existing techniques on price forecasting can be found in [4,16]. A literature review including the amount of the recently published papers can be found in [17]. These studies attempt to reveal both the similarities and the differences between current techniques. According to [4] the existing approaches can be distinguished in three major categories: Game theory models, simulation models and time series models. The latter can be further categorized to stochastic models, artificial intelligence models and regression models. Time series models, such as ARIMA and GARCH, are a popular approach; they can serve as benchmark models for further model comparison and can be combined with other models leading to the formation of hybrid models. Their widespread usage is due to the fact that the mathematical formulation that refers to is comprehensive. Time series models require historical values of the quantity under prediction and they assume that the quantity evolution follows a specific pattern. The prediction is accomplished through the pattern's extension to a pre-defined future time period [18,19]. A comparison of various time series models like AR, ARMA and ARIMA can be found in [20]. Various sub-models are built (i.e. AR(1) with jumps, AR(1) in logs with jumps, AR(1) with time variant mean and others) and tested on the LPX market. Another comparative simulation study is conducted in [21] between a k-factor GIGARCH process and a SARIMA-GARCH model. The test study is applied on one month data of EEX market and the models include only lagged price values.

The time series models category includes the ANN based models. Representative bibliography is registered in Table 1. The ANNs that have been proposed in the literature are the following: Feed-Forward Neural Networks (FFNNs) and specifically Multi-Layered Perceptrons (MLPs) and Radial Basis Function Networks (RBFNs), Support Vector Machines (SVMs), Fuzzy Neural Networks (FNNs), Recurrent Neural Networks (RNNs), Probabilistic Neural Networks (PNN) and Self-Organizing Maps (SOMs). MLPs are the most commonly used networks due to their simplicity, training speed and reported effectiveness. MLPs are implemented as the sole forecaster in [9,10,19,22–67] or in combined models utilizing another time series models [68–71]. Other approaches propose of using the same ANN for both load and price forecasting or combining the ANN with a data mining technique to select appropriate days for training [72–75]. While the majority of the studies refer to Day-ahead predictions, the MLPis also utilized in hour-ahead time framework [75,76]. The role of the MLP in a combined model is to enhance the prediction produced by the conventional time series model (for example an ARIMA). Apart from price values, the ARIMA can be used to predict other variables like stored energy in reservoirs, inflow energy in reservoirs, total hydro generation and system load [77]. The predictions are used as inputs in the ANN system. Also, time series approaches combining ARIMA and GARCH have been proposed in the literature [78]. The ANN is compared with different AR and ARIMA models in [79].

RBFNs is the other type of FFNNs and are utilized in the works of [80–83]. An RBF network involves a hidden and an output layer. The RBF holds the role of the activation function of the hidden layer. This type of ANNs is able to simulate complex relationships underlying the data and can adapt fast to possible changes of these relationships.

SVMs provide a non-linear mapping of the original data into high dimensional space [84–90]. The boundaries of the new space are demarcated using linear function. SVMs provide a global solution to a problem unlike MLPs who can operate within local minima of their objective function. This fact has been also recognized in many researches in the load forecasting area [2]. A comparison between SVM and ANN in NYISO is the focus of [91]. Also, the SVM is used for estimating the prediction intervals which quantify the uncertainty related to forecasts by estimating the ranges of the targeted quantities [92].

Another approach in price forecasting is the synergetic operation of Fuzzy Logic (FL) and ANNs. This part of the literature can be further classified into two categories: Studies that utilize FL and ANN in the same system (i.e. neuro-fuzzy systems like ANFIS) and the studies where FL and ANN are separated forecasters that are combined into a two-part forecaster [83,93–105]. In the latter, the forecasting is accomplished via a linguistic description of the relationships between the input data. On the other hand, usually in ANFIS implementations the inputs involve only historical price values that are close to the days of the prices that are to be predicted. In [106], the forecasts of three models, i.e. ANN, ANFIS and ARIMA are fused by a Kalman filter to provide the forecast for the Spanish electricity system. The same models are also combined in [107]. The modified ordered weighted average algorithm is used to fuse the predictions produced by each model and the algorithms are tested on four representative weeks in the system of Spain. In [108], a hybrid system is used to predict prices in the Spanish market. The inputs are selected via the mutual information method and fed into the ANFIS. The optimal parameters of ANIS are determined by the evolutionary particle swarm optimization method.

RNNs rely their operation on the memory of the previous stages of the network, a concept that leads to a dynamic operation [98,109,110]. The output layer provides a feedback to the units of the hidden layer that process information from external signals and for a set of context units; the latter have fixed weights and serve as the memory units.

Electricity spike prices are abnormal values within price time series that can cause considerable economical effects on market participants. There are some researches that have focused on this problem proposing a series of techniques [111–120]. The most common approach is the PNNs implementation. PNNs are rapidly trained feed forward networks with a single output. Through a preliminary analysis, the threshold value of price that is recognized as spike is extracted. The PNN is used for the occurrence prediction (i.e. spike or not spike) or combined with a ANN, apart from the occurrence, the spike value can also be predicted.

SOMs are unsupervised machine learning neural networks that have found applications in various clustering problems. In load forecasting applications, SOMs are combined with ANN (MLPs or SVMs) or another time series model to form hybrid systems [121,122]. The SOM is occupied for the clustering of the training set in homogenous clusters. The training patterns within the same cluster display higher similarity compared with the rest. Then a number of forecasting models equals to the number of clusters is employed. This type of hybrid systems is not very common in price forecasting [60,99,123–125]. In [123] the SOM is combined with a SVM for the sort-term price forecasting in New England market. Again the hybrid system in [99] is composed by SOM and SVM examining the PJM market. The particle swarm optimization (PSO) algorithm is used for the optimal selection of SVM parameters via trial and error. To predict the next hour price h, the SVM receives as inputs the loads of hours h, h-1, h-2, h-3, the price of same hour one day before, the minimal and the maximum price of one day before. Also there is a distinction between weekends (entering "0" as input) and working days (entering "1" as input). In [124,125] the SOM is combined with ANN for load and price forecasting in NYISO, Australia and Spain markets. For each quantity, no exogenous parameters are used for both clustering and forecasting. The SOM is applied to cluster the training data and extract the cluster labels. The ANN is used to predict next day's cluster label. The predicted cluster label is associated with the topological coordinates of the map and the time series of load and price are obtained. Authors of [126] involve a pattern sequence similarity approach. The clustering is used both for grouping the data and forecasting. No neural network is considered. The K-means algorithm is employed for the purpose of grouping and labeling the samples from the dataset. Next, the pattern sequence prior to the day to be predicted is extracted. This sequence is searched in the historical data and the forecasted values are calculated by averaging all the samples immediately after the matched sequence. In [60] the Fuzzy C-Means (FCM) algorithm is used to cluster the day period into three zones (peak, normal and off-peak hours). There also a distinction of these zones between working days, Saturdays and Sundays. The clustering is combined with a time series model in [127], i.e. the clustering is not combined with ANN. In [128] the FCM clusters into three clusters and for each one a RNN is trained and tested in PJM market. Finally, other ANNs that have been proposed in the literature are the fuzzy Adaptive Resonance Theory (ART), Bayesian and Extreme Learning Machine (ELM) [104,129,130]. The ELM is combined with the maximum likelihood method for predicting the prices of ANEM [131]. A comparison takes place between the proposed model and three other models, i.e. the persistence approach, bootstrap ANN and ELM-Bootstrap approach. Also, another type of ANN used in price forecasting is the ELMAN neural network. According to [132], it outperforms the FFNN and the RBN when tested on NSW, Australia. The authors test various configurations differ in terms of number of inputs and number of neurons in the hidden layer.

The ANNs are suitable in problems that a general model or function expressing the relationship between the data and the parameters affecting them is absent or not straight-forward. Since the selection of input data is crucial to the model' s successful

Table 1
Selective bibliography focusing on price forecasting via ANNs application.

Reference	Model	Market	Characteristics
[22]	MLP	PJM; Spain	4 test weeks and the last week of all months of 1 year for PJM 4 test weeks for Spain
[8]	MLP	Ontario	Application of the discrete wavelet transform 6 test weeks
[23]	MLP	New York; Spain	Only historical values of price and load are used 2 test months for NYISO; 4 test weeks for Spain
[24]	MLP	PJM; Spain; Ontario	Only historical values of price and load are used 4 test weeks for PJM; 4 test weeks for Spain; 3 test months for Ontario
[71]	ARIMA/MLP	Ontario; New England; Italy	Only historical values of price and load are used 6 test weeks for Ontario; 11 test months for New England; 11 test months for Italy
[26]	MLP	PJM	Application of the discrete wavelet transform 4 test weeks
[27]	MLP	PJM; Spain	Only historical values of price and load are used 19 test days for PJM; 4 test weeks for Spain
[28]	MLP	NEMMCO; Spain	Only historical values of price and load are used 4 test weeks for NEMMCO; 4 test weeks for Spain
[68]	ARIMA/MLP	PJM; Spain	Only historical values of price and load are used 19 test days for PJM; 4 test weeks for Spain
[111]	PNN	PJM; QLD	Price spike forecasting Application of the discrete wavelet transform 1 test year for PJM; 2 test months for QLD
[112]	PNN	PJM	Price spike forecasting 3 test days
[29]	MLP	New York; Spain	Only historical values of price are used 4 test weeks for New York; 4 test weeks for Spain
[30]	MLP	New York; Spain	Only historical values of price are used 4 test weeks for New York; 4 test weeks for Spain
[32]	MLP	Iran	Only historical values of price and load are used 219 test days
[34]	MLP	Spain; California	Only historical values of price are used 4 test weeks for Spain; 1 test week for California
[35]	MLP	Spain	Only historical values of price are used 4 test weeks
[36]	MLP	Spain	Only historical values of price are used Application of the discrete wavelet transform 4 test weeks
[94]	ANFIS	Spain	Only historical values of price are used Application of the discrete wavelet transform 4 test weeks for Spain
[36]	MLP	Spain; California	Application of the discrete wavelet transform Only historical values of price are used 273 test days for Spain; 1 test year for California
[96]	FL/MLP	Australian New-South Wales	The clustering method is utilized Only historical values of price are used
[80]	RBF	Spain	Only historical values of price are used 24 test days
[18]	MLP	PJM	Only historical values of price and load are used 4 test weeks
[38]	MLP	ANEM	Only historical values of price and load are used
[9]	MLP	*Not mentioned*	2 test years
[98]	FL/RNN	PJM	1 test week
[41]	MLP	PJM	Only historical values of price and load are used 4 test weeks
[125]	SOM/MLP	Spain; New York; ANEM	The clustering method is utilized Only historical values of price and load are used 4 test weeks for Spain; 1 test year for NYISO; 1 test year for ANEM
[42]	MLP	New England	Only historical values of price and load are used 75 test days
[43]	MLP	PJM; Spain	Only historical values of price and load are used 4 test weeks for PJM; 4 test weeks for Spain
[70]	GARCH/MLP	Victoria (Australia); New York	Only historical values of price are used 3 test months for Victoria; 3 test months for New York
[81]	PNN and MLP	PJM	17 test days
[113]	RBF	PJM	30 test days
[114]	MLP	ANEM	Application of the discrete wavelet transform 3 test months
[45]	MLP	Victoria (Australia)	1 test month
[48]	MLP	NYISO; ANEM; Spain	21 test months for NYISO; 21 test month for ANEM; 9 test months for Spain
[99]	SOM/SVM	PJM	The clustering method is utilized

(*continued on next page*)

Table 1 (*continued*)

Reference	Model	Market	Characteristics
			1 test month
[108]	ANFIS	PJM; Spain	Application of the discrete wavelet transform Only historical values of price are used 2 test weeks for PJM; 4 test weeks for Spain
[49]	MLP	EEX	Only historical values of price are used 3 test months
[52]	MLP	Spain	Only historical values of price and load are used 12 test days
[101]	ANFIS	Spain	Only historical values of price are used 4 test weeks
[103]	ANFIS	Spain	Only historical values of price are used 4 test weeks
[82]	ARIMA/RBF	Spain	Application of the discrete wavelet transform Only historical values of price are used 4 test weeks
[86]	SVM	Iran; Ontario; Spain	Application of the discrete wavelet transform Only historical values of price and load are used 4 test weeks for Iran; 6 test weeks for Ontario; 4 test weeks for Spain
[87]	SVM	New York; PJM; New South Wales	Application of the discrete wavelet transform Only historical values of price and load are used 4 test weeks for New York; 1 test week for PJM; 4 test weeks for New South Wales
[130]	ELM	Ontario; PJM; New York; Italy	Application of the discrete wavelet transform Only historical values of price are used 6 test weeks for Ontario; 4 test weeks for PJM; 11 test months for New York; 11 test months for Italy
[7]	MLP	*Not mentioned*	6 test days
[59]	MLP	New England; Alberta; Spain	4 test weeks for New England; 4 test weeks for Alberta; 4 test weeks for Spain
[60]	FCM/MLP	PJM	The clustering method is utilized Only historical values of price and load are used 4 test weeks
[115]	MLP	Nord Pool	Price spike forecasting 1 test year
[116]	MLP	Victoria (Australia)	29 test days
[62]	MLP	England and Wales Power Pool	3 test days
[64]	MLP	California	1 test week
[66]	ARIMA/SVM	New South Wales	Application of the discrete wavelet transform 4 test weeks
[90]	SVM	PJM; Spain	Application of the discrete wavelet transform 4 test weeks and 3 test days for PJM; 4 test weeks for Spain
[89]	ARIMA/SVM	New Wales	Application of the discrete wavelet transform Only historical values of price are used 4 test weeks

operation, various input selection techniques has been presented, for example the mutual information and correlation coefficient [43,111]. Historical price and load values are sorted based on the highest similarity in a list with the current hour's MCP and the top listed values are fed into the ANN. Another aspect in ANN based price forecasting is the importance of network training. Some researches propose nature inspired algorithms like genetic algorithms, invasive weed optimization, cuckoo search and particle swarm optimization as helping techniques for better neural network training [26,94,111,133–135]. A popular approach is the wavelet decomposition of the MCP series. For each wavelet-domain indicator a dedicated ANN is applied. The predictions on the wavelet domain of each forecaster are synthesized to obtain the MCP time series [8,26,36,37,82,87,89,90,94,100,114].

The majority of the papers focus on PJM and Spain markets. The test sets correspond to 4 representative weeks, 1 per season. These sets are used as unofficial benchmarks for the models comparison by many researchers. No universal competition in MCP forecasting providing test sets, guidelines and evaluation framework have been taken place, contrary to load forecasting [122]. This fact has been also discussed in [16].

Based on the above brief literature survey, it is obvious that the price forecasting problem has been tackled by many approaches involving sole and hybrid forecasting systems, various input selection techniques and algorithms for advancements in neural network training. The objectives of the present study, aiming at filling gaps and supplementing research in the literature, can be summarized in the following:

(a) This study is structured around the problem of working with no pre-processed data. The aim is to test the robustness of various MLP topologies in the case of data sample with missing entries and null values. According to the brief literature survey conducted on the previous, it is evident that most studies assess their respective models on four representative weeks of the year. To fully examine the performance of the proposed models, our test set involves a four month period covering working days, weekends and holidays. To limit the effect of null and abnormal values, authors of [130] discuss some modified version of Mean Absolute Percentage Error (MAPE). This scheme leads to smooth MAPE values. Contrary to this paper, in this study the evaluation framework involves the conventional definition of the MAPE, which is the most common indicator in load and price forecasting applications.

(b) Various types of inputs are investigated. Proportionally to other studies, we are also concerned of using only historical load and price values. The influence of MCP of other countries' markets that are interconnected with Italy is checked.

(c) According to [7] another classification of the literature can be held considering the type of the output of the ANN. The output can refer to next hour's price, the price of several hours ahead, the next day's peak price (spike), the next day's average on-peak price and next day's average price. For example, in next hour's (denoted with h) price forecasting, the ANN can receive as inputs the prices of previous hour and two hour before, h-1 and h-2, respectively. This is followed in the majority of the studies. Authors of [7] states that the next day's complete price profile is the less popular group of studies. Our study belongs to this category. The forecasting is focused in the next day's prices; this means that the prices of h-1 and h-2 cannot be used since they are unknown.

(d) Another objective is to analyze the potential of the clustering tool in the data sample under study. The clustering algorithm and the conventional forecaster are hybridized to combined models. Through the clustering algorithm, the historical MCP data are separated into well-separated and homogenous groups. Next, for each group a dedicated ANN is trained and applied. It should be noted that since the ANN is trained with a different sub-set of the historical data, the optimal parameters of each network obtained by the training phase may differ. While hybrid models have been successfully applied in load forecasting problems, so far their utilization in MCP is relatively limited; it is represented by only six studies [60,99,123–125,127].

To sum up, the paper proposes ANN and hybrid ANN models for the Day-ahead Market price forecasting, working with no pre-processed data, elaborating historical load and price data and analyzing the potential of the clustering algorithm, which separates historical data in well-separated and homogeneous groups. The proposed ANN models aim at providing evidence so as to be considered as useful and robust forecasting tools to the actual needs of market participants, including the traditional generation companies and self-producers, but also the retailers/suppliers and aggregators.

1.4. Structure of the study

The rest of this paper is organized as follows: The general price forecasting framework is presented in Section 2. The models description as included in this section. In Section 3 the models are compared using various indicators. The purpose is to examine the models within a general evaluation background in order to reach into safe conclusions regarding the potential and drawbacks of each model. The main findings of the present work are presented in Section 4.

2. Price forecasting framework

2.1. Data collection

The general framework followed in the study is graphically presented in Fig. 1. It is based on a feature selection stage, models training and test stages. Due to the absence of previous works on price forecasting in the SUD Italy, an investigation on the selection of the number and type of inputs of the ANN takes place. Each developed model refers to different types of inputs. It should be noted that the current data set present many spikes, null values and missing entries.

The price time series of the period referring to the training set is illustrated in Fig. 2. It can be observed that the prices vary greatly. The price time series has no constant mean and displays high volatility. It represents a non-stationary time series. The time series includes many zero values and also, many low values below 15 €/MW h. The peak prices are close to 210 €/MW h and there are many observations close to 160 €/MW h. The peak price of the test accounts for 136.13 €/MW h which is considerably lower than the training test peak.

Table 2 presents some statistical characteristics of the hourly prices of the training data set. The highest prices are observed at evening peak hours (18:00–20:00). The minimum prices in all hours are zeros while the rest have values very close to zero. According to mean and median prices, it is noticeable that prices are low in early morning hours and from 08:00 to 10:00 there is an continuous increment. The prices are low between 11:00 and 14:00 and afterwards prices increase until 22:00. Sixth column give indications about the variance. Most variance is met at 20:00. The coefficient of variation is an indication about the dispersion of values around the means. Low values correspond to low variability. The values vary between 0.22 and 0.46. This means that hourly values are characterized by high variability. A further exploration of the price time series is held though the plotting of extreme value distribution of the training data in Fig. 3. The red line corresponds to the peak of the distribution. The Figure shows the relationship between the minimum value (i.e. zero) and the rest in terms of the frequency of occurrences. Most prices lie within the [40,50] range.

2.2. Feature selection

According to [136] electricity price is a nonlinear function of many candidate inputs including its past values as well as past and forecast values of the exogenous variables such as load demand and available generation. To deal with this fact, various researches have proposed feature selection techniques. For example, the selection includes time domain and wavelet components of the price time series. In this study, the optimal input selection is based on trial and error. This means that a series of models are examined in order to reach safe conclusions about the forecasting approach that needs to be followed for the current data set. In the majority of forecasting problems historical values of the parameter under study are always input candidates. To examine the price time series periodicity, the Pearson correlation coefficient is used to measure the degree of dependence between current values and values up to 1 week before [137]. The correlation coefficient curve is illustrated in Fig. 4. The data used for the correlation analysis refer to the average values of hours $h = 1, 2, \ldots, 168$ of the training set. It is shown that current hour price shows high correlation with those of hour h-24 and h-168, a fact that indicates daily and weekly periodicity. Prices of the days prior to the target day and the same day 1 week before are selected as inputs. Also, short-term periodicity is evident (hours h-2, h-3) but the corresponding prices are not selected as inputs since, as previously described, in the next day forecasting problem formulation are not known. Recall that in this paper the forecasting is employed once in a day to predict the whole price profile of the next day.

In MCP analysis works, the most influential external variable is considered to be the load. In the present paper, historical load values are also considered as candidate inputs. Additionally, we assume that next day's forecasted load is available. Fig. 5 shows the correlation between price and load of the training set period.

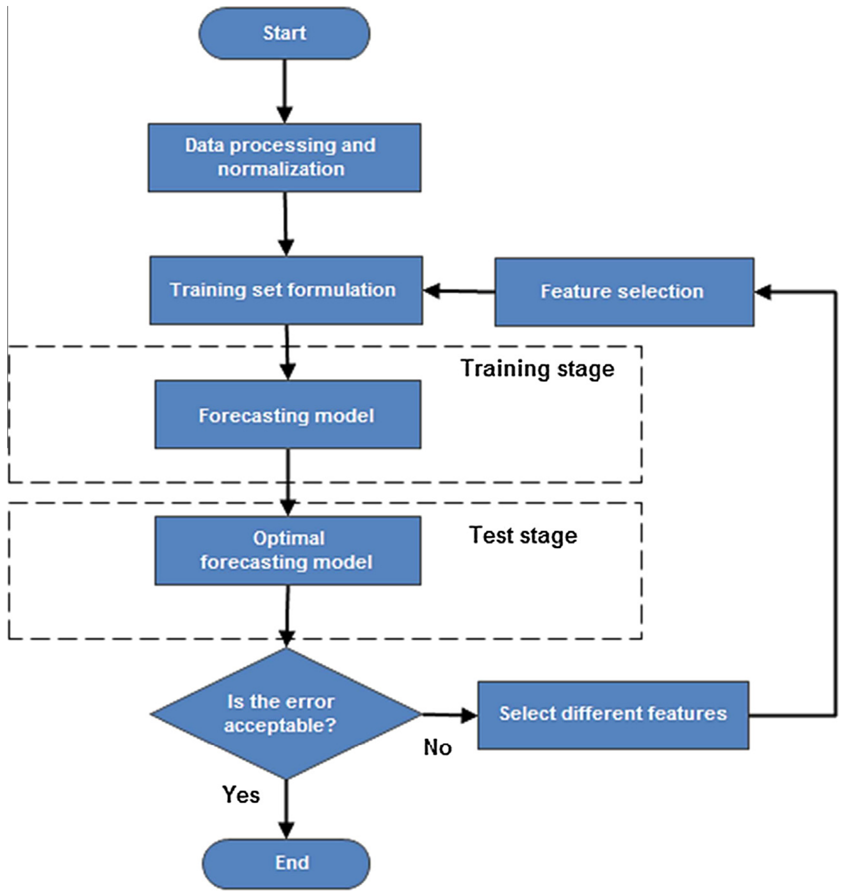

Fig. 1. Flow-chart of the proposed methodology.

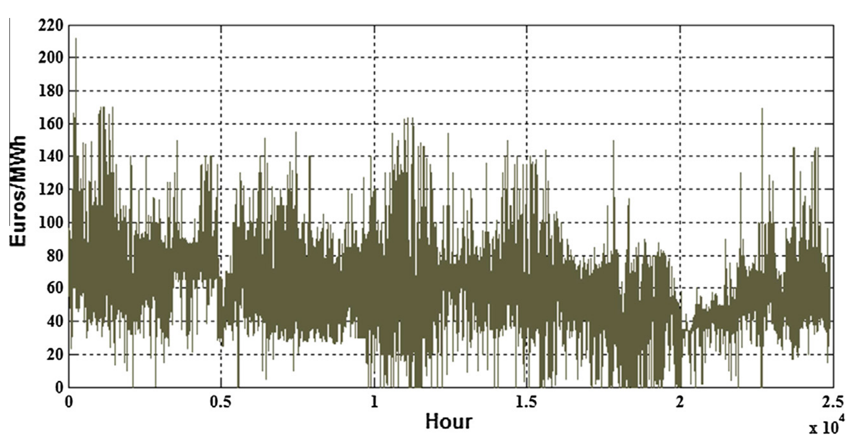

Fig. 2. Electricity price times series corresponding to the period of the training set.

There is an analogy between price and load values. While the load level rises, a constant increase of price is observed. To keep a consistent framework on the inputs relationships, the hourly loads of the days prior to the target day and the same day 1 week before are used. With respect to the MCP dependence on the temperature variations, Fig. 6 displays the correlation between these two variables. It is observed that no strong correlation exists. This means that no temperature or other climatic parameters are used as inputs. Finally, day type identification and holiday codes are used to support the robustness of the ANN learning capability.

2.3. Models description

2.3.1. Neural network mathematical background

The basic architecture of the back-propagation ANN is depicted in Fig. 7. It is composed by the input layer l_1 with neurons a_1, a_2 and a_3, hidden layer l_2 with neurons a_4, a_5 and a_6 and output layer l_3 with neurons a_7 and a_8. Each output signal in layer l_3 is compared against the corresponding target output signal in order to calculate the error information. Let t_k and y_k be the target and actual targets signal of the kth neuron, respectively. The error information of the kth neuron δ_k is given by the following expression:

Table 2
Statistical indices of the training data set.

Hour	Maximum	Minimum	Mean	Median	Standard deviation	Coefficient of variation
1:00	94.98	5	54.86	54.13	14.28	0.26
2:00	88.10	0	47.96	46.33	14.27	0.29
3:00	88.06	0	43.75	41.37	14.39	0.32
4:00	87	0	41.48	39.19	14.23	0.34
5:00	85.65	0	41.49	39.55	14.02	0.33
6:00	88.03	0	45.16	43.91	13.55	0.30
7:00	85.09	0	52.90	55.76	14.66	0.27
8:00	119.38	0	60.44	63	16.72	0.27
9:00	149.24	0	63.98	65.56	19.08	0.29
10:00	140	0	61.13	63.46	18.69	0.30
11:00	140	0	54.90	58.86	18.89	0.34
12:00	135	0	50.46	54.67	19.60	0.38
13:00	105.03	0	45.83	48.05	19.58	0.42
14:00	98.03	0	42.95	43.33	20.16	0.46
15:00	115	0	45.64	48	20.72	0.45
16:00	140	0	50.62	55	20.13	0.39
17:00	140	0	57.95	62	19.67	0.33
18:00	186.14	0	67.33	67.91	23.44	0.34
19:00	212	0.10	73.38	72.91	22.83	0.31
20:00	191.04	0.10	79.94	78	24.31	0.30
21:00	170	0.10	79.68	76.38	23.60	0.29
22:00	163.05	0.10	71.94	69	20.15	0.28
23:00	140.12	0.10	64.24	64.10	15.04	0.23
24:00	91.43	0.10	58.38	59.50	13.33	0.22

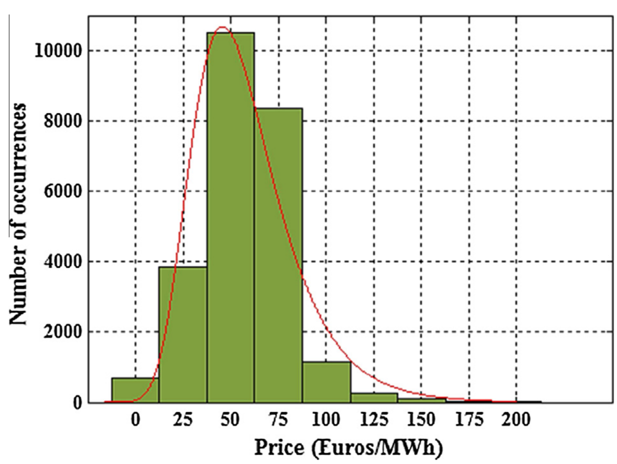

Fig. 3. Extreme value distribution of the training set.

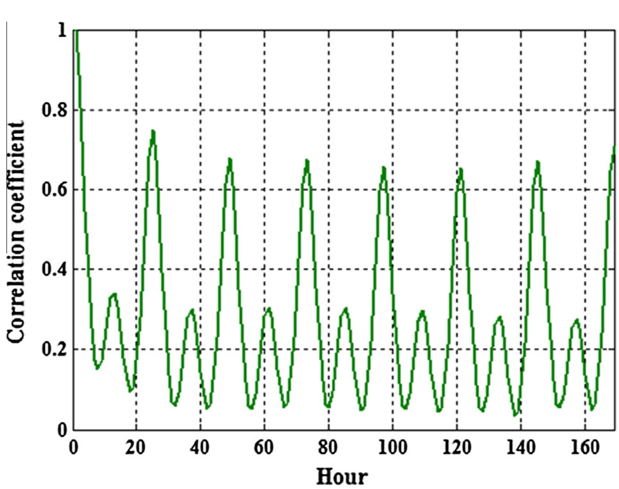

Fig. 4. Correlation coefficient between the current MCP up to the value of hour h-168.

$$\delta_k = (t_k - y_k) f'(y_{in_k}) \tag{1}$$

where y_{in_k} is the activation value of the kth output neuron and f' is the activation function. Next, δ_k is used to update the neuron of the output layer and propagates back to the hidden layer. The weight correction information is expressed as:

$$\Delta w_{jk} = \mu \delta_k z_j \tag{2}$$

where w_{jk} is the weight connecting jth hidden neuron and the kth output neuron, μ is the learning rate and z_j is the output signal of the jth neuron. The weight correction information used to update the neuron is given by:

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk} \tag{3}$$

where w_{jk}(new) and w_{jk}(old) denote the updated weight connecting jth hidden neuron and the kth output neuron, the initial weight connecting the jth hidden neuron and the kth output neuron, respectively.

For the neurons in layer l_2, the error information is obtained from the connected neurons in the input layer. The incoming error information is given by:

$$\delta_{in_j} = \sum_{k=1}^{m} \delta_k w_{jk} \tag{4}$$

where δ_{in_j} is the denotes the incoming error information of the jth neuron. The error information of the jth neuron δ_j is given by:

$$\delta_j = \delta_{in_j} f'(z_{in_j}) \tag{5}$$

The weight correction information and the weight updating process of the neurons of the hidden are the same as the output layer. To provide a more detailed description, the weight of a neuron in the output layer is updated via the following formula:

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \mu\left(\left(k - f\left(\sum_{i=1}^{\tau} w_i x_i\right)\right) f'\left(\sum_{i=1}^{\tau} w_i x_i\right)\right)(x_j) \tag{6}$$

where τ is the number of patterns, w_{jk} is the weight connecting the neuron and the jth neuron, k is the target output signal, f is the activation function, f' is the derivative activation function and x_k is the kth input. If the neuron is in the hidden layer, the weight of the neuron is given by:

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \mu\left(\delta_j f'\left(\sum_{i=1}^{\tau} w_i x_i\right)\right)(x_j) \tag{7}$$

where δ_j is the sum of error information gathered from neurons in higher layers which uses information from the jth neuron [138]. Different types of activation functions can be used for the specific ANN [139].

2.3.2. Model A

The objective of the present model is to examine a single configuration that does not include exogenous variables. This approach is followed in some studies, for instance in [34,35,101]. The model refers to feed-forward ANN (MLP) trained by the Levenberg–Marquard algorithm. The input vectors are present sequentially to the network. Specifically, column vectors are used. The output layer has one neuron referring to the price of hour h. This means that the forecasting is repeated 24 times, one for each hour of the next day. There is no need to set the forecast in every hour; it can be held any time in the previous day. The reason for this approach is the size of the network. According to the experimental results, the Levenberg–Marquard algorithm becomes suitable for small sized networks, i.e. small number of input and output

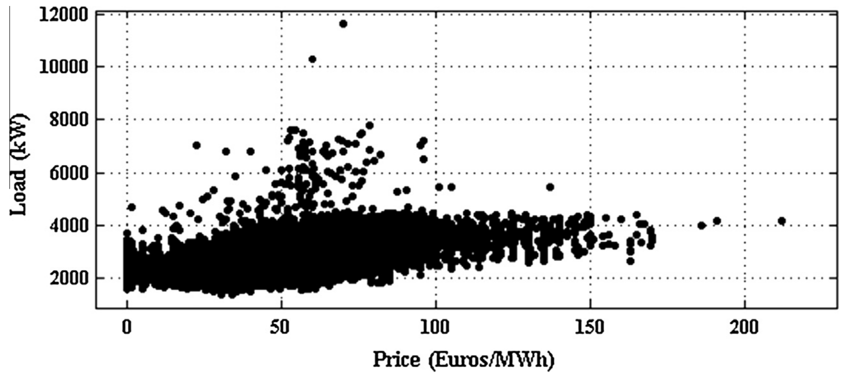

Fig. 5. Correlation between load and MCP of the training set.

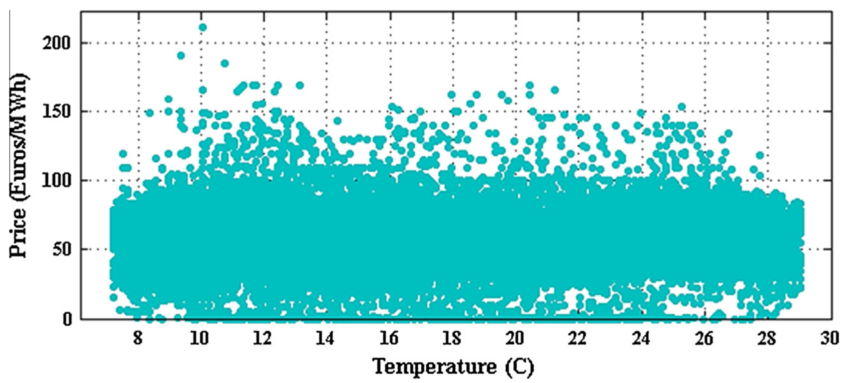

Fig. 6. Correlation between MCP and temperature.

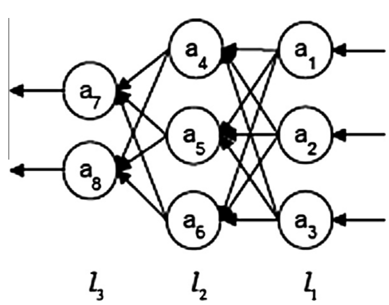

Fig. 7. Back-propagation neural network structure [138].

neurons. Not only the training performance is high, but the required training time is relatively small time. This allows us to set more experiments and evaluate different ANN topologies. Apart from price values, distinction indicators of the target day are fed into the neural network (inputs 1–3). Let *d*, *d*-1 and *d*-7 be the indicators that refer to the target day, the day prior to it and the same day 1 week before. Model A receives the following inputs:

- Input 1: Day type indicator, i.e. "1" for Sunday, "2" for Monday and so on.
- Input 2: Hour indicator, i.e. "1", "2",…,"24".
- Input 3: Holiday indicator, i.e. "1" for holidays and "0" for working days and weekends.
- Input 4: Hourly price of day *d*-1 (i.e. price *h*-24), $P(d\text{-}1, h)$.
- Input 5: Hourly price of day *d*-7 (i.e. price *h*-168), $P(d\text{-}7, h)$.

Model A will be used as a benchmark for the models comparative analysis.

2.3.3. Model B

Model B is a variation of the benchmark model. The historical prices up to 1 week before are sorted based on their highest correlation with the target price. Then the 10 top listed values are selected. These correspond to hours *h*-24, *h*-25, *h*-47, *h*-48, *h*-72, *h*-96, *h*-120, *h*-144, *h*-167 και *h*-168. The corresponding prices are used as inputs. Model B has the following inputs:

- Input 1: Day type indicator, i.e. "1" for Sunday, "2" for Monday and so on.
- Input 2: Hour indicator, i.e. "1", "2",…,"24".
- Input 3: Holiday indicator, i.e. "1" for holidays and "0" for working days and weekends.
- Input 4: Price of hour *h*-24.
- Input 5: Price of hour *h*-25.
- Input 6: Price of hour *h*-47.
- Input 7: Price of hour *h*-48.
- Input 8: Price of hour *h*-72.
- Input 9: Price of hour *h*-96.
- Input 10: Price of hour *h*-120.
- Input 11: Price of hour *h*-144.
- Input 12: Price of hour *h*-167.
- Input 13: Price of hour *h*-168.

The operation of Model B aims is to build a more representative function between inputs and outputs that may lead to better ANN performance.

2.3.4. Model C

Model C takes into account exogenous variables such as load, renewable energy sources (RES) generation and natural gas prices. The RES generation refers to the sum of wind turbines and photovoltaics generations. The available RES capacity is a significant factor in the setting of the prices in electricity wholesale markets. Our purpose is to examine if the information about the RES affects the MCP daily patterns. Furthermore, natural gas prices determine, in some degree, the operation scheduling of gas fired and combined cycle plants. Load data are entered as 4 different values: the load of hour *h* obtained by a prediction, loads of hours *h*-24 and *h*-168 and previous day average load. The latter is considered in order to reflect the daily variation. We also consider the previous day average price. The inputs are described below:

- Input 1: Day type indicator, i.e. "1" for Sunday, "2" for Monday and so on.
- Input 2: Hour indicator, i.e. "1", "2", ..., "24".
- Input 3: Holiday indicator, i.e. "1" for holidays and "0" for working days and weekends.
- Input 4: Predicted hourly load of target day (day *d*), *L*(*d*, *h*).
- Input 5: Hourly load of day *d*-1, *L*(*d*-1, *h*).
- Input 6: Hourly load of day *d*-7, *L*(*d*-7, *h*).
- Input 7: Previous day average load, *L*avg(*d*-1).
- Input 8: Hourly price of day *d*-1, *P*(*d*-1, *h*).
- Input 9: Hourly price of day *d*-7, *P*(*d*-7, *h*).
- Input 10: Previous price average load, *P*avg(*d*-1).
- Input 11: Predicted hourly RES generation of target day (day *d*), *GRES*(*d*, *h*).
- Input 12: Natural gas price of target day (day *d*), *P*NG(*d*).
- Input 13: Natural gas price of previous day (day *d*-1), *P*NG(*d*-1).

For the natural gas prices, only 1 value per day is available in the SUD Italian market. Inputs 12 and 13 can provide indirect information about the gas and combined cycle plants daily scheduling.

2.3.5. Model D

This configuration examines the impact of SMP of the main countries that Italy imports electricity. Specially, next day price forecasts of Greece, France, Germany and Switzerland are available and are led to the input neurons. Model D has the 13 inputs of Model C plus 4 inputs corresponding to the prices of the aforementioned countries. Note that this approach is first introduced in the price forecasting literature by the present study. Fig. 8 shows the MCP time series of the five countries. The most similar curve to Italy is the one of Greece. The latter displays the higher values among all. There are many instances where the Greek SMP reaches its higher value (i.e. 150 MW h), a value that is restricted by the Greek power market legislation. Germany's SMP presents many negative values.

2.3.6. Model E

Model E is structured as a cascaded neural network. It combines 2 networks in order to increase the overall performance. The 1st network has the same inputs with Model D. The output which is an initial estimation of the hourly price of day *d*, *P*est(*d*, *h*), is transferred into the 2nd network which also receives the inputs of Model D. The role of the 2nd network is to smooth the initial prediction. The number of its inputs is 14. The final prediction is obtained at the output neuron of the 2nd network. The topology of Model E is presented in Fig. 9.

2.3.7. Model F

Model E is a hybrid model that involves 2 general stages, namely the unsupervised and supervised machine learning stages. The aim is to improve the performance capability of the single ANN by generating smaller training set with more similar patterns via the application of clustering. The respective topology is shown in Fig. 10. The 1st stage includes the data preprocessing part and the application of clustering. The 2nd stage is the application of the neural network. The aim of the preprocessing stage is to formulate the data set into the suitable form for the clustering operation, i.e. scaling into the appropriate range. This is necessary in order to obtain meaningful and exploitable results from the clustering. The price magnitude is not taken into account in the clustering stage; the purpose is to track similarities and trends of the daily price curves, by comparing their shapes. Every daily price curve is represented with a *D*-dimensional vector $p_m = \{P_{m1}, \ldots, P_{mD}\}$, where $m = 1, \ldots, M$ is the number of the available vectors or patterns and $i = 1, \ldots, D$ is the dimension. The elements of the vectors refer to the hourly price, hence it is $D = 24$. The scaling is held via mean and standard deviation values. Every pattern is normalized according to the following formula:

$$x_m = \frac{p_m - \bar{p}_m}{std_m} \tag{8}$$

where x_m is the normalized vector, $\bar{p}_m$ and std_m are the mean value and the standard deviation of p_m, respectively. Eq. (8) normalizes the patterns in order to have zero mean and unity standard deviation. The set of the normalized patterns is denoted as $X = \{x_m, m = 1, \ldots, M\}$. First of all, the normalized price curves of days *d*-1 of the training set form a $M \times D$ matrix which is led into the clustering algorithm. In this study, the K-means algorithm is used. K-means is the most common algorithm in load profiling problems [140]. Through an iterative process, the algorithm tends to minimize the sum of squared errors, i.e. the sum of Euclidean distances between the patterns and the clusters centroids. The term centroid corresponds to the averaged pattern of all patterns in the same clusters. The algorithm terminates when there are no transpositions of patterns from cluster to cluster during the successive iterations. The patterns of each cluster present higher similarity than those of the other clusters. K-means partitions the training set matrix into *k* clusters. Each cluster (i.e. subset of the training set) is represented by its centroid. For each subset, a separate ANN (Model D) is employed. Hence, *k* ANNs are trained separately with the data of the corresponding clusters. During the test phase, the appropriate ANN should be chosen. The test matrix is built. The rows correspond to the daily price curves of the day prior to the target one. Each row of the test matrix is compared with *k* centroids via the Euclidean distance. The chosen ANN corresponds to the most similar centroid, i.e. the one that corresponds to the smallest Euclidean distance. This procedure is continued until all the rows are compared with *k* centroids and in every case, the most suitable ANN is selected. The final forecasting error is calculated by taking into account the errors generated from each ANN.

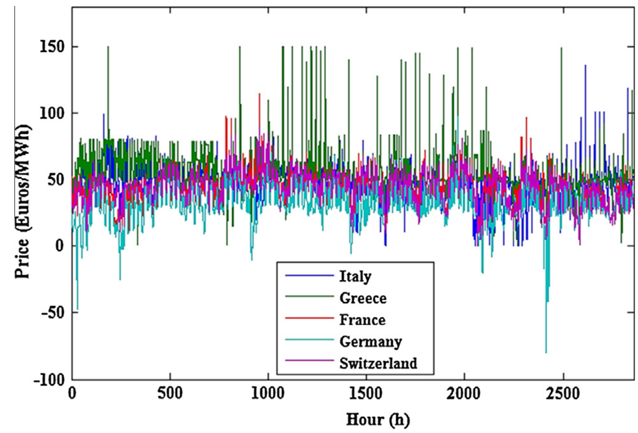

Fig. 8. Price time series of the test set of the 5 countries.

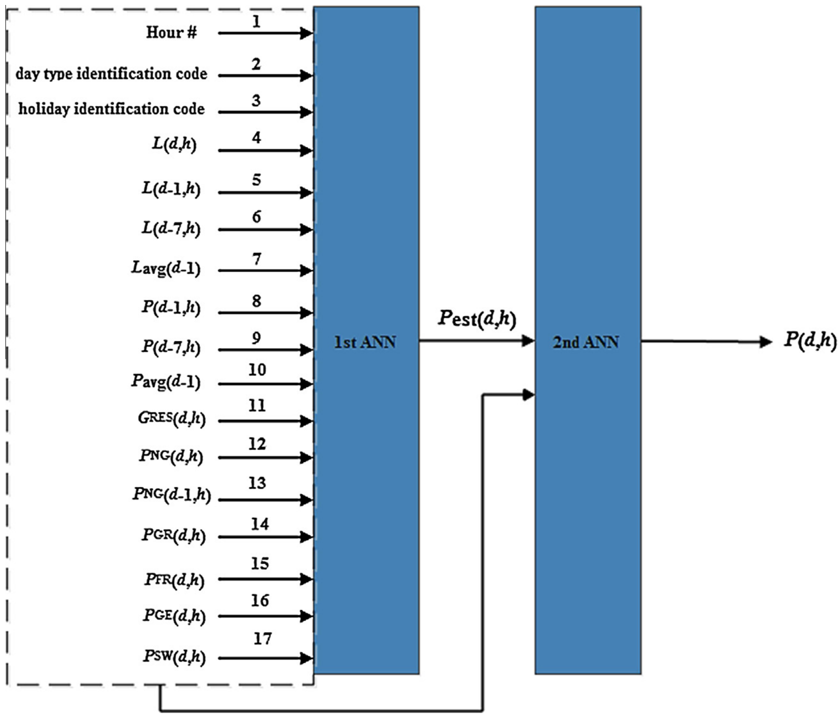

Fig. 9. The configuration of Model D.

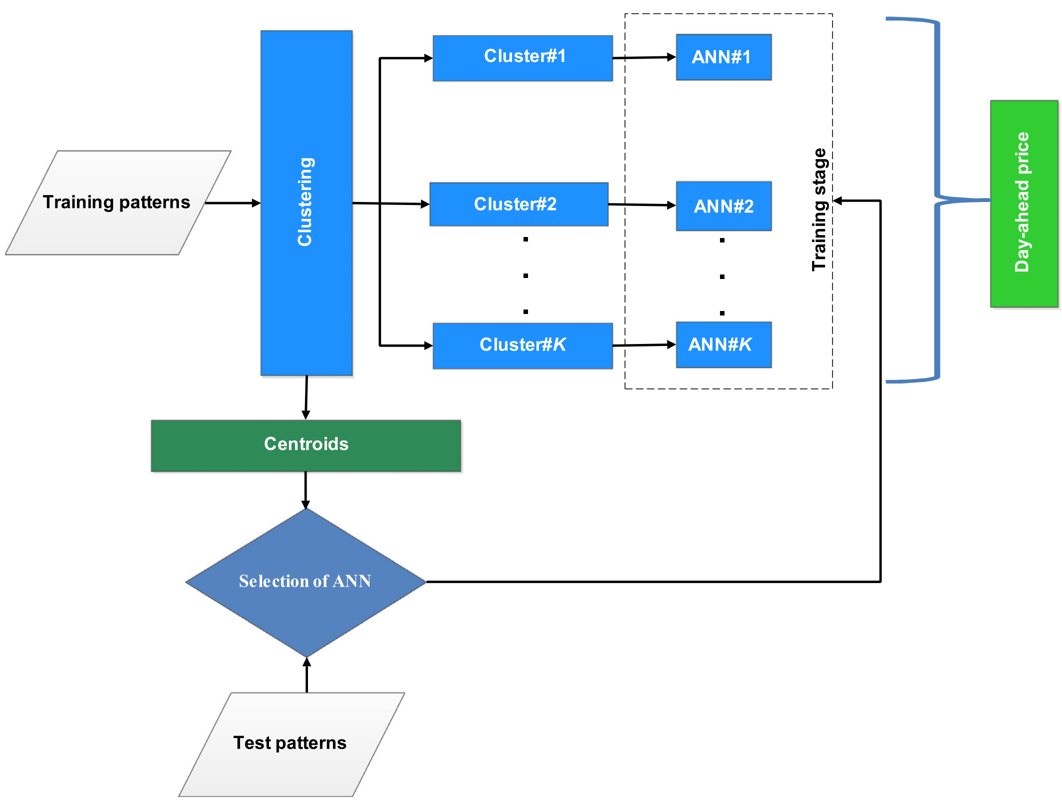

Fig. 10. The configuration of Model F.

3. Experimental results

3.1. General aspects

FFNNs have adjustable parameters providing a framework for experimentation on trial and error. A series of experiments are required to define the optimal ANN configuration of each model. The parameters that need to be properly determined are the number of hidden layers, the number of neurons of the hidden layer(s) and the type of transfer functions of the neurons of the hidden and the output layers. Also, the maximum number of training epochs of the Levenberg–Marquardt algorithm is a parameter under investigation. For Model F, the optimal number of clusters is an additional parameter. The models performances are evaluated with the Mean Absolute Percentage Error (MAPE), the most common indicator in load and price forecasting studies:

$$\mathrm{MAPE} = \frac{1}{M}\sum_{m=1}^{M}\frac{\left|P_m^{\mathrm{a}} - P_m^{\mathrm{f}}\right|}{P_m^{\mathrm{a}}} \times 100 \tag{9}$$

where P_m^{a} and P_m^{f} are the actual and predicted load of the m-th day, respectively. Furthermore, the Absolute Percentage Error (APE) provides information about the errors dispersion around zero. It is defined as follows:

$$\mathrm{APE} = \sum_{m=1}^{M}\frac{\left|P_m^{\mathrm{a}} - P_m^{\mathrm{f}}\right|}{P_m^{\mathrm{a}}} \times 100 \tag{10}$$

Also the differences between the forecasted and actual loads are examined simply defined as "Errors":

$$\mathrm{Errors} = \sum_{m=1}^{M} P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \tag{11}$$

To keep the comparisons accurate, the same optimal transfer function and number of hidden layers are considered. Additionally, we consider other performance metrics such as the Mean Absolute Error (MAE), Theil U Statistic 1 and U Statistic. The MAE is given by:

$$\mathrm{MAE} = \frac{1}{M}\sum_{m=1}^{M}\left|P_m^{\mathrm{a}} - P_m^{\mathrm{f}}\right| \tag{12}$$

Theil U Statistic 1 receives values between 0 (denoting good accuracy) and 1 (denoting poor accuracy). It is expressed by:

$$U_1 = \frac{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(P_m^{\mathrm{a}} - P_m^{\mathrm{f}}\right)^2}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(P_m^{\mathrm{a}}\right)^2} + \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(P_m^{\mathrm{f}}\right)^2}} \tag{13}$$

Theil U Statistic 2 can take values greater than 1 which denote poor forecasting accuracy. It is given by the following equation:

$$U_2 = \frac{\sqrt{\frac{1}{M}\sum_{m=1}^{M-1}\left(\frac{P_{m+1}^{\mathrm{a}} - P_{m+1}^{\mathrm{f}}}{P_m^{\mathrm{a}}}\right)^2}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\frac{P_{m+1}^{\mathrm{a}} - P_m^{\mathrm{a}}}{P_m^{\mathrm{a}}}\right)^2}} \tag{14}$$

3.2. Models comparison

According to the findings of the paper's analysis, a working practice is proposed to further exploit the developed models in real time applications. This practice is composed by the three general phases: "data-preprocessing", "model selection and training" and "expert knowledge incorporation". In the first phase, the characteristics of the data set should be examined by a preliminary study, for instance a statistical analysis. The abnormal data should be tracked and determine if these data outliers or metering failures. The latter should be excluded from the data set since they do not correspond to the accurate problem formulation and statement. The "model selection and training" phase seeks answers to the selection of the type of the model and its availability problems. The analysts should consult a related literature for the optimal selection of the model (i.e. time series model, ANN-based model, hybrid model or others). After the selection, the training of the model takes place. Usually, the rule for data splitting ratio is "80/20", 80% of the available data are used for models training and the rest 20% for models comparison and validation. In the case of ANN based models, the best performance is reached by the random feeding of the training patterns. This approach is the opposite of the sequential representation and leads to better generalization capabilities for the ANN. As for the availability of the models implementations, the analyst should aim on available and tested solutions. Time series models and ANN codes are available from commercial software packages and third-party distributors. These implementations provide a user friendly interface that is easily executed in common PC configurations. The third phase refers to the expert' s feedback. When expertise knowledge is available, it should be used accordingly. The analyst should consult the data provider (i.e. utility, retailer, transmission system operator) for the special characteristics of the data. Additionally, the findings of the analysis should be presented to the data provider seeking recommendations. If possible, every experimental result should be discussed with the data provider.

The test set is composed by 119 days; 1 day is missing from the data. The total number of hours is 2856. Note that no preprocessing of the data took place. The SUD Italy market functions as a pool, a fact that leads to many low and zero hourly price values. At least theoretically, the presence of low values affects negatively the forecasting performance. For example, if the network is fed with a typical pattern and a typical output is expected, if the real output corresponds to low value, MAPE indicator on the specific case will extremely high. In order to examine ANN based model in a realistic paradigm, all the available data have been included. Table 3 presents the low values distribution of the test set. Considering the 20 €/MW h as the threshold of low values, it is shown that low values represent almost 5% of the data set.

Each model is separately trained and applied using the training and test sets, respectively. Three different transfer functions are considered namely the logistic sigmoid, hyperbolic tangent sigmoid and linear. Various parameter combinations have been examined. Simulation results indicate that the optimal configuration refers to 1 hidden layer and tangent both in hidden and output layers. With respect to the influence of the maximum number of training epochs, the increase of them results in better training but also higher training times. The simulation results point out that 500 training epochs is a reasonable selection. It should be noted that in most cases, the maximum number of epochs is not reached. The ANN completes its operation in less epochs, i.e. when the improvement of the objective function threshold of the Levenberg–Marquardt algorithm between subsequent epochs is reached.

The optimal number of neurons in the hidden layer is an important parameter in network's performance. For each model, the number of neurons varies from 2 to 30 with an increasing step of 2. Larger networks do not necessarily increase efficiency. Fig. 11 presents the models comparison using the MAPE of the test set. No correlation is observed between the MAPE values and the increasing number of neurons. The MAPE curves do not present monotonic decreasing behavior. Model E corresponds to the higher prediction accuracy. This is justified by the examination the overall performance. In 3 cases, Model D outperforms Model E. The lowest MAPEs for Model A, Model B, Model C, Model D, Model E and Model F are 25.06%, 24.56%, 19.08%, 18.40%, 18.33% and 21.74%, respectively. The lowest MAPE are not met in the same number of neurons for all models. Model A and Model B have comparable performance. Model B presents some extreme error peaks. This indicates that, in the case of the current data set, using only

Table 3
Low values distribution.

Values (€/MW h)	Total	Percentage (%)
0	16	0.56
0–10	65	2.28
10–20	48	1.68
0–20	113	4.96

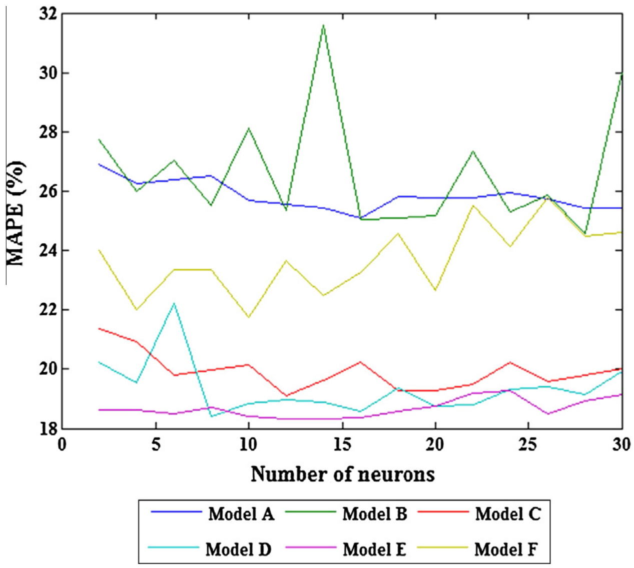

Fig. 11. Models comparison considering the MAPE indicator.

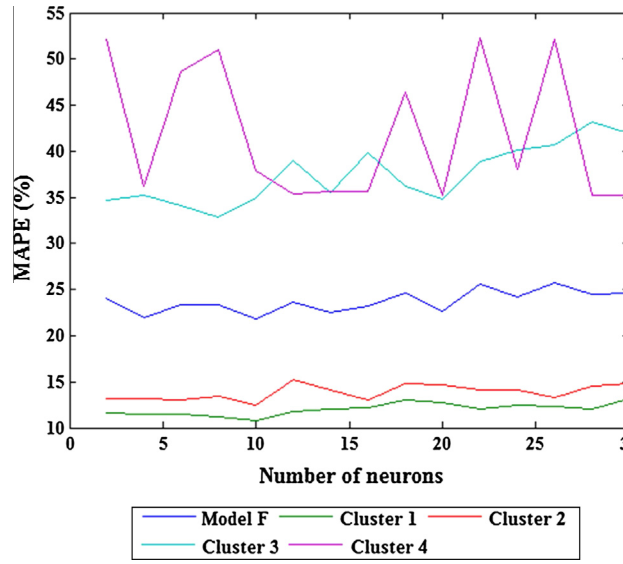

Fig. 12. MAPEs of Model F and for each cluster.

historical prices as inputs is not recommended. The models should include exogenous variables such as load and interconnected countries market prices. The consideration of the RES generation and natural gas prices are also important. Moreover, there is no correlation between the performance and the number of inputs. The most important role is the types of inputs that are used in every case.

The hybrid Model F leads to better results compared to Model A and Model B. The results show that the clustering stage does not improve the sole operation of Model E. Recall that Model F employs k Models E. This is justified by a further exploration of the clustering outcome. In Fig. 11, Model E is employed with $k = 4$. Table 4 shows the clusters compositions. All clusters include days from all months, January to April. Fig. 12 illustrates the MAPE curves of Model E and those of every cluster separately. MAPE of Model E is drawn as a combination of the 4 MAPEs. The lowest MAPE values of cluster 1, 2, 3 and 4 are 10.78%, 12.42%, 32.80% and 35.17%, respectively. Excluding MAPE of cluster 4, the rest curves have a more smooth shape compared with the curves of the rest of the models. MAPEs of clusters 1 and 2 are lower compared to the Model E, leading to the conclusion that the utilization of clustering improves the forecasting accuracy in some subsets of the main test set. The authors believe that this is also the case when using the test sets of Spain, PJM and other markets, according to the literature survey of the previous sections. The models of other studies are tested on 4 representative weeks of the year. Clusters 1 and 2 are composed mainly by working days. Cluster 4 which includes mainly weekends corresponds to higher errors.

Table 4
Clusters composition.

Cluster	Days
1	3rd and 4th Sundays of January, 3 working days of January, All Sundays of February, All Mondays and Tuesdays of March, 3 working days of April
2	2 Thursdays of January, 1st Friday of January, 4 days of February, 2 Saturdays and 1 Sunday of March, 3 Saturdays and 3 Sundays of April
3	2 Saturdays and 1 Sunday of January, 13 working days of January, 8 working days of February, 5 working days of March
4	1 Sunday of January, 1st and 2nd Sundays of February, 1 Saturday and 2 Sundays of March, 7 days of April

A further comparison of the models is held via the APE distributions. MAPE indicator receives one value for a specific prediction, for example for a given number of neurons in the hidden layer. It is essential to examine the error distribution over a focusing period. Using the APE indicator, the analysis can be scaled to hours. This concept strengthens the conclusions drawn from the comparison. Figs. 13 and 14 present the hourly and daily boxplots of the APEs of the models in logarithmic scale, respectively. The boxplots findings confirm the models hierarchy as concluded by the analysis of Fig. 11. The discrete outliers correspond to large error values, which are considered as indicators of the models failure. The outliers correspond to zero and low price values. The authors believe that forecasting failure is mainly due to the atypical values present in the test set. Boxplots are helpful to examining the forecasting efficiency per hour and day for all the test set. According to Fig. 13, the difficulties in the models operation are found at the periods 10:00–12:00 and 14:00–16:00 which correspond to morning and early afternoon peak hours. The lowest errors are met in 20:00–24:00 period. Early morning hours present relatively low errors but this is not the case for Model A and Model B. Model D and Model E limits the errors in the majority of the hours. Model C is more reliable forecaster compared to Model A and Model B. The boxplots of APEs per day are illustrated in Fig. 14. The indicators $1, 2, \ldots, 7$ correspond to $\text{Sunday}, \text{Monday}, \ldots, \text{Saturday}$, respectively. As it can be observed from the figures, weekends present the higher fluctuations. All models present poor results in Sundays. The lowest weekend errors are produced by Model E. The models present robust performance in working days and especially in Mondays and Thursdays. Wednesdays' prices are also easily to be predicted, excluding Model B, where it presents some limitations on the specific day. Among the working days, prices of Fridays are difficult to effectively be predicted in most cases. Fridays' errors peaks are higher than those of Saturdays and comparable with those of Sundays. Model A presents low efficiency in Fridays. This is also the case for Model B but in general terms appears a forecaster of higher capacity. Model F has satisfactory operation but shows some limitations in Sundays and Fridays. There are cases of high errors but are observed within a limited range.

Through the Errors indicator, it is feasible to examine the errors dispersion around zero. The Errors histograms are shown in Fig. 15 and present a visual assessment of the models performance. The

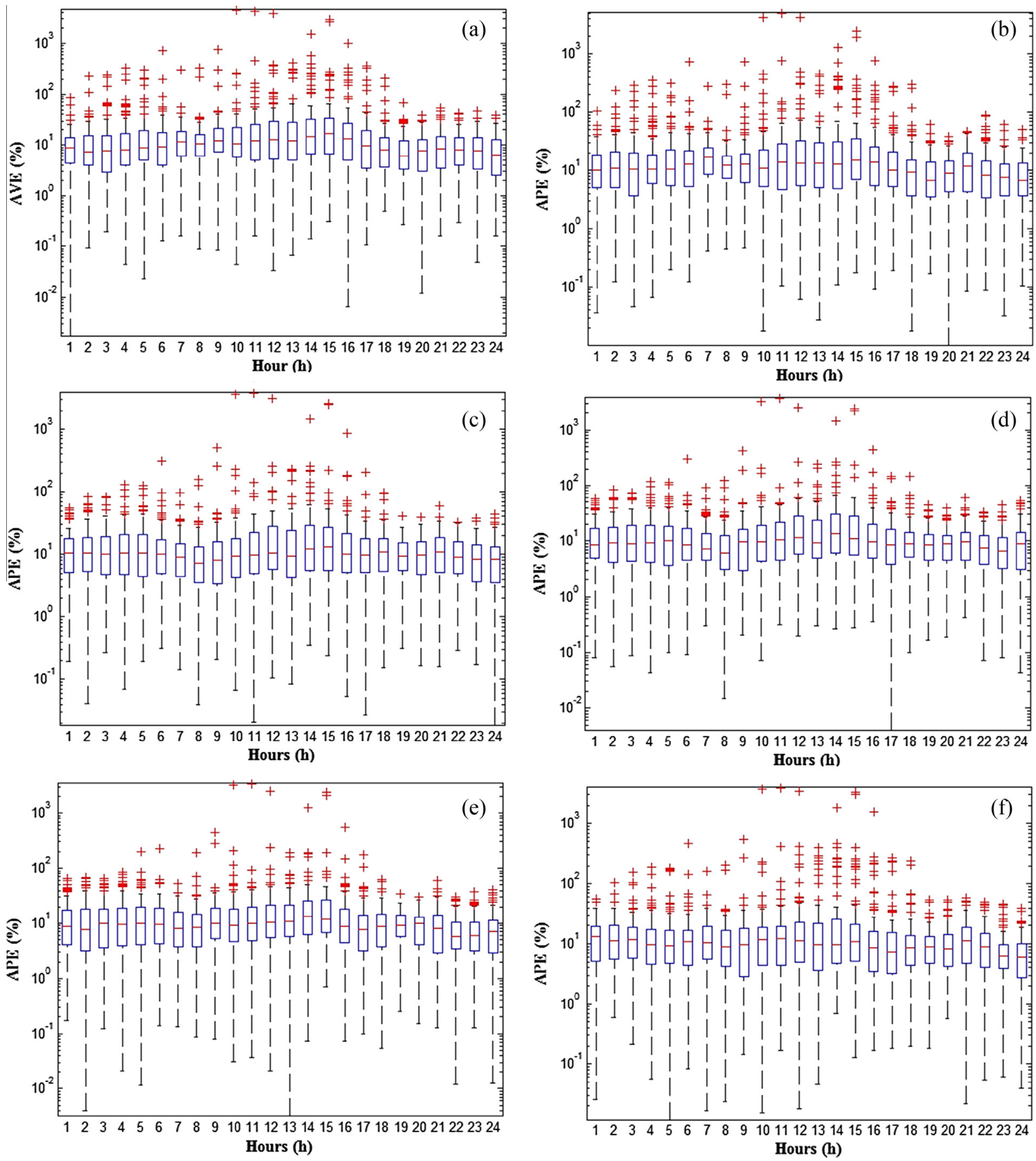

Fig. 13. Comparison of the models based on the hourly boxplots: (a) Model A, (b) Model B, (c) Model C, (d) Model D, (e) Model E and (f) Model F.

vertical axe display the number of instances for each error value. Values around zero refer to accurate forecasts. In all cases most Error values are limited in the $[-30\%, 30\%]$ range. By the examination of the histogram of Model A, it is observed some instances distributed beyond 60% while many instances are gathered between -10% and -20% and between 10% and 20%. On the contrary Model B limits the large positive values but has large negative values that are gathered around -60%. Again it is shown that Model E and Model D correspond to the more efficient forecasters for the problem under study.

As an illustrative example of the models operation behavior, Fig. 16 presents the actual and forecasted daily price curves of 2 holidays and 2 regular days of the test set. One day from every month has been selected. The holidays refer to the Theophany and Easter Sunday. The regular days refer to 2 working days of February and March, respectively. It is observed that in all cases the Model E lead to forecasts more similar with the actual price and this is more evident in the cases of the holidays loads.

Table 5 registers the models score on the MAE, U_1 and U_2 indicators as calculated on the test set data. The MAE cannot provide information on accuracy relative to the scales of the times series under investigation. This information is provided by MAPE and APE indicators. According to the results, Model E leads to the lowest MAE. According to this indicator, Model A leads to poor performance. Models C, D and E result in almost similar U_1 (i.e., the difference lies in the 3rd decimal point value). The U_2 statistic will

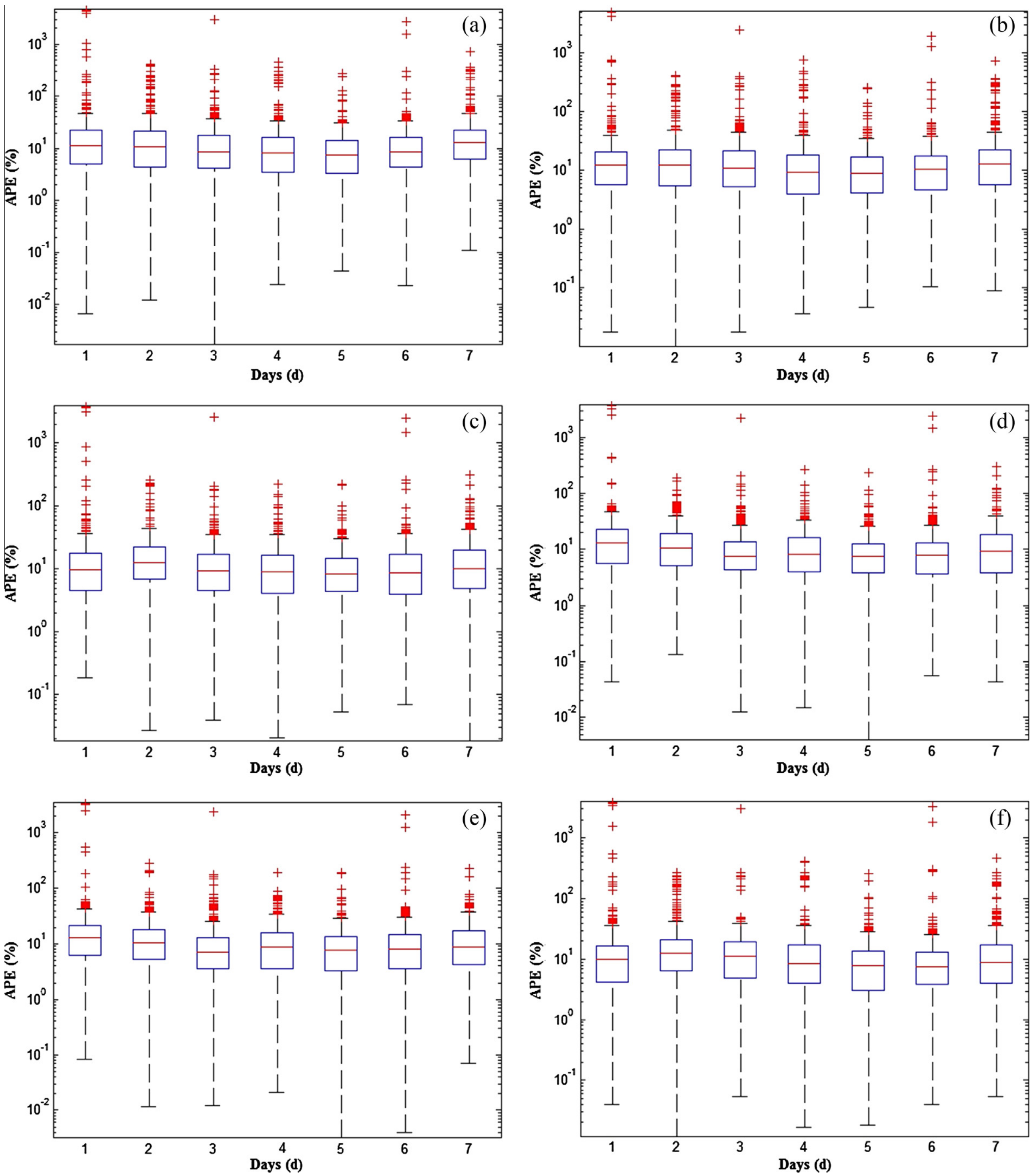

Fig. 14. Comparison of the models based on the daily boxplots: (a) Model A, (b) Model B, (c) Model C, (d) Model D, (e) Model E and (f) Model F.

receive the value 1 under the nave forecasting method. Using this indicator, the models performance comparison is more distinct. All models lead to values greater than 1. Again Model F results in better forecasting performance.

The proposed ANN based models are also compared with time series models. Specifically, the VAR, ARMA and GARCH models have been implemented. To keep the comparison feasible, we considered the following lag values: 24 and 168 and 24, 25, 47, 48, 72, 96, 120, 144, 167, 168. This lag values correspond to the historical prices used in Model A and Model B. The models are denoted as VAR Model A, ARMA Model A, ARMA-GARCH Model A, VAR Model B, ARMA Model B and ARMA-GARCH Model B, respectively. Since GARCH can only predict the conditional variance of a time series, it is hybridized we ARMA. A more accurate model estimation will involve a set of tests for parameter identification, i.e. using the Akaike Information Criterion. However, our main focus is the direct comparison with the structure of the basic ANN used in this study. After of series of experiments, it has been concluded that it is preferable to decreases the size of the training set. Recall that the total number of hours of the training set is 24914 corresponding to period between 01/02/2012 and 31/12/2014. All time series models lead to increased errors using large data sets for training. The set has been reduced to the 5192 h that are more close to the period covering the test set. VAR Model A and VAR Model B result to MAPE above 40%. These large values is due to the not optimal selection of lagged values in order to keep the comparison fair

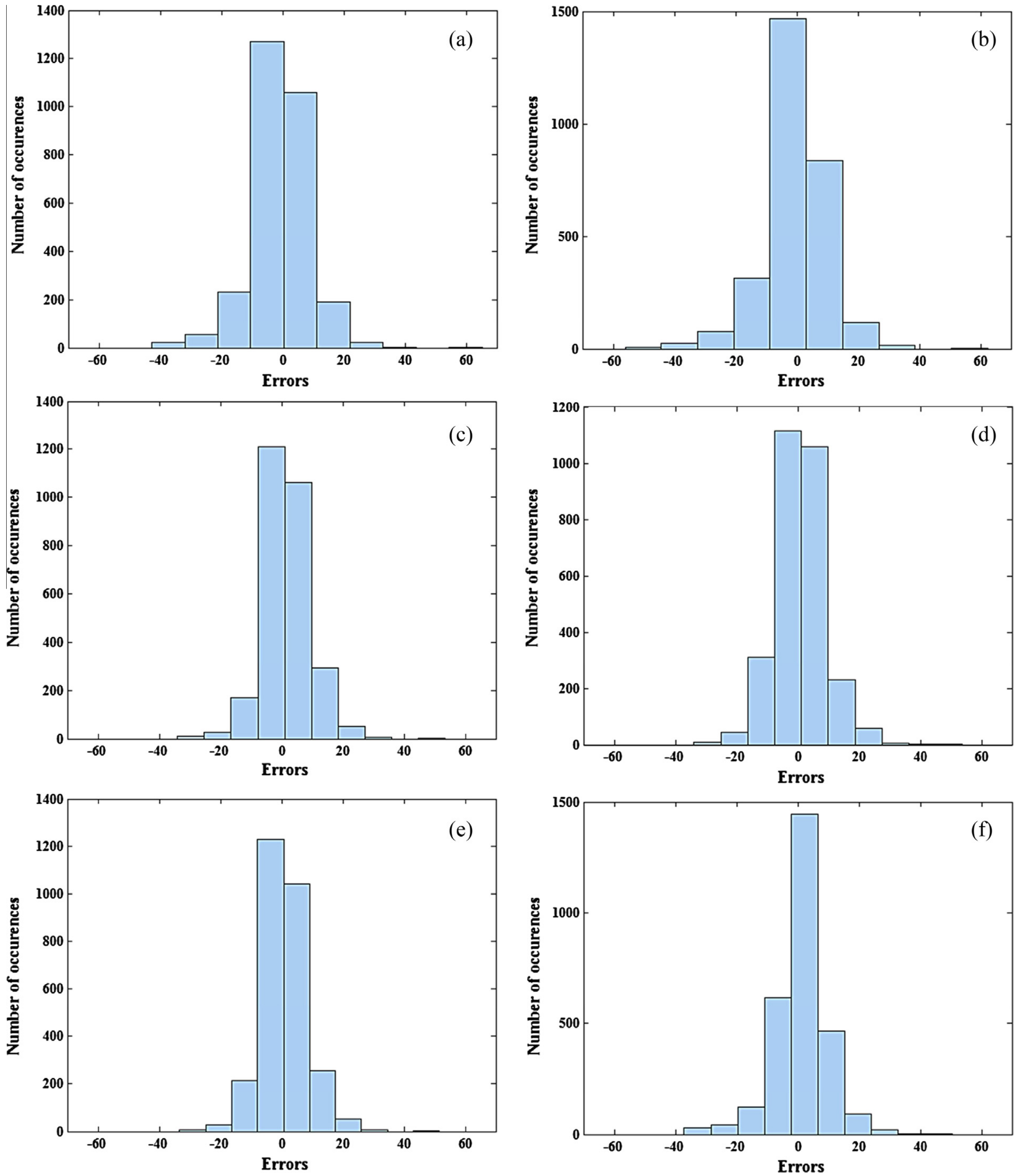

Fig. 15. Comparison of the models based on the errors distributions: (a) Model A, (b) Model B, (c) Model C, (d) Model D, (e) Model E and (f) Model F.

with the ANN models and due to the high non-linearity of the present data set. A VAR model works more accurately in time series with less volatility. ARIMA Model A and ARIMA Model B leads to test set MAPEs equal to 35.58% and 36.16%, respectively. This leads to the conclusion that are not recommended for utilization for the current data set. Almost similar performance present ARMA-GARCH Model A and ARMA-GARCH Model B. The respective MAPE are 36.16% and 37.46%, respectively. Table 6 shows the MAE, U_1 and U_2 values of the time series models. In comparison with the ANNs, MAEs of time series are considerably higher. This fact is more evident in the cases of VAR models. The sole and hybrid ARMA models lead to similar results when evaluated with the U_1 indicator. The sole ARMA models are superior to the others when the assessment is held with the U_2 indicator.

4. Discussion and concluding remarks

Generation electricity trends are largely affected by the market clearing price patterns. The importance of accurate predictions of market prices is manyfold. For instance, the Day-ahead strategy of a producer can be built and carried on reliable forecasts, leading

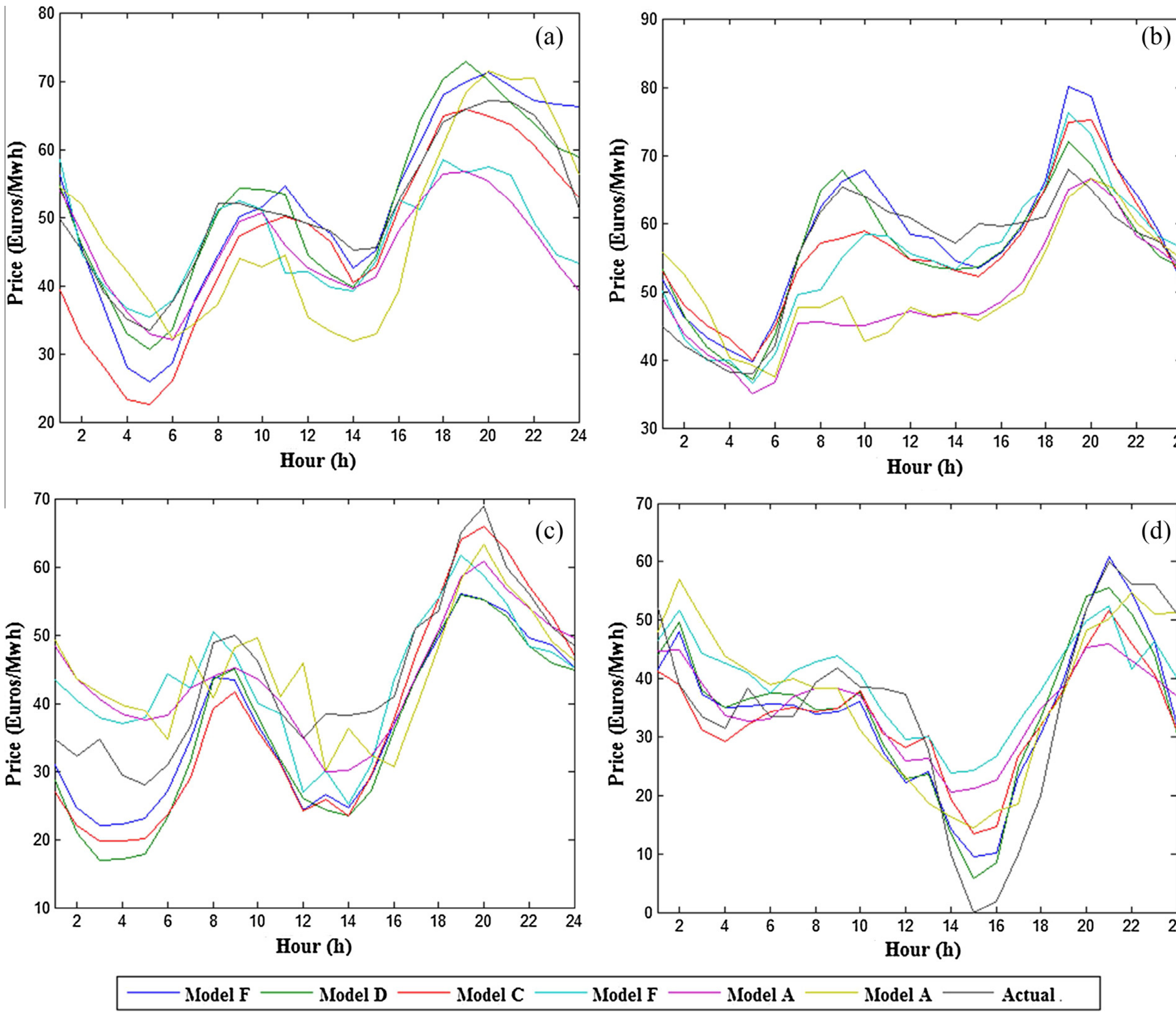

Fig. 16. Actual and forecasted prices loads of: (a) Theophany 06/01/2015, (b) Monday 16/02/2015, (c) Friday 13/03/2015 and (d) Easter Sunday 05/04/2015.

Table 5
Models comparison considering the MAE, U_1 and U_2 indicators.

Model	MAE	U_1	U_2
A	7.18	0.10	5.22
B	6.46	0.09	5.14
C	5.81	0.07	4.49
D	5.80	0.07	4.31
E	5.69	0.07	3.12
F	6.09	0.08	4.62

Table 6
Time series models comparison considering the MAE, U_1 and U_2 indicators.

Model	MAE	U_1	U_2
VAR Model A	14.23	0.18	7.22
VAR Model B	16.52	0.21	8.40
ARMA Model A	7.57	0.11	6.20
ARMA Model B	7.70	0.11	6.56
ARMA-GARCH Model A	9.05	0.12	7.04
ARMA-GARCH Model B	8.65	0.12	7.32

to increased profits, economic viability, limited risks and competiveness. While load forecasting counts many years of research efforts and application, the price forecasting literature is relatively more limited. Many markets were structured as oligopolies, hence the key players were the market's dominant influencing significantly the level of prices. As the energy markets landscape continually support competition, price forecasters become important for all market participants and there is need in the research community to present novel approaches. As the newcomers enter the wholesale and retail markets, price forecasting becomes a challenging contemporary engineering task.

The reported errors in load forecasting problems are already below 3%. This is not the case for price time series. The basic attribution of market clearing prices are their high volatility. System load depends mainly on weather conditions and seasonal effects while prices are influenced by a larger and diverse set of parameters. Due to market competition and regulation, many of these parameters may not be available to researchers or engineers. Thus, various combination of inputs should examine.

This paper serves as an introductory study in the Day-ahead price forecasting centered around the case of working with atypical data. Under this condition, several ANN based models are implemented. ANNs are a common technique for capturing and simulating the non-linear relationships between input data and target parameters. The aim is to test models with different input requirements when dealing with no pre-processed data sets. These kind of sets represent more accurately the Day-ahead Market transactions. According to the numerical results, using raw data will lead to high errors close to 20%. Hence, it can be conlcuded that more effort should be placed on models construction and evaluation. The models comparison point out that cascaded neural networks are the optimal selection among alternatives. Apart from load entries, the optimal training set should include natural gas prices, prices of other energy markets and renewable generation capacity. Models using only historical price values have not presented reliable performance. Among the aims of the present work is the implementation of a hybrid forecasting model, considering a two stage process. The clustering tool is combined with the cascaded ANN. The aim is to take advantage of the prons of both unsupervised and unsupervised machine learning. The hybrid model have not improved the efficiency of the cascaded network. However, the hybrid model led to the lowest errors in subsets of the main test set. The focus of the analysis of the present paper is next day's price curve. This means that only prices one day before the target day has been used. This is not followed by the majority of the related literature.

The paper fills gaps and supplements existing research in the literature, concerning the evolving issue of the Day-ahead Market price forecasting. The proposed ANN based models provide evidence that it could be considered as useful and robust forecasting tools to the actual needs of market participants, including the traditional generation companies and self-producers, but also the retailers/suppliers and aggregators.

Modern power systems design focusing more and more in the optimal balancing of generation and demand. The smart grid concept has risen due to the needs of scheduling more efficiently the energy sources and affecting the demand patterns. Usually, these needs are treated in parallel using advanced automation, metering and storage technologies. The consumers become increasingly important in power systems operations. Thus, energy service packages (such as real-time pricing) are served to the consumers in order to modify and smooth their demand patterns. In order to efficiently design, implement and evaluate the various demand response measures, robust short-term price forecasting in the Day-ahead transactions is needed. However, the price forecasting problem is a very challenging task. The price time series is volatile and is influenced by a diverse set of parameters (for example, weather conditions, hydro capacity, fossil fuel prices and others). The future knowledge gaps and challenges of the price forecasting problem lie to the following topics: (a) To build and test new computational intelligence based models, i.e. model design studies, (b) to consider other clustering algorithms, i.e. algorithm adoption studies and (c) to test the proposed models in other no preprocessed data sets, i.e. model test studies.

References

[1] Seifi H, Sepasian MS. Electric power systems planning: issues, algorithms and solutions. Berlin: Springer-Verlag; 2011.
[2] Hahn H, Meyer-Nieberg S, Pickl S. Electric load forecasting methods: Tools for decision making. Eur J Oper Res 2009;199:902–7.
[3] Yu S, Wang K, Wei YM. A hybrid self-adaptive particle swarm optimization-genetic algorithm-radial basis function model for annual electricity demand prediction. Energy Convers Manage 2015;2015(91):176–85.
[4] Aggarwal SK, Saini LM, Kumar A. Electricity price forecasting in deregulated markets: A review and evaluation. Int J Electr Power Energy Syst 2009;31:13–22.
[5] Bask M, Widerberg A. Market structure and the stability and volatility of electricity prices. Energy Econ 2009;31:278–88.
[6] Ziel F, Steinert R, Husmann S. Efficient modeling and forecasting of electricity spot prices. Energy Econ 2015;47:98–111.
[7] Singhal D, Swarup KS. Electricity price forecasting using artificial neural networks. Int J Electr Power Energy Syst 2011;2011(33):550–5.
[8] Aggarwal SK, Saini LM, Kumar A. Electricity price forecasting in Ontario electricity market using wavelet transform in artificial neural network based model. Int J Control Autom Syst 2008;6:639–50.
[9] Gareta R, Romeo LM, Gil A. Forecasting of electricity prices with neural networks. Energy Convers Manage 2006;47:1770–8.
[10] Priddy KL, Keller PH. Artificial neural networks: an introduction. Bellingham: SPIE Press; 2005.
[11] Bosco B, Parisio L, Pelagatt M. Deregulated wholesale electricity prices in Italy: an empirical analysis. Int Adv Econ Res 2007;13:415–32.
[12] Gianfreda A, Grossi L. Zonal price analysis of the Italian wholesale electricity market. In: Proceedings of the 2009 European energy markets conference, Belgium. p. 407–12.
[13] Gestore Mercati Energetici (GME). Annual report 2013. Rome, Italy; 2013.
[14] Ahmadi A, Charwand M, Aghaei J. Risk-constrained optimal strategy for retailer forward contract portfolio. Int J Electr Power Energy Syst 2013;53:704–13.
[15] Carrión M, Conejo AJ, Arroyo JM. Forward contracting and selling price determination for a retailer. IEEE Trans Power Syst 2007;22:2105–14.
[16] Weron R. Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast 2014;30:1030–81.
[17] Cerjan M, Krzelj I, Vidak M, Delimar M. A literature review with statistical analysis of electricity price forecasting methods. In: Proceedings of the 2013 EUROCON. p. 756–63.
[18] Conejo AJ, Contreras J, Espinola R, Plazas MA. Forecasting electricity prices for a day-ahead pool-based electric energy market. Int J Forecast 2005;21:435–62.
[19] Conejo AJ, Plazas MA, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;2005(20):1035–42.
[20] Cuaresma CJ, Hlouskova J, Kossmeier S, Obersteiner M. Forecasting electricity spot-prices using linear univariate time-series models. Appl Energy 2004;77:87–106.
[21] Diongue AK, Guégan D, Vignal B. Forecasting electricity spot market prices with a k-factor GIGARCH process. Appl Energy 2009;4:505–10.
[22] Abedinia O, Amjady N, Shafie-khah M, Catalão JPS. Electricity price forecast using Combinatorial Neural Network trained by a new stochastic search method. Energy Convers Manage 2015;105:642–54.
[23] Amjady N, Daraeepour A. Mixed price and load forecasting of electricity markets by a new iterative prediction method. Electr Power Syst Res 2009;79:1329–36.
[24] Amjady N, Daraeepour A. Design of input vector for day-ahead price forecasting of electricity markets. Expert Syst Appl 2009;36:12281–94.
[25] Amjady N, Hemmati M. Day-ahead price forecasting of electricity markets by a hybrid intelligent system. Eur Trans Electr Power 2009;19:89–102.
[26] Amjady N, Keynia F. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. Int J Electr Power Energy Syst 2008;30:533–46.
[27] Amjady N, Keynia F. Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. IEEE Trans Power Syst 2009;24:306–8.
[28] Amjady N, Keynia F. Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique. Energy Convers Manage 2009;50:2976–82.
[29] Anbazhagan S, Kumarappan N. Classification of day-ahead deregulated electricity market prices using DCT-CFNN. In: Panigrahi BK, Suganthan PN, Das S, Dash SS, editors. Swarm, evolutionary, and memetic computing, 4th international conference, SEMCCO 2013, Chennai, India, December 19–21, 2013, Proceedings, Part II, Springer International Publishing Switzerland.
[30] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price classification using neural network input featured by DCT. Int J Electr Power Energy Syst 2012;37:103–9.
[31] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT. Energy Convers Manage 2014;78:711–9.
[32] Bigdeli N, Afshar K, Amjady N. Market data analysis and short-term price forecasting in the Iran electricity market with pay-as-bid payment mechanism. Electr Power Syst Res 2009;79:888–98.
[33] Bunn DW. Forecasting loads and prices in competitive power markets. Proc IEEE 2000;88:163–9.
[34] Cataláo JPS, Mariano SJPS, Mendes VMF, Ferreira LAFM. Short-term electricity prices forecasting in a competitive market: a neural network approach. Energy Convers Manage 2007;77:1297–304.
[35] Cataláo JPS, Mariano SJPS, Mendes VMF, Ferreira LAFM. An artificial neural network approach for short-term electricity prices forecasting. Eng Int Syst 2007;1:15–23.
[36] Cataláo JPS, Pousinho HMI, Mendes VMF. Neural networks and wavelet transform for short-term electricity prices forecasting. Eng Int Syst 2010;2:85–92.
[37] Chakravarty S, Dash PK. Dynamic filter weights neural network model integrated with differential evolution for day-ahead price forecasting in energy market. Expert Syst Appl 2011;38:10974–82.
[38] Dev P, Martin MA. Using neural networks and extreme value distributions to model electricity pool prices: Evidence from the Australian National Electricity Market 1998–2013. Energy Convers Manage 2014;84:122–32.
[39] Gao F, Guan X, Cao XR, Papalexopoulos A. Forecasting power market clearing price and quantity using a neural network method. In: Proceedings of the 2000 IEEE power engineering society summer meeting, Seattle. p. 2183–8.
[40] Guo JJ, Luh PB. Improving market clearing price prediction by using a committee machine of neural networks. IEEE Trans Power Syst 2004;19:1867–76.
[41] Hong YY, Wu CP. Day-ahead electricity price forecasting using a hybrid principal component analysis network. Energies 2012;5:4711–25.

[42] Karsaz A, Rajabi Mashhadi H, Mirsalehi MM. Market clearing price and load forecasting using cooperative co-evolutionary approach. Int J Electr Power Energy Syst 2010;32:408–15.
[43] Keynia F. A new feature selection algorithm and composite neural network for electricity price forecasting. Eng Appl Artif Intell 2012;25:1687–97.
[44] Lora AT, Santos JMR, Exposito AG, Ramos JLM, Santos JCR. Electricity market price forecasting based on weighted nearest neighbors techniques. IEEE Trans Power Syst 2007;22:1294–301.
[45] Mandal P, Senjyu T, Funabashi T. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. Energy Convers Manage 2006;47:2128–42.
[46] Mandal P, Senjyu T, Funabashi T, Srivastava AK. A novel approach to forecast electricity price for PJM using neural network and similar days method. IEEE Trans Power Syst 2007;22:2058–65.
[47] Menniti D, Scordino N, Sorrentino N. Forecasting next-day electricity prices by a neural network approach. In: Proceedings of the 8th international conference on European Energy Market, Zagreb. p. 209–15.
[48] Neupane B, Perera KS, Aung Z, Woon WL. Artificial neural network-based electricity price forecasting for smart grid deployment. In: Proceedings of the 2012 international conference on computer systems and industrial informatics, Sharjah. p. 1–6.
[49] Pao HT. Forecasting electricity market pricing using artificial neural networks. Energy Convers Manage 2007;48:907–12.
[50] Peng TM, Hubele NF, Karady GG. Advancement in the application of neural networks for short-term load forecasting. IEEE Trans Power Syst 1992;7:250–7.
[51] Pindoriya NM, Singh SN, Singh SK. An adaptive wavelet neural network-based energy price forecasting in electricity markets. IEEE Trans Power Syst 2008;23:1423–31.
[52] Pino R, Parreno J, Gomez A, Priore P. Forecasting next-day price of electricity in the Spanish energy market using artificial neural networks. Eng Appl Artif Intell 2008;21:53–62.
[53] Ranjbar M, Soleymani S, Sadati N, Ranjbar AM. Electricity price forecasting using artificial neural network. In: Proceedings of 2006 international conference on the power electronics, drives and energy systems, New Delhi. p. 1–5.
[54] Rodriguez CP, Anders GJ. Energy price forecasting in the Ontario competitive power system market. IEEE Trans Power Syst 2004;19:366–74.
[55] Sapeluk A, Ozveren CS, Birch AP. Pool price forecasting: a neural network application. In: Proceedings of the 1994 UPEC conference, Calway. p. 840–3.
[56] Sestito S, Leung S, Dillon TS. Short term load forecasting using an adaptive neural network. Int J Electr Power Energy Syst 1991;13:186–92.
[57] Szkuta BR, Sanabria LA, Dillon TS. Electricity price short-term forecasting using artificial neural networks. IEEE Trans Power Syst 1999;14:851–7.
[58] Toyama H, Senjyu T, Areekul P, Chakraborty S, Yona A, Funabashi T. Next-day electricity price forecasting on deregulated power market. In: Proceedings of the 2009 transmission & distribution conference & exposition: Asia and Pacific, Seoul. p. 1–4.
[59] Unsihuay-Vila C, Zambroni de Souza AC, Marangon-Lima JW, Balestrassi PP. Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. Int J Electr Power Energy Syst 2010;32:108–16.
[60] Vahidinasab V, Jadid S, Kazemi A. Day-ahead price forecasting in restructured power systems using artificial neural networks. Electr Power Syst Res 2008;78:1332–42.
[61] Wang Z, Liu F, Wu J, Wang J. A hybrid forecasting model based on bivariate division and a backpropagation artificial neural network optimized by chaos particle swarm optimization for day-ahead electricity price. Abstr Appl Anal 2014:1–31.
[62] Wang A, Ramsay B. Prediction of system marginal price in the UK power pool using neural networks. In: Proceedings of the international conference on neural networks 1997, Houston. p. 2116–40.
[63] Wang AJ, Ramsay B. A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays. Neurocomputing 1998;23:47–57.
[64] Wu L, Shahidehpour M. A hybrid model for day-ahead price forecasting, power systems. IEEE Trans Power Syst 2010;25:1519–30.
[65] Yamin HY, Shahidehpour SM, Li Z. Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets. Int J Electr Power Energy Syst 2004;26:571–81.
[66] Zhang L, Luh PB, Kasiviswanathan K. Energy clearing price prediction and confidence interval estimation with cascaded neural network. IEEE Trans Power Syst 2003;18:99–105.
[67] Zhang L, Luh PB. Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method. IEEE Trans Power Syst 2005;20:59–66.
[68] Amjady N, Keynia F. Application of a new hybrid neuro-evolutionary system for day-ahead price forecasting of electricity markets. Appl Soft Comput 2010;10:784–92.
[69] Areekul P, Senjyu T, Toyama H, Yona A. A hybrid ARIMA and neural network model for short-term price forecasting in deregulated market. IEEE Trans Power Syst 2010;25:524–30.
[70] Khosravi A, Nahavandi S, Creighton D. A neural network-GARCH-based method for construction of Prediction Intervals. Electr Power Syst Res 2013;96:185–93.
[71] Amjady N, Daraeepour A, Keynia F. Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. IET Gener Transm Distrib 2010;4:432–44.
[72] Hassan S, Khosravi A, Jaafar J, Raza MQ. Electricity load and price forecasting with influential factors in a deregulated power industry. In: Proceedings of the 2014 9th international conference on system of systems engineering (SOSE). p. 79–84.
[73] Sandhu HS, Fang L, Guan L. Forecasting day-ahead electricity prices using data mining and neural network techniques. In: Proceedings of the 2014 11th international conference on service systems and service management (ICSSSM). p. 1–6.
[74] Sahay KB, Tripathi MM. Day ahead hourly load and price forecast in ISO New England market using ANN. In: Proceedings of the 2013 annual IEEE india conference (INDICON). p. 1–6.
[75] Agarwal A, Ojha A, Tewari SC, Tripathi MM. Hourly load and price forecasting using ANN and fourier analysis. In: Proceedings of the 2014 6th IEEE power india international conference (PIICON). p. 1–6.
[76] Sahay KB. One hour ahead price forecast of Ontario electricity market by using ANN. In: Proceedings of the 2015 international conference on energy economics and environment (ICEEE). p. 1–6.
[77] Filho JCR, Affonso CM, Oliveira RCL. Energy price forecasting in the North Brazilian market using NN – ARIMA model and explanatory variables. In: Proceedings of the 2014 IEEE symposium on computational intelligence for engineering solutions (CIES). p. 171–5.
[78] Tan Z, Zhang J, Wang J, Xu J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87:3606–10.
[79] Keles D, Scelle J, Paraschiv F, Fichtner W. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. Appl Energy 2016;162:218–30.
[80] Coelho LDS, Santos AAP. A RBF neural network model with GARCH errors: Application to electricity price forecasting. Electr Power Syst Res 2011;81:74–83.
[81] Lin WM, Gow HJ, Tsai MT. An enhanced radial basis function network for short-term electricity price forecasting. Appl Energy 2010;87:3226–34.
[82] Shafie-khah M, Parsa Moghaddam M, Sheikh-El-Eslami MK. Price forecasting of day-ahead electricity markets using a hybrid forecast method. Energy Convers Manage 2011;52:2165–9.
[83] Yun Z, Quan Z, Caixin S, Shaolan L, Yuming L, Yang S. RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. IEEE Trans Power Syst 2008;23:853–8.
[84] Yan-Gao C, Guangwen M. Electricity price forecasting based on support vector machine trained by genetic algorithm. In: Proceedings of the third international symposium on intelligent information technology application, China. p. 292–5.
[85] Elattar E, Shebin EK. Day-ahead price forecasting of electricity markets based on local informative vector machine. IET Gener Trans Distrib 2013;7:1063–71.
[86] Shayeghi H, Ghasemi A. Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. Energy Convers Manage 2013;74:482–91.
[87] Shayeghi H, Ghasemi A, Moradzadeh M, Nooshyar M. Simultaneous day-ahead forecasting of electricity price and load in smart grids. Energy Convers Manage 2015;95:371–84.
[88] Zhao JH, Dong ZY, Xu Z, Wong KP. A statistical approach for interval forecasting of the electricity price. IEEE Trans Power Syst 2008;23:267–76.
[89] Zhang J, Tan Z, Yang S. Day-ahead electricity price forecasting by a new hybrid method. Comput Ind Eng 2012;63:695–701.
[90] Zhang J, Tan Z. Day-ahead electricity price forecasting using WT, CLSSVM and EGARCH model. Int J Electr Power Energy Syst 2013;45:363–8.
[91] Shiriz A, Afsharx M, Rahimi-Kian A, Maham B. Electricity price forecasting using Support Vector Machines by considering oil and natural gas price impacts. In: Proceedings of the 2015 IEEE international conference on smart energy grid engineering (SEGE). p. 1–5.
[92] Shrivastava NA, Khosravi A, Panigrahi BK. Prediction interval estimation for electricity price and demand using support vector machines. In: Proceedings of the 2014 international joint conference on neural networks (IJCNN). p. 3995–4002.
[93] Amjady N. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. IEEE Trans Power Syst 2006;21:887–96.
[94] Cataláo JPS, Pousinho HMI, Mendes VMF. Short-term electricity prices forecasting in a competitive market by a hybrid intelligent approach. Energy Convers Manage 2011;52:1061–5.
[95] Cataláo JPS, Pousinho HMI, Mendes VMF. Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. IEEE Trans Power Syst 2011;26:137–44.
[96] Chogumaira EN, Hiyama T. Short-term electricity price forecasting using a combination of neural networks and fuzzy inference. Energy Power Eng 2011;3:9–16.
[97] Ghodsi R, Zakerinia M, Jokar M. Neural network and fuzzy regression model for forecasting short term price in Ontario electricity market. In: Proceedings of the 41st international conference on computers & industrial engineering. p. 954–9.
[98] Hong YY, Lee CF. A neuro-fuzzy price forecasting approach in deregulated electricity markets. Electr Power Syst Res 2005;73:151–7.

[99] Niu D, Liu D, Wu DD. A soft computing system for day-ahead electricity price forecasting. Appl Soft Comput 2010;10:868–75.
[100] Osório GJ, Matias JCO, Cataláo JPS. Electricity prices forecasting by a hybrid evolutionary-adaptive methodology. Energy Convers Manage 2014;80:363–73.
[101] Pousinho HMI, Mendes VMF, Catalão JPS. Hybrid PSO-ANFIS approach for short-term electricity prices prediction. In: Proceedings of the 2010 PES general meeting, Michigan. p. 1–6.
[102] Zhou H, Wu XH, Li XG. An ANFIS model of electricity price forecasting based on subtractive clustering. In: Proceedings of the 2011 IEEE power and energy society general meeting, San Diego. p. 1–5.
[103] Pousinho HMI, Mendes VMF, Catalão JPS. Short-term electricity prices forecasting in a competitive market by a hybrid PSO-ANFIS approach. Int J Electr Power Energy Syst 2012;39:29–35.
[104] Vahidinasab V, Jadid S. Bayesian neural network model to predict day-ahead electricity prices. Eur Trans Electr Power 2008;20:231–46.
[105] Vanaja L, Fung C, Gedeon T. A fuzzy-neural approach to electricity load and spot price forecasting in a deregulated electricity market. In Proceedings of the 2003 Conference on Convergent Technologies for Asia-Pacific Region, Bangalore. p. 1479–82.
[106] Bashari M, Darudi A, Raeyatdoost N. Kalman Fusion algorithm in electricity price forecasting. In: Proceedings of the 2014 14th international conference on environment and electrical engineering (EEEIC). p. 313–7.
[107] Darudi A, Bashari M, Javidi MH. Electricity price forecasting using a new data fusion algorithm. IET Gener Transm Distrib 2015;9:1382–90.
[108] Osorio GJO, Matias JCO, Catalao JPS. Hybrid evolutionary-adaptive approach to predict electricity prices and wind power in the short-term. In: Proceedings of the 2014 power systems computation conference (PSCC). p. 1–7.
[109] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price forecasting using recurrent neural network. IEEE Syst J 2012;7:866–72.
[110] Niu D, Liu D, Xing M. Electricity price forecasting using generalized regression neural network based on principal components analysis. J Cent South Univ Technol 2009;15:316–20.
[111] Amjady N, Keynia F. Electricity market price spike analysis by a hybrid data model and feature selection technique. Electr Power Syst Res 2010;80:318–27.
[112] Amjady N, Keynia F. A new prediction strategy for price spike forecasting of day-ahead electricity markets. Appl Soft Comput 2011;11:4246–56.
[113] Lin WM, Gow HJ, Tsai MT. Electricity price forecasting using enhanced probability neural network. Energy Convers Manage 2010;51:2707–14.
[114] Lu X, Dong ZY, Li X. Electricity market price spike forecast with data mining techniques. Elect Power Syst Res 2005;73:19–29.
[115] Voronin S, Partanen J. Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spike frameworks. Energies 2013;6:5897–920.
[116] Wang QQ, Dong ZY, Li X, Zhao J, Wong KP. An effective approach to predicting electricity market price spikes. In: Proceedings of the 2007 IEEE power engineering society general meeting, Tampa. p. 1–7.
[117] Wu W, Zhou J, Mo L, Zhu C. Forecasting electricity market price spikes based on Bayesian expert with support vector machine. In: Li X, Zaïane OR, Li Z, editors. Advanced data mining and applications. Volume 4093 of the series lecture notes in computer science. Springer; 2006. p. 205–12.
[118] Zao JH, Dong ZY, Li X, Wong KP. A general method for electricity market price spike analysis. In: Proceedings of the 2005 IEEE power engineering society general meeting, San Francisco. p. 286–93.
[119] Zao JH, Dong ZY, Li X, Wong KP. A framework for electricity price spike analysis with advanced data mining methods. IEEE Trans Power Syst 2007;22:376–85.
[120] Zao JH, Dong ZY, Li X. Electricity market price spike forecasting and decision making. IET Gener Transm Distrib 2007;1:647–54.
[121] Kim CL, Yu IK, Song YH. Kohonen neural network and wavelet transform based approach to short-term load forecasting. Electr Power Syst Res 2002;63:169–76.
[122] Nagi J, Yap KS, Nagi F, Tiong SK, Ahmed SK. A computational intelligence scheme for the prediction of the daily peak load. Appl Soft Comput 2011;11:4773–88.
[123] Fan S, Mao C, Chen L. Next-day electricity-price forecasting using a hybrid network. IET Gener Trans Distrib 2007;1:176–82.
[124] Jin CH, Pok G, Lee Y, Park HW, Kim KD, Yun U, et al. A SOM clustering pattern sequence-based next symbol prediction method for day-ahead direct electricity load and price forecasting. Energy Convers Manage 2015;90:84–92.
[125] Jin CH, Pok G, Paik I, Ryu KH. Short-term electricity load and price forecasting based on clustering and next symbol prediction. IEJJ Trans Electr Electron Eng 2015;10:175–80.
[126] Martinez-Alvarez F, Troncoso A, Riquelme JC, Aguilar-Ruiz JS. Energy time series forecasting based on pattern sequence similarity. IEEE Trans Knowl Data Eng 2011;23:1230–43.
[127] Bompard E, Ciwei G, Napoli R, Torelli F. Dynamic price forecast in a competitive electricity market. IET Gener Transm Distrib 2007;1:776–83.
[128] Hong YY, Hsiao CY. Locational marginal price forecasting in deregulated electricity markets using artificial intelligence. IET Proc Gener Transm Distrib 2002;149:621–6.
[129] Mandal P, Haque AU, Meng J, Srivastava A, Martinez R. A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting. IEEE Trans Power Syst 2013;28:1041–51.
[130] Shrivastava NA, Panigrahi BK. A hybrid wavelet-ELM based short term price forecasting for electricity markets. Int J Electr Power Energy Syst 2014;55:41–50.
[131] Wan C, Xu Z, Wang Y, Dong ZY, Wong KP. A hybrid approach for probabilistic forecasting of electricity price. IEEE Trans Smart Grid 2014;5:463–70.
[132] Singh NK, Singh AK, Tripathy M. Short-term load/price forecasting in deregulated electric environment using ELMAN neural network. In: Proceedings of the 2015 international conference on energy economics and environment (ICEEE). p. 1–6.
[133] Safari MIKM, Dahlan NY, Razali NS, Rahman TKA. Electricity prices forecasting using ANN hybrid with invasive weed optimization (IWO). In: Proceedings of the 2013 IEEE 3rd international conference on system engineering and technology (ICSET). p. 275–80.
[134] Sarada K, Bapiraju V. Comparison of day-ahead price forecasting in energy market using neural network and genetic algorithm. In: Proceedings of the 2014 international conference on smart electric grid (ISEG). p. 1–5.
[135] Kim MK. Short-term price forecasting of Nordic power market by combination Levenberg–Marquardt and Cuckoo search algorithms. IET Gener Transm Distrib 2015;9:1553–63.
[136] Amjady N. Short-term electricity price forecasting. In: Catalão JPS, editor. Electric power systems: advanced forecasting techniques and optimal generation scheduling. Boca Raton: CRC Press, Taylor & Francis Group; 2012. p. 4-1–4-58.
[137] Xu R, Wunsch D. Clustering. New Jersey: John Wiley & Sons. Inc.; 2006.
[138] Sugunnasil P, Somhom S, Jumpamule W, Tongsiri N. Modelling a neural network using an algebraic method. SCIENCEASIA 2014;40:94–100.
[139] Graupe P. Principles of artificial neural networks. 2nd ed. Massachusetts: World Scientific; 2007.
[140] Panapakidis IP, Alexiadis MC, Papagiannis, GK. Load profiling in the deregulated electricity markets: a review of the applications. In: Proceedings of the 9th international conference on the European energy, Italy; 2012. p. 1–8.