

[Browse Conference Publications](#) > [Fuzzy Systems \(FUZZ\), 2013 IE ...](#) ?

# Short-term electric load forecasting using computational intelligence methods

 **Full Text**  
[Sign-In or Purchase](#)**5**  
Author(s)Jurado, Sergio ; Sensing & Control Systems, Aragón 208-210, 08011 Barcelona, Spain ; Peralta, Juan ; Nebot, Angela ; Mugica, Francisco  
more authors

Abstract	Authors	References	Cited By	Keywords	Metrics	Similar
----------	---------	------------	----------	----------	---------	---------

## Author(s)

Jurado, Sergio

Sensing &amp; Control Systems, Aragón 208-210, 08011 Barcelona, Spain

Peralta, Juan ; Nebot, Angela ; Mugica, Francisco ; Cortez, Paulo

0

Like

0

Tweet

0

Share

[Sign In](#) | [Create Account](#)

### IEEE Account

[Change Username/Password](#)[Update Address](#)

### Purchase Details

[Payment Options](#)[Order History](#)[Access Purchased Documents](#)

### Profile Information

[Communications Preferences](#)[Profession and Education](#)[Technical Interests](#)

### Need Help?

**US & Canada:** +1 800 678 4333**Worldwide:** +1 732 981 0060[Contact & Support](#)[About IEEE Xplore](#) | [Contact](#) | [Help](#) | [Terms of Use](#) | [Nondiscrimination Policy](#) | [Site Map](#) | [Privacy & Opting Out of Cookies](#)A not-for-profit organization, IEEE is the world's largest professional association for the advancement of technology.  
© Copyright 2013 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.

# Short-term Electric Load Forecasting Using Computational Intelligence Methods

Sergio Jurado\*, Juan Peralta†, Àngela Nebot‡, Francisco Mugica‡ and Paulo Cortez§

\*Sensing & Control Systems

Aragó 208-210, 08011 Barcelona, Spain

Email: sergio.jurado@sens-ingcontrol.com

†Universitat Autònoma de Barcelona

Campus UAB Edificio O, 08193 Bellaterra, Spain

Email: jperalta@cvc.uab.es

‡Technical University of Catalonia

Jordi Girona Salgado, 1-3, 08034 Barcelona, Spain

Email: {angela,fmugica}@lsi.upc.edu

§Centro Algoritmi, University of Minho,

Guimaraes, Portugal

Email: pcortez@dsi.uminho.pt

**Abstract**—Accurate time series forecasting is a key issue to support individual and organizational decision making. In this paper, we introduce several methods for short-term electric load forecasting. All the presented methods stem from computational intelligence techniques: Random Forest, Nonlinear Autoregressive Neural Networks, Evolutionary Support Vector Machines and Fuzzy Inductive Reasoning. The performance of the suggested methods is experimentally justified with several experiments carried out, using a set of three time series from electricity consumption in the real-world domain, on different forecasting horizons.

**Keywords**—Artificial Neural Networks, Evolutionary Computation, Support Vector Machines, Random Forest, Time Series, Forecast.

## I. INTRODUCTION

The integration of time-series forecasting into existing energy grids and new Smart Grids brings up some major technical challenges that have to be addressed. The interaction between advanced communications infrastructures, mathematical modeling techniques, and numerical simulation environments is a promising approach in this area. This also holds for the potential storage capacity for both electrical and thermal energy within energy networks, which can be achieved by intelligent demand side management. A major requirement in the Smart Grid is to schedule the trading of energy between different consumers and producers (a.k.a. prosumers). According to the Smart Grids European Technology Platform, a large fraction of the generation capacity in 2035 will be stochastic and/or intermittent [1].

Both challenges have to be handled by the following options: "stand by" power plants, storage, matching over large geographic areas and improved demand-side management in industry and households. Forecasting permits better and more efficient management of all options.

Electric load forecasting has been studied largely since the early '50s, because it is crucial to the utilities and the

power industry as it determines the required future expansion of generation, transmission and distribution of electric energy. In this scenario, it is essential to integrate forecasting models of consumptions and/or productions that dwellings, public buildings, industries and possibly other structures, are going to generate. In order to achieve this goal, several approaches have been studied in the last decades; from the most classical statistical models, mainly based on regression methods and statistical analysis, to dynamic regression techniques, such as autoregressive moving average [2] or autoregressive distributed-lag models, among others, that have traditionally been used in short-term electric load forecasting ([3], [4]).

These techniques, although reliable, do not give very good performance when dealing with the non-linearity associated to the electric load time series. On the other hand, a large variety of Artificial Intelligent (AI) techniques have been applied in the field of short-term electricity consumption (electric load) forecasting, showing a better performance than classical techniques. Specifically, Machine Learning and Soft Computing techniques have been proven to represent electric consumption uncertainties with very good detail. For instance, in [5] Alon and colleagues have analysed Artificial Neural Network (ANN) versus other traditional methods such as ARIMA, Winters exponential smoothing, Box-Jenkins and multivariate analysis, on retail sales data. Khamis proposes in [6] a multilayer perceptron neural network to predict the electricity consumption for a small scale power system, obtaining a better performance than with traditional methods, while Marvuglia and colleagues consider Elman neural network for the short forecasting of the household electric consumption with interesting prediction errors under 5% [7]. Also in [8] a study of electric load forecasting is carried out with CART and other soft computing techniques obtaining again better results than classical approaches. Soft computing and hybrid techniques have become popular in this field, for instance there are several works that combines fuzzy logic and expert systems ([9], [10]). Moreover, fuzzy logic and expert systems have been also integrated with ANN in other interesting research ([11],

[12]). These hybrid techniques offer the advantages of each of the individual approaches, reducing their weaknesses.

Large scale studies for comparing machine learning and soft computing tools have focused on the classification domain [13]. On the contrary, a very few extensive studies can be found in the regression domain. In [14] Nesrren and colleagues carried out a large scale comparison of machine learning models for time series forecasting. The study includes techniques such as K-nearest neighbours, CART regression trees, multilayer perceptron networks, Support Vector Machines, Gaussian processes, Bayesian Neural Networks and radial basis functions. The research reveals significant differences between the methods studied and concludes that the best techniques for time series forecasting are multilayer perceptron and Gaussian regression when applied on the monthly M3 time series competition data (a thousand time series) [15]. Moreover, in [16] an empirical comparison of regression analysis, decision trees and ANN techniques for the prediction of electricity energy consumption is carried out, concluding that the decision tree and neural network models perform slightly better than regression analysis in the summer and winter phases, respectively. However, the differences between the three types of models are quite small in general, indicating that the three modeling techniques are comparable when predicting energy consumption.

These studies are confined to machine learning techniques and do not expand to a general comparison between models with different characteristics such as novel hybrid machine learning or hybrid soft computing approaches. Some of the most recent approaches are based on dynamic regression techniques, ANN ([17]), Neuro-Fuzzy approaches [18], input-output hidden Markov models [19] and Rain Forest models [20]. In this study we compare the performance of four of these approaches in the electricity forecasting scenario: Random Forest, Nonlinear Autoregressive Neural Networks, Evolutionary Support Vector Machine and Fuzzy Inductive Reasoning Methodology.

In addition, the quality of the algorithms has to be evaluated not only by the accuracy of forecasting, but also by other criteria, e.g. is the algorithm able to deliver forecasts in real time for grid operations? Can a forecasting method cope with incomplete measurement/sensor data? Does the algorithm have any special strength, like forecasting signals of great variability? Within this paper, we present the above mentioned Computational Intelligence (CI) techniques to predict time series forecasting of electricity consumptions that can deal with all these issues. Apart from that, the data used in our work is from real building consumptions. Thus, our results can provide a better understanding on the demand side management and buildings profiling.

The integration of forecasting models of consumptions and/or productions are among the main motivations for this paper, which has a twofold goal: 1) to provide readers with knowledge of distinct methods that may serve as a first step to find generalizable patterns and corresponding models within the energy consumption data, which may be used afterwards for early predictions on real time data in the Smart Grid; and 2) to compare the models proposed and challenge prior evidence on the forecasting accuracy of some CI techniques to predict electricity consumptions [21].

The paper is organized as follows. Section II describes briefly the CI methodologies used in this research for short-term electric load forecasting. In section III we describe the seasonal datasets, introduce the model forecasting evaluation and present and discuss the obtained results. Finally, section IV concludes the paper.

## II. COMPUTATIONAL INTELLIGENCE TECHNIQUES FOR TIME SERIES FORECASTING

In this research, we compare four CI approaches for multistep seasonal Time Series Forecasting (TSF): random forest, nonlinear autoregressive neural networks, evolutionary support vector machines and fuzzy inductive reasoning. The last two are hybrid methods that combine different soft computing approaches.

### A. Random Forest

RF is a set of classification and regression trees (CART), that was first put forward by Breiman (2001) [22]. In RF, the training sample set for a base classifier is constructed by using the Bagging algorithm [23]. When building a base classifier, inner nodes are spitted with a random candidate attribute set. The final classification rule or regression function is the simple majority voting method or the simple average method.

In traditional CART, each inner node is a subset of initial data set and the root node contains all the initial data. Random forests for regression are formed by growing trees depending on a random vector such that the tree predictor takes on numerical values as opposed to class labels. The output values are numerical and we assume that the training set is independently drawn from the distribution of the random vector. The random forest predictor is formed by taking the average over  $k$  of the trees.

The importance of a feature in RF is based on a very simplistic rule: if a variable is important, then random exchanging the value of this feature significantly reduces the prediction accuracy:

- 1) Calculate Out-Of-Bag (OOB) accuracy [24].
- 2) Change the value of a feature randomly, and then compute again the OOB accuracy, named as noised OOB accuracy.
- 3) Calculate the difference of the original OOB accuracy and the noised OOB accuracy. This difference is taken as the measure of the importance for this feature.
- 4) The importance of feature is ranked according the difference of original OOB accuracy and the noised OOB accuracy.

Breiman has proved that for both random forest classification and regression, the generalization error converges to a limit as the number of trees becomes larger [22]. Random forest algorithm has many advantages, for instance, providing accurate results, generalizing well and learning fast. In addition, it is suitable to handle the “missing data” problem and provides a tree structured method for regression [25].

## B. Nonlinear Autoregressive Neural Networks

Many time-series models are based on nonlinear Autoregressive Neural Networks (ARNN) [26]. It consists in a single input which is the present value and the  $p - 1$  past values of the series of returns, while the output is the value for the next time period, computed as described in equation 1.

$$y(t) = f(y(t-1), \dots, y(t-d)) \quad (1)$$

where  $d$  is the number of past values of  $y(t)$ .  $f$  is a nonlinear function approximated by a multilayer feedforward neural network (FNN) [27]. Unlike Recurrent Neural Networks (RNN), which are very powerful but can be extremely complicated and slower than the conventional networks due to the loops, the autoregressive multi-context recurrent neural network improves the speed of the training session, due to a reduction of the recurrent connections and is an appropriate method for approximating daily peak load [28].

## C. Evolutionary Support Vector Machine

The problem of designing Support Vector Machines (SVM) for time series forecasting can be seen as a search problem into the space of all possible solutions. While several evolutionary computation methods could be used for this search, we adopt the Estimation of Distribution Algorithm (EDA), since it has outperformed the standard genetic algorithms in previous works [29]. When designing a Evolutionary Support Vector Machine (ESVM), there are three crucial issues: i) setting the solution's space, i.e., what information of the SVM is previously set and what is included into the chromosome; ii) how each solution is codified into a chromosome, i.e., encoding schema; and iii) what are we optimizing, i.e., defining the fitness function. When designing a SVM, there are also three important issues to take into account: the type of SVM to use, the selection of the kernel function and tuning the parameters associated with the two previous selections. Since TSF is a particular regression case, for the SVM type and kernel, we selected the popular  $\varepsilon$ -insensitive loss function (known as  $\varepsilon$ -SVR) and Gaussian kernel combination, as implemented in the LIBSVM tool [30]. In SVM regression [31], the input  $\mathbf{y} = (y_{t-k_I}, \dots, y_{t-k_2}, y_{t-k_1})$ , for a SVM with  $I$  inputs, is transformed into a high  $m$ -dimensional feature space, by using a nonlinear mapping ( $\phi$ ) that does not need to be explicitly known but that depends on a kernel function. Then, the SVM algorithm finds the best linear separating hyperplane, tolerating a small error ( $\varepsilon$ ) when fitting the data, in the feature space:

$$\hat{y}_t = w_0 + \sum_{j=1}^m w_j \phi_j(\mathbf{y}) \quad (2)$$

This model requires setting three parameters:  $\gamma$  – the Gaussian kernel parameter,  $\exp(-\gamma \|x - x'\|^2)$ ,  $\gamma > 0$ ;  $C$  – a trade-off between fitting the errors and the flatness of the mapping; and  $\varepsilon$  – the width of the  $\varepsilon$ -insensitive tube.

In this paper, an evolving hybrid system that uses EDA and SVM, is adopted. Following the suggestion of the LIBSVM authors [30], SVM parameters are searched in terms of an exponentially growing scale. We also take into account the number of input values of the time series ( $I$ ) used to train

the SVM. Therefore, we adopt a direct encoding scheme, using a numeric representation with 8 genes, according to the chromosome  $g_1 g_2 g_3 g_4 g_5 g_6 g_7 g_8$ , such that:

$$\begin{aligned} I &= \text{round}(\alpha \cdot n \cdot \frac{10g_1 + g_2 + 1}{100}) \\ g_1, g_2 &\in \{0, 1, \dots, 9\} \\ \gamma &= 2^{(g_3 + \frac{g_4}{10}) - 5} \\ g_3, g_4 &\in \{-9, -8, \dots, 9\} \\ C &= 2^{(g_5 + \frac{g_6}{10}) + 5} \\ g_5, g_6 &\in \{-9, -8, \dots, 9\} \\ \varepsilon &= 2^{(g_7 + \frac{g_8}{10}) - 8} \\ g_7, g_8 &\in \{-9, -8, \dots, 9\} \end{aligned} \quad (3)$$

where  $g_i$  denotes the  $i$ -th gene of the chromosome. The input search range includes only integer numbers, due to the use of round function, and depends on  $n$ , the length of the time series (in-samples or training data), scaled by a constant  $\alpha$  factor.

The evolutionary process consists of the following steps:

- 1) First, a randomly generated population is obtained.
- 2) The phenotypes (SVM model) and fitness value of each individual of the actual generation are obtained. This includes the steps: a) The phenotype of an individual of the actual generation is first obtained (using LIBSVM [30]); b) For each model, training and validation subsets are obtained from time series data depending on the number of inputs nodes; c) The model is fitted using the Sequential Minimal Optimization algorithm, as implemented in LIBSVM. The fitness for each individual is given by the mean squared validation error during the learning process. The aim is to reduce extreme errors (e.g., outliers) that can highly affect multi-step ahead forecasts.
- 3) Once the fitness values for the whole population have been already obtained (with no dependencies between variables), operators (selection, estimation of the empirical probability distribution and sampling solutions) are applied in order to generate the population of the next generation [29].
- 4) Steps 2 and 3 are iteratively executed until a maximum number of generations is reached.

## D. Fuzzy Inductive Reasoning Methodology

The conceptualization of the FIR methodology arises of the General System Problem Solving (GPS) approach proposed by Klir [32]. This methodology of modeling and simulation is able to obtain good qualitative relations between the variables that compose the system and to infer future behavior of that system. It has the ability to describe systems that cannot easily be described by classical mathematics or statistics, i.e. systems for which the underlying physical laws are not well understood. The Fuzzy Inductive Reasoning (FIR) methodology, offers a model-based approach to predicting either univariate or multivariate time series ([33], [34]). A FIR model is a qualitative, non-parametric, shallow model based on fuzzy logic. Visual-FIR is a tool based on the Fuzzy Inductive Reasoning (FIR) methodology (runs under Matlab environment), that offers a new perspective to the modeling and simulation of complex systems. Visual-FIR designs process blocks that allow the treatment of the model identification and prediction phases

of FIR methodology in a compact, efficient and user friendly manner [35]. The FIR model consists of its structure (relevant variables) and a set of input/output relations (history behavior) that are defined as if-then rules. Feature selection in FIR is based on the maximization of the models' forecasting power quantified by a Shannon entropy-based quality measure. The Shannon entropy measure is used to determine the uncertainty associated with forecasting a particular output state given any legal input state. The overall entropy of the FIR model structure studied,  $H_s$ , is computed as described in equation 4.

$$H_s = - \sum_{\forall_i} p(i) \cdot H_i, \quad (4)$$

where  $p(i)$  is the probability of that input state to occur and  $H_i$  is the Shannon entropy relative to the  $i^{th}$  input state. A normalized overall entropy  $H_n$  is defined in equation 5.

$$H_n = 1 - \frac{H_s}{H_{max}} \quad (5)$$

$H_n$  is obviously a real-valued number in the range between 0.0 and 1.0, where higher values indicate an improved forecasting power. The model structure with highest  $H_n$  value generates forecasts with the smallest amount of uncertainty.

Once the most relevant variables are identified, they are used to derive the set of input/output relations from the training data set, defined as a set of if-then rules. This set of rules contain the behaviour of the system. Using the five-nearest-neighbors (5NN) fuzzy inferencing algorithm [36] the five rules with the smallest distance measure are selected and a distance-weighted average of their fuzzy membership functions is computed and used to forecast the fuzzy membership function of the current state, as described in equation 6.

$$Memb_{out_{new}} = \sum_{j=1}^5 w_{rel_j} \cdot Memb_{out_j} \quad (6)$$

The weights are based on the distances and are numbers between 0.0 and 1.0. Their sum is always equal to 1.0. It is therefore possible to interpret the relative weights as percentages.

The FIR methodology is, therefore, a modeling and simulation tool that is able to infer the model of the system under study very quickly and is a good option for real time forecasting. Moreover, it is able to deal with missing data as has been already proved in a large number of applications ([33], [34]).

### III. EXPERIMENTS AND RESULTS

#### A. Time Series and Evaluation

The data used for the experiments covers three different locations of electric consumptions: the whole campus of the Technical University of Catalonia (UPC), one office in Barcelona of  $200m^2$  and a second office of  $50m^2$ , also located in Barcelona. Data is recorded every hour; therefore there are

TABLE I. DAYS SELECTED FOR THE ELECTRIC CONSUMPTION FORECASTING (MODELS TEST DATA SETS)

	Day 1	Day 2	Day 3	Day 4
UPC	01/10/11	18/01/12	25/04/12	20/07/12
Office 1	NA	15/02/12	12/04/12	04/07/12
Office 2	NA	23/02/12	16/04/12	12/07/12

24 recordings per day and per location. With the data used for this paper, it is pretended to cover three different locations, with different features in terms of climatology, consumption patterns and schedules. The first set of data from the UPC comprises a whole year of electricity consumptions and the models generated with the different methodologies are evaluated in four different days separated by approximately four months, thus, models are tested in the four different seasons. The second and third data sets from offices 1 and 2, comprise seven and six months of electric consumptions records, respectively. For this reason it is considered appropriate to choose three test days, in three representative seasons, instead of four. Table I shows the days that have been chosen as test data for the electricity load forecasting.

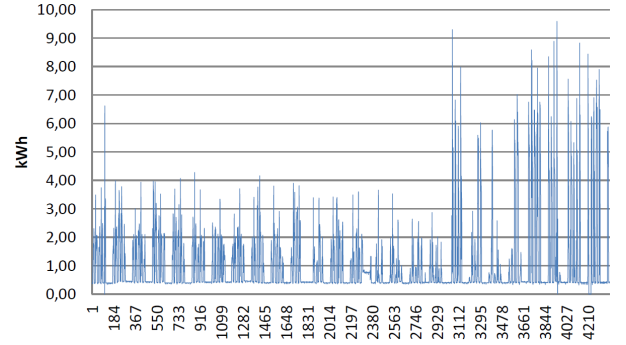


Fig. 1. Electricity consumptions in Office 1 used in the current study (from 23/01/2012 to 23/07/2012 ).

In Figure 1 the whole data set available for Office 1 that is used in the current study is presented. There is not a clear pattern in the consumption, with a lot of peaks and unexpected changes. These issues are probably due to the characteristics of Office 1, with approximately 12 people working inside, where electricity consumptions are low and turning on a single device can produce unexpected peaks, being, therefore, quite complex data.

The Normalized Mean Squared Error ( $NMSE$ ), described in equation 7, is used as the error measure to evaluate the forecasted results.

$$NMSE = \frac{1}{N} \sum_{i=1}^N [(y_{r_i}(t) - y_{f_i}(t))]^2 / var(y_r) \quad (7)$$

where  $y_{r_i}$  and  $y_{f_i}$  are real and forecast electric consumptions, respectively and  $var(y_t)$  is the variance of the real electric consumptions in the test data set.

#### B. Models

In this study the modeling scheme shown in Figure 2 has been used to obtain the different models for each data set

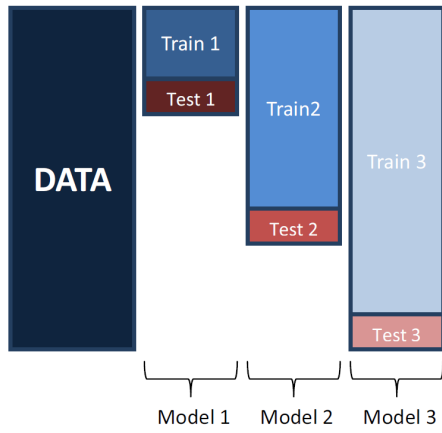


Fig. 2. Scheme of the training and test data distribution for each model. For the UPC data set, four models are trained and tested, instead of the three shown in the figure. In this case Model 1 has been trained using 2184 data points, Model 2 was trained with 4800, Model 3 with 7138 and for Model 4, 9202 data points were available for training. In the case of offices 1 and 2 three models have been derived. For Office 1, 528, 1896 and 3888 data points were available to train Models 1, 2 and 3, respectively. For Office 2, 528, 1896 and 3888 data points were available to train Models 1, 2 and 3, respectively.

available. For Model 1, that is tested by means of Test 1 data set (corresponding to Day 1 in Table I), only the Train 1 data set is used to derive the model. For obtaining Model 2, all the previous accumulated data (Train 2 data set) is used, and this model is then used to forecast the Test 2 data set (corresponding to Day 2 in Table I), and so on. It has been considered that building up one model that predicts electricity consumptions one day ahead in a specific season could give us enough information about the prediction capacity of the different models.

It is important to mention that the ARNN, ESVM and FIR methodologies have used only the univariate time series of the electricity consumption to obtain the models for the UPC and both offices presented in this study. However, the RF approach, due to its own nature, needs more information in order to derive models able to forecast electricity consumptions of future days. To this end, new indicators are also used in this case. In particular the time of day, two Booleans indicating if it is a working day and if it is a working hour, the day of the week, the electricity consumption of the previous day and week at the same hour and, finally, the electricity consumption average of the previous 24 hours are included in the train data sets used only by the RF technique. It is clear that, due to this reason, a fair comparison can not be made between the prediction results obtained by ARNN, ESVM and FIR models and those obtained by RF models. However, we have a special interest to work with random forest and fuzzy inductive reasoning techniques, because there are few works that address these concrete methodologies in the area of electric load forecasting, as can be seen in the introduction of this paper.

The data studied do not present a clear trend and no deseasonal preprocessing has been applied because we want to study the capacity of the different methodologies to obtain generic models. In the near future we plan to study the modeling of each season separately versus the deseasonalization of the data. In case of RF, due to its special characteristics, it was necessary

to prepare additional input data as mentioned above.

Concerning the lags of the electric load series for the ARNN, ESVM and FIR models, it has been considered the previous 72 hours, which is the minimum past values to observe a change between weekday and weekend consumptions. The selection of the number of lags is an important issue that can affect those methods that are more sensitive to the course of dimensionality. In this research we have selected three days, i.e. 72 hours, instead of a full week because we think that the models proposed are able to capture the consumption during the week in a reasonable way, being the transitions from Friday to Saturday (working/non working day) and Sunday to Monday (non working/working day) the more difficult to be captured. The idea is to find the simplest possible models capable of predicting electric load reasonably well. Remember that the final goal is the integration of these models in a smart grid.

The RF model for the data studied in this research starts with the previous identification of the most important input features through a simple correlation. The correlation analysis performed for the data at hand shows that the most deterministic input variables are time of day, if it is a working day, if it is a working hour and day of the week. Moreover, it has been selected 20 as the minimum number of observations per tree leaf. It is used in the tree building process, by ignoring any split that leads to nodes containing fewer than this number of training set instances. By imposing this limit helps to reduce variance in predictions at leaves. Finally, number of trees has been established in 20, a higher number does not improve experiments performance.

In relation to the number of hidden layers in ARNN model, after several experiments, it has been concluded that no weighty improvement is achieved with more than one hidden layer, while time to build up the model increases significantly; therefore, this parameter will remain as one. Additionally, the suitable number of neurons in the hidden layer is selected by calculating the NMSE for each model from 1 to 30 neurons, and selecting the model and number of neurons with the best performance. The search has been considered up to 30 neurons, since with more neurons the cost of time increases significantly and the performance does not improve.

Based on a previous research using evolving artificial neural network for TSF [29], we set the ESVM models for the task at hand with an  $\alpha$  value of 0.45. Thus, the ranges for the search space are:  $I$ , depends on the series length;  $\gamma \in 2^{[-14.9, 4.9]}$ ;  $C \in 2^{[-4.9, 14.9]}$ ; and  $\varepsilon \in 2^{[-17.9, 1.9]}$ .

The initial random population is composed of 50 individuals, i.e. the value of 50 is decided from prior experimentation. The parameters of the estimation of distribution algorithm (EDA) are set to the values proposed in [29]. Since the EDA works as a second order optimization procedure, the tuning of its internal parameters is not a critical issue.

To obtain FIR models for the electrical consumptions data addressed in this research, causal correlations up to 72 hours in the past have been studied. Remember that FIR methodology perform feature selection using a Shannon entropy-based quality measure. The feature selection process has found that almost all the models need only information of the last 24 hours to predict the consumption of the next day. With respect the fuzzyfication parameters used, the electrical consumption

TABLE II. *NMSE* FOR EACH TEST DAY OF THE THREE DIFFERENT LOCATIONS (UPC, OFFICE 1 AND OFFICE 2), OBTAINED BY MEANS OF THE RANDOM FOREST (RF), AUTOREGRESSIVE NEURAL NETWORK (ARNN), EVOLUTIONARY SUPPORT VECTOR MACHINE (ESVM) AND FUZZY INDUCTIVE REASONING (FIR) METHODOLOGIES

		Day 1 (Fall)	Day 2 (Winter)	Day 3 (Spring)	Day 4 (Summer)	Average
RF	UPC	5.4591	0.0091	0.9945	0.0857	1.6371
	Office 1	-	0.2339	0.565	0.4154	0.4048
	Office 2	-	0.2016	1.0014	0.137	0.4467
ARNN	UPC	175.237	0.0269	0.0756	0.0801	43.8549
	Office 1	-	1.2327	1.0033	1.1359	1.124
	Office 2	-	0.6065	1.0401	0.6475	0.6270
ESVM	UPC	94.7816	0.5708	0.1328	0.1117	23.8992
	Office 1	-	0.2152	0.3238	0.3485	0.2958
	Office 2	-	0.1446	1.2715	0.8101	0.7421
FIR	UPC	0.6769	0.0073	0.0058	0.0670	0.1893
	Office 1	-	0.1647	0.2698	0.2024	0.2123
	Office 2	-	0.0916	0.9336	0.1383	0.3878

variable has been discretized into three classes following in all the identified models the equal frequency partition algorithm to obtain the shape of the membership function for each class.

### C. Results

Table II presents the *NMSE* obtained for each methodology and data set. In this manner, representative results for the whole year are provided. The experiments results show a better performance of the FIR models, followed by RF, ESVM and ARNN models.

Figures 3 to 6 show the forecasting of all the methodologies studied for each test day for UPC data set. Figures 7 to 9 present the forecasting results for each test day for Office 1 data set. Finally, figures 10 to 12 show the prediction results for each test day for Office 2 data set. As can be seen from the figures all the approaches, except the ARNN, forecast quite well the minimum electricity consumptions. However, in some cases the picks are not reached (see figures 9 and 11). Both RF and FIR can deal with pronounced changes, while the SVM performs smooth curves.

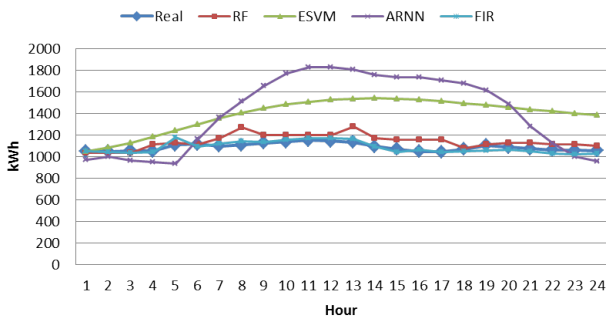


Fig. 3. Prediction for day 01/10/2011 (test data set - Day 1) for the UPC with all the methodologies.

To sum up, results obtained denote a really good performance of FIR, followed by RF that is in advantage since it uses more input variables than those models that only use implicit information of the time series. On the other hand, the results of the ESVM models do not show a good accuracy according to the cost of the models construction, pointing out that it may not be the best solution for this data. ARNN performs

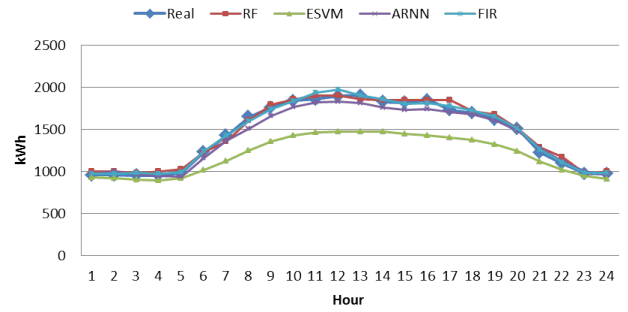


Fig. 4. Prediction for day 18/01/2012 (test data set - Day 2) for the UPC with all the methodologies.

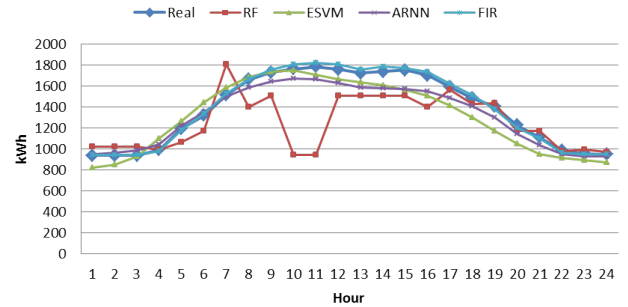


Fig. 5. Prediction for day 25/04/2012 (test data set - Day 3) for the UPC with all the methodologies.

the worst model but due to the ability of improvement of neural networks and its demonstrated good results with other time series forecasting experiments reported in the literature, cannot be discarded as a model for 1-day-ahead electric load forecasting. The better performance of the FIR models is probably due to the characteristics of FIR methodology that after performing a feature selection holds explicitly the previous experience in the form of pattern rules. This is a different view from other methodologies like NN that the previous experience is held implicitly in a mathematical function. FIR models are synthesised rather than trained, which speeds up the modeling phase in comparison with other inductive modeling techniques, such as NN. Definitely the ESVM is the

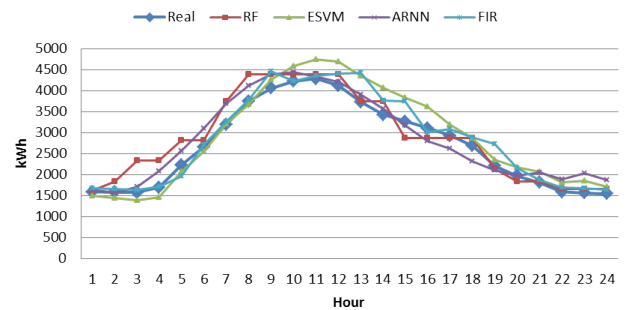


Fig. 6. Prediction for day 20/07/2012 (test data set - Day 4) for the UPC with all the methodologies.

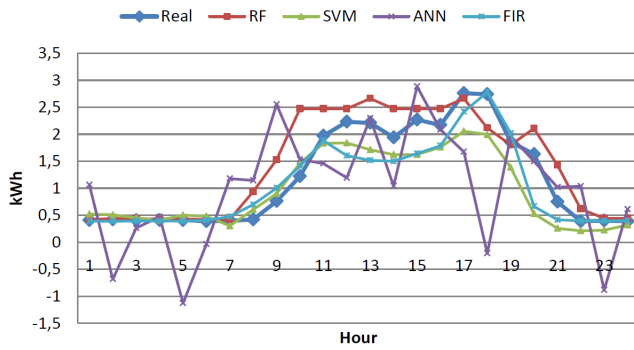


Fig. 7. Prediction for day 15/02/2012 (test data set - Day 1) for the Office 1 with all the methodologies.

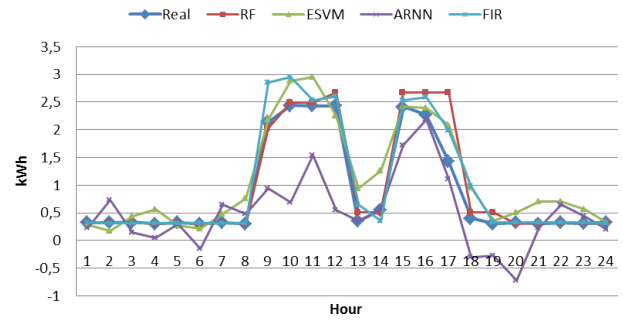


Fig. 10. Prediction for day 23/02/2012 (test data set - Day 1) for the Office 2 with all the methodologies.

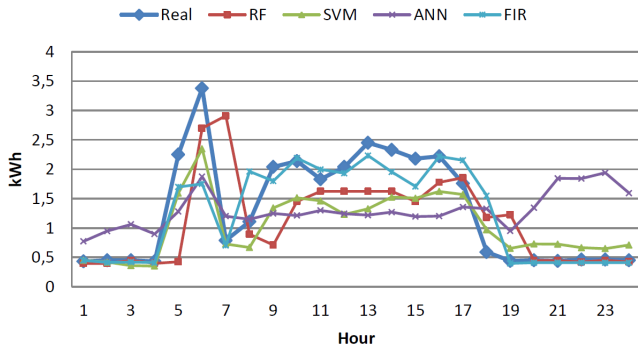


Fig. 8. Prediction for day 12/04/2012 (test data set - Day 2) for the Office 1 with all the methodologies.

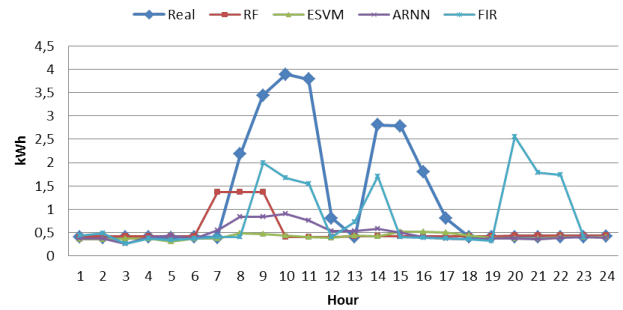


Fig. 11. Prediction for day 16/04/2012 (test data set - Day 2) for the Office 2 with all the methodologies.

method with highest computational cost due to the evolutionary approach used to find the SVM parameters. The ARNN have higher computational complexity than RF and FIR methods, and needs more time to derive the models from the data. However, the three methodologies are able to perform the prediction quickly enough to be used in real time.

Therefore, we plan to study the performance of RF and FIR methodologies in more depth, working with different number of lags and adding the features that RF uses to the FIR models. Other type of neural networks will be also included in the study.

#### IV. CONCLUSIONS AND FUTURE WORK

In this work different AI promising methodologies, i.e. random forest, autoregressive neural networks, evolutionary support vector machines and fuzzy inductive reasoning, are proposed to perform short-term electric load forecasting (24 hours). These approaches could help inside the future Smart Grid framework providing predictions of electricity consumptions in different type of buildings, making possible a better distribution plan.

Three real locations have been used in this research to study the prediction performance of the proposed methods. The first one correspond to the whole campus of the Technical

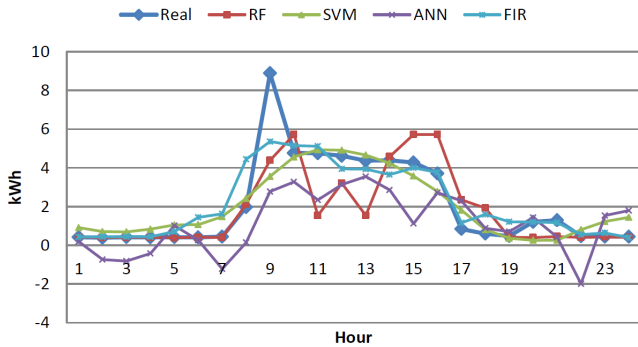


Fig. 9. Prediction for day 04/07/2012 (test data set - Day 3) for the Office 1 with all the methodologies.

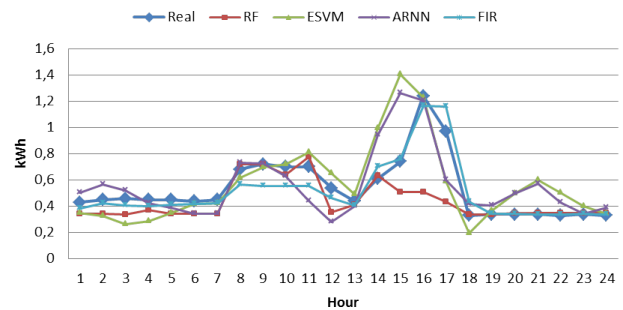


Fig. 12. Prediction for day 12/07/2012 (test data set - Day 3) for the Office 2 with all the methodologies.



University of Catalonia (UPC). The second one is an office of  $200m^2$  and the third location is another office of  $50m^2$ , both located in Barcelona. With the data used for this paper, it is pretended to cover three different locations, with different features in terms of climatology, consumption patterns and schedules. Data is recorded every hour, therefore there are 24 recordings per day and per location.

Based on this study, fuzzy inductive reasoning (FIR) is the methodology that performs a better forecast followed by the random forest (RF), the evolutionary support vector machines and the autoregressive neural networks. From the results it can be concluded that FIR and RF are promising methodologies for the task of predicting electric load and, therefore, should be studied more deeply.

The next step is to include additional input variables, such as the ones used by the RF approach, and to study new methodologies such as Nonlinear Autoregressive Exogenous Neural Network that are characterized by using more external inputs than previous values of the series of returns. A second phase is to incorporate and test these models in real buildings and prove that they are useful for both, consumption and production profiles. Therefore, they are also useful with prosumers, a key participant inside all the parties concerned in the Smart Grid framework.

#### REFERENCES

- [1] Strategic Energy Technologies Information System, SmartGrids Strategic Research Agenda (SRA) for RD&D needs towards 2035, *European Technology Platform SmartGrids*, 2012.
- [2] J. Contreras, R. Espinola, F. J. Nogales, A. J. Conejo, ARIMA Models to Predict Next-Day Electricity Prices, *Power Engineering Review, IEEE*, vol.22, no.9, pp.57, Sept. 2002.
- [3] Metaxiotis, K. and Kagiannas, A. and Askounis, D and Psattas, J., Artificial intelligence in short term electric load forecasting: a state of the art survey for the researcher, *Energy Conversion & Management*, vol. 44, p.p.1525-1534, 2003.
- [4] Kennedy, Peter E., Forecasting with dynamic regression models, *International Journal of Forecasting*, vol. 8 (4), p.p. 647-648, 1992.
- [5] I. Alon, M. Qi, and R. J. Sadowski, Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods, *Journal of Retailing and Consumer Services*, 8, 147156, 2001.
- [6] M.F.I Khamis, Z. Baharudin, N.H. Hamid, Electricity forecasting for small scale power system using artificial neural network, *Power Engineering and Optimization Conference (PEOCO)*, 2011 5th International, pp.54-59, 6-7 June 2011.
- [7] A. Marvuglia, A. Messineo, Using Recurrent Artificial Neural Networks to Forecast Household Electricity Consumption, *Energy Procedia*, vol.14, pp. 45-55, 2012.
- [8] C. Tranchita, A. Torres, Soft computing techniques for short term load forecasting, *Power Systems Conference and Exposition*, 2004. IEEE PES , vol., no., pp. 497- 502 vol.1, 10-13 Oct. 2004.
- [9] Park Y, Park J. An expert system for short-term load forecasting by fuzzy decision, *Proceedings of ESAPS*, p. p. 24450, 1989.
- [10] Ranaweera D, Hubele N, Karady G., Fuzzy logic for short term load forecasting, *Electr Power Energ Syst*, vol. 18(4), p.p. 21522, 1995.
- [11] Kim K, Park J, Hwang K, Kim S., Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems, *IEEE Trans Power Syst*, vol. 10(3), p.p. 60310, 1995.
- [12] Tamimi M, Egbert R. Short term electric load forecasting via fuzzy neural collaboration, *Electric Power Syst Res*, vol. 56, p.p. 2438, 2000.
- [13] R. Caruana, A. Niculescu-Mizil, An Empirical Comparison of Supervised Learning Algorithms, *Proceedings of the 23rd International Conference on Machine Learning (ICML2006)*, pp.161-168, June 2006.
- [14] Ahmed, Nesreen K.; Atiya, Amir F.; Gayar, Neamat el; El-Shishiny, Hisha, An empirical comparison of machine learning models for time series forecasting, *Econometric reviews*, Vol. 29, pp. 594-621, 2010.
- [15] <http://www.forecasters.org/data/m3comp/m3comp.htm>
- [16] Geoffrey K.F. Tso, Kelvin K.W. Yau, Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks, *Energy*, vol. 32, pp. 17611768, 2007.
- [17] Bishop, Christopher M., *Neural Networks for Pattern Recognition*, Oxford University Press, USA, 1996.
- [18] Faria, A., Forecasting: Methods and Applications, *International Journal of Forecasting*, vol. 18(1), p.p. 158-159, 2002.
- [19] Gonzalez, A.M. and Roque, A.M.S. and Garcia-Gonzalez, J., Modeling and forecasting electricity prices with input/output hidden Markov models, *Power Systems, IEEE Transactions on*, vol. 20(1), p.p. 13-24, 2005.
- [20] Criminisi, A. and Shotton, J. and Konukoglu, E., Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning, *Found. Trends. Comput. Graph. Vis.*, vol. 7, p.p. 81-227, 2012.
- [21] Crone, S. F. and Hibon, M. and Nikolopoulos, K., Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction, *International Journal of Forecasting*, vol. 27 (3), p.p. 635-660, 2011.
- [22] Breiman, L., Random Forests, *Mach. Learn.*, vol. 45(1), p.p. 5-32, 2001.
- [23] Breiman, L., Bagging predictors, *Machine Learning*, vol. 24 (2), p.p. 123140, 1996.
- [24] Linkai, L. and Weihang, L. and Renshuo, Z., An empirical study on balance of trade in China based on random forest regression model, *Computer Science and Education (ICCSE), 2010 5th International Conference on*, p.p. 561-564, 2010.
- [25] Yali Li and Shengjin Wang and Xiaoqing Ding, Person-independent head pose estimation based on random forest regression, *Image Processing (ICIP), 2010 17th IEEE International Conference on*, p.p. 1521 -1524, 2010.
- [26] Ferreira De Souza, A. and Daros Freitas, F. and Coelho de Almeida, A. G., High performance prediction of stock returns with VG-RAM weightless neural networks, *IEEE Workshop on High Performance Computational Finance (WHPCF)*, p.p. 1-18, 2010.
- [27] Chow, T.W.S. and Cho, S.Y., *Neural Networks and Computing: Learning Algorithms and Applications*, vol. 7, 2007.
- [28] Rashid, T. and Huang, B. Q. and Kechadi, M-T. and Gleeson, B., Auto-regressive Recurrent Neural Network Approach for Electricity Load Forecasting, *International Journal of Computational Intelligence*, vol. 3(1), p.p. 1-9, 2006.
- [29] Peralta, J. and Gutierrez, G. and Sanchis, A., Time series forecasting by evolving artificial neural networks using genetic algorithms and estimation of distribution algorithms, *Neural Networks (IJCNN), The 2010 International Joint Conference on*, p.p. 1-8, 2010.
- [30] Chang, Chih Chung and Lin, Chih Jen, LIBSVM: A library for support vector machines, *ACM Transactions on Intelligent Systems and Technology*, vol. 2(3), 2011.
- [31] A. Smola and B. Schölkopf, A tutorial on support vector regression, *Statistics and Computing*, p.p. 199-222, vol. 14, 2004.
- [32] Klir, J. and Elias, D., *Architecture of Systems Problem Solving*, 2nd Ed., Plenum Press, NY, USA, 2002.
- [33] Nebot, A. and Mugica, F. and Cellier, F., and Vallverdú, M., Modeling and Simulation of the Central Nervous System Control with Generic Fuzzy Models, *Simulation*, vol. 79(11), 2003.
- [34] Carvajal, R. and Nebot, A., Growth Model for White Shrimp in Semi-intensive Farming using Inductive Reasoning Methodology, *Computers and Electronics in Agriculture*, vol. 19 (11), 1998.
- [35] Escobet, A. and Nebot, A. and Cellier, F.E., Visual-FIR: A tool for model identification and prediction of dynamical complex systems, *Simulation Modelling Practice and Theory*, vol. 16, 2008.
- [36] Cellier, F.E. and Nebot, A. and Mugica, F. and De Albornoz, A., Combined qualitative/quantitative simulation models of continuous-time processes using fuzzy inductive reasoning techniques, *Simulation Modelling Practice and Theory*, vol. 24(1-2), 1995.