Energy Conversion and Management 103 (2015) 1040–1051

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Day-ahead load forecast using random forest and expert input selection

A. Lahouar [a,b,*], J. Ben Hadj Slama [b]

[a] National Engineering School of Tunis, University of Tunis El Manar, BP 37, Le Belvedere, Tunis 1002, Tunisia
[b] Advanced Systems in Electrical Engineering, National Engineering School of Sousse, University of Sousse, BP 264, Erriadh, Sousse 4023, Tunisia

ARTICLE INFO

Article history:
Received 8 April 2015
Accepted 16 July 2015
Available online 28 July 2015

Keywords:
Short term load forecast
Random forest
Expert input selection
Online learning
Variable importance
Prediction of holidays

ABSTRACT

The electrical load forecast is getting more and more important in recent years due to the electricity market deregulation and integration of renewable resources. To overcome the incoming challenges and ensure accurate power prediction for different time horizons, sophisticated intelligent methods are elaborated. Utilization of intelligent forecast algorithms is among main characteristics of smart grids, and is an efficient tool to face uncertainty. Several crucial tasks of power operators such as load dispatch rely on the short term forecast, thus it should be as accurate as possible. To this end, this paper proposes a short term load predictor, able to forecast the next 24 h of load. Using random forest, characterized by immunity to parameter variations and internal cross validation, the model is constructed following an online learning process. The inputs are refined by expert feature selection using a set of if–then rules, in order to include the own user specifications about the country weather or market, and to generalize the forecast ability. The proposed approach is tested through a real historical set from the Tunisian Power Company, and the simulation shows accurate and satisfactory results for one day in advance, with an average error exceeding rarely 2.3%. The model is validated for regular working days and weekends, and special attention is paid to moving holidays, following non Gregorian calendar.

1. Introduction

Electrical load forecast is a challenging task for power operators. It is an old research theme that followed the evolution of power installations and computational techniques. Also, it is getting more important by the beginning of the 21th century due to the emergence of renewable energy resources and smart grids. Load forecast means the prediction of the future evolution of the electric load signal of an individual apartment, a local grid, a region or even a whole country. This prediction is performed for a period of time called forecast horizon through one or several time steps. An accurate power prediction of one or several hours ahead is very important to load dispatch, unit commitment and energy exchange decisions. Predicting load for larger time horizon is also useful for maintenance planning and energy management policies. Increasing the forecast error of 1% may entail a spectacular increase in operational costs [1], thus the tiniest improvement in accuracy is interesting. The error may be an underestimation or an overestimation of power, and both entail difficulties to balance supply and demand. The concept of smart grid implies utilization of intelligent computing techniques, including power forecast, to manage supply, in order to match the demand in real time. The supply management is closely related to spinning reserve, which is the total synchronized capacity, minus losses and load. Then, forecasting load means predicting the spinning reserve, which is very important in cases of sudden huge demand, outage or failure of some generators. When the forecast is accurate, the spinning reserve becomes ready to offset rapidly any deficiency. For longer horizons, the prediction of the load profile determines the amount of capacity to add to the overall network, in order to prevent any contingency.

Load forecast becomes harder than before due to two main reasons. First, privatization and deregulation of electricity market in many countries mean that energy consumers are free to choose their provider among several operators. Second, high penetration of intermittent resources in the grid, namely wind and solar energy, increases the degree of uncertainty due to their non-regular behavior. The deregulation of the market entails varying electricity price, which pushes customers to consume when the energy cost is low, and therefore new forecasting schemes are required [2]. In this case, price forecast is performed along with load forecast [3]. Having sparse customers is also a consequence of deregulation. All these aspects may lead to load curves that are non-smooth and poorly correlated with weather variables. This is not the case of the Tunisian electricity market, subject of

* Corresponding author at: National Engineering School of Tunis, University of Tunis El Manar, BP 37, Le Belvedere, Tunis 1002, Tunisia.

http://dx.doi.org/10.1016/j.enconman.2015.07.041
0196-8904/

this research, where privatization is still limited. In 2013, only 3046 GW h was generated by private companies, while 13,947 GW h was produced by the government corporation. The price is therefore fixed and the deregulation problem does not arise. However, intermittent sources are integrated in the grid by the presence of two wind farms. The wind power generation was 357.8 GW h in 2013, which means 2.6% of the total production. Despite this tiny contribution, intermittent energy may affect the grid even with low degree of penetration [4]. To this end, prediction of the wind generation should be carried out along with load forecast, and this was subject of a previous work for the Tunisian wind farms [5]. These two types of forecast, when coupled together, ensure better management of the residual load which must be covered by classical power plants. The small area of the country and the installed power limited to 4425 MW make it possible to forecast the entire country load demand as a whole. It is also possible to use the weather measurements of one station located in Tunis, the capital and the most consuming city, instead of considering many scattered weather stations [6,7].

In order to face the above mentioned challenges, this paper proposes a short term prediction model using the random forest technique. This model is designed for day-ahead prediction by a step of one hour, with respect to the specifications of the Tunisian power installation; such as small area, warm weather, no deregulation and presence of intermittent energy. The main contribution of this paper is to demonstrate the flexibility of random forests when associated with expert selection, to handle any load profile, and in particular to fit with complex customers behavior. The proposed approach shows high accuracy and effectiveness in the four seasons and for particular days, such as weekends and holidays, whether they are moving or not. The remainder of the paper is organized as follows: Section 2 gives a bibliographic overview of existing forecast methods, Section 3 details the necessary mathematical development, Section 4 analyzes the load profile and clarifies the prediction strategy, results and comparisons are given in Section 5 and finally, Section 6 concludes the paper.

2. Literature review

2.1. General approaches for short term forecast

Prediction is a regression analysis applied to time series, which means studying the relationship between several variables, namely future and past samples. The load signal is a time series, and a predictor should estimate its future evolution in terms of past samples and eventually some exogenous variables affecting the future load. Based on this concept, several forecast models were elaborated in the literature for the very short term; less than one hour, the short term; one hour to several days, the medium term; one to several months, and the long term; one to several years in advance. It is possible to classify approaches to conventional statistical methods, artificial intelligence methods and hybrid methods.

2.2. Conventional statistical methods

Statistical methods are white-box models where outputs are explicitly related to inputs through mathematical equations. This family of methods includes the simplest linear regression [8], multiple regression [9,10], the well-known Box–Jenkins models; autoregressive moving average (ARMA) [11] and autoregressive integrated moving average (ARIMA) [12,13], exponential smoothing (Holt–Winters) [14] and Kalman filter [15]. These methods are simple to implement and well adapted for the short term, but unable to handle the non-linearity existing in the load series. This fact pushes towards using intelligent methods.

2.3. Artificial intelligence methods

Artificial intelligence methods are black-box models where the internal dynamic is unknown. This family includes three main approaches, namely fuzzy inference (expert) systems (FIS) [16,17], artificial neural networks (ANN) [18–21] and support vector machines (SVM) [22,23,7]. Here, the relationship between inputs and outputs is determined through a set of linguistic rules for fuzzy systems or training process for learning machines (ANN and SVM). Apart from these three main approaches, little attention was paid to random forest (RF) [24,25], which is also a machine learning technique requiring a training phase.

These methods have the great advantage of non linear estimation. A three-layer neural network is able to achieve any accuracy of continuous function mapping [26]. However, ANN has problems of under-fitting and over-fitting, in addition to the local optimal solution. The SVM uses the empirical risk minimization principle, and overcomes the problems caused by ANN [27,26], which makes all its strength. Expert systems have the advantage of giving good interpretability of the system [28], while the main advantage of RF is low sensitivity to parameter values [24]. All those aspects make these methods very powerful, but they have their limitations such as optimal architecture and parameter tuning. These limits are surpassed by hybridization.

2.4. Hybrid methods and metaheuristics

Metaheuristics are stochastic algorithms that try to find a sufficiently good solution to a hard optimization problem, by sampling an objective function. They include evolutionary algorithms such as genetic algorithms (GA) and differential evolution (DE), as well as particle swarm optimization (PSO), ant colony (AC) and simulated annealing (SA). They are commonly used to tune ANN and SVM parameters or for training purpose. In a similar manner, signal processing techniques, especially wavelet transform (WT), are used in hybrid methods.

In general, almost all recent forecast techniques are combinations of the three main approaches and their derivatives, or hybridization with metaheuristics or signal processing techniques. For example, some derivatives of ANN are spiking [1], abductive [29], structural [30], Elman recursive [31] and generalized neural networks [32]. ANN may be hybridized with exponential smoothing [33], grey theory [34], wavelet transform [35,36] or evolutionary algorithms [37]. A class of ANN called self organizing map (SOM) or Kohonen network is also elaborated [28]. This wide range of applications makes the ANN the most commonly used method for load forecast. Likewise, SVM may be used along with ant colony [38], particle swarm [39], adaptive neuro fuzzy inference system (ANFIS) [40] and wavelet transform [41,3]. Fuzzy systems may also be combined with evolutionary algorithms [42–44]. Exponential smoothing, despite not among intelligent methods, may be combined with WT and achieve results as accurate as other hybrid methods [45].

Hybrid methods that combine artificial intelligence and metaheuristic optimization are the most effective and accurate approaches according to many researchers. Nevertheless, limitations always exist, such as consuming time and resources, varying accuracy according to the context and available data.

2.5. Medium and long term approaches

For longer time horizon forecast, literature is scarcer. The majority of researchers utilize hybrid ANN for both medium term [46,47,27,48] and long term [49,50]. Here, the factors driving electricity consumption are significantly different, such as electricity tariff, manufacturing value added, prevailing fuel prices, the number of employees [51].

3. Mathematical preliminaries

3.1. Decision tree

A decision tree, or classification and regression tree (CART) is a statistical model for solving classification and regression problems [52]. It describes the different classes or values that an output may take in terms of a set of input features. In general, a tree is a set of nodes and edges organized according to a hierarchy with no loops. A decision tree is a tree where each split node stores a test function to be applied to the incoming data. The final nodes are called leaves of the tree. Each leaf stores the final test result, or answer. The tree is binary if each internal node has exactly two outgoing edges; called left child and right child. Obviously, prediction is a regression problem. In the case of load forecast, if load is function of maximum temperature, day type and season, it is possible to construct as an example the regression tree of Fig. 1. The decision tree is robust, immune to irrelevant inputs and provides a good interpretability.

The remainder of this section is then valid only for regression problems. Let's call X the input vector containing p features, Y the output scalar or label and S_n a training sample of n observations (X_i, Y_i).

$$S_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}, \;\; X \in \mathbb{R}^p, \;\; Y \in \mathbb{R} \tag{1}$$

The decision tree needs training and testing phases. During training, an algorithm is driving the inputs split at each node, so that parameters of split functions are optimized to fit the training test S_n. The principle is to split recursively the input space X by searching optimal sub-partitions. More precisely, the first step of the CART algorithm consists in splitting at best the root into two different child nodes according to:

$$\{X^j < d\} \cup \{X^j > d\} \tag{2}$$

where $j \in \{1, \ldots, p\}$ and $d \in \mathbb{R}$. To select the best split, the couple (j, d) should minimize a cost function, which is generally the variance of child nodes. The variance of a node t is defined as:

$$Var(t) = \sum_{i:X_i \in t} (Y_i - \bar{Y}_t)^2 \tag{3}$$

where $\bar{Y}_t$ is the mean of observations Y_i present at the node t. Then, the child nodes are also divided according to the same procedure. The expansion of the tree is stopped by a termination criterion. It is common to stop the tree when a maximum number of levels has been reached, or when a node contains less that a defined number of observations. At the end of the training, a prediction function $\hat{h}(X, S_n)$ is constructed over S_n.

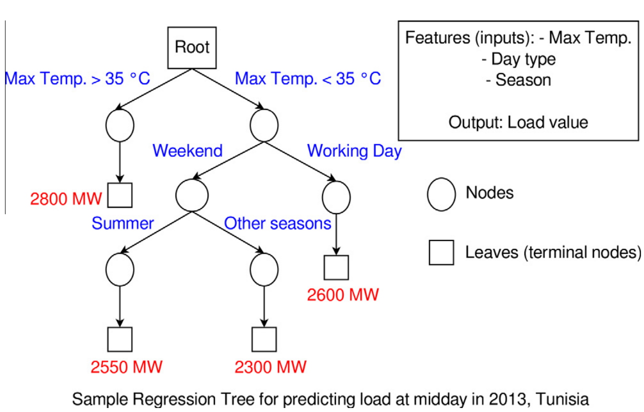

Fig. 1. Regression tree.

The testing process determines an estimation $\widehat{Y}'$ of the output Y' corresponding to any unknown new input X' using the constructed $\hat{h}(X, S_n)$.

$$\widehat{Y}' = \hat{h}(X', S_n) \tag{4}$$

Starting from the root, each split node applies its associated split function to X'. Depending on the result of the binary test the data is sent to the right or left child. This process is repeated until the data point reaches a leaf node.

3.2. Random forest

The random forest (RF) is an ensemble method that combines the prediction of many decision trees [53]. The main principle is called bagging; where a sample of size n from the training set S_n is selected randomly and fitted to a regression tree. This sample is called bootstrap, and is chosen with replacement, i.e. the same observations (X_i, Y_i) may appear many times.

A bootstrap sample is get by selecting randomly n observations with replacement from S_n , each observation has the probability of $1/n$ to be selected. The independent identically distributed random variables Θ_l represent this random selection. The bagging algorithm selects several bootstrap samples $(S_n^{\Theta_1}, \ldots, S_n^{\Theta_q})$, applies the CART algorithm to them to get a collection of q predicting trees $(\hat{h}(X, S_n^{\Theta_1}), \ldots, \hat{h}(X, S_n^{\Theta_q}))$, and then aggregates the output of all these predictors.

In addition to bagging, to split a node, only a predefined number *mtry* of the p features are selected, and the RF algorithm tries to find the best cutting among only the *mtry* selected features. The selection at each node is uniform, thus each feature has the probability of $1/p$ to be selected. The number *mtry* is the same for all trees, and it is recommended to be the square root of the features number p.

$$mtry = \lfloor \sqrt{p} \rfloor \tag{5}$$

The remainder is the same as the CART algorithm, the best (j, d) couple for cutting is set by minimizing a cost function, and the procedure continues until all the trees are fully developed.

The aggregation is made by averaging the output of all trees. Hence, the output estimation $\hat{Y}'$ of any unknown new input X' is:

$$\widehat{Y}' = \frac{1}{q} \sum_{l=1}^{q} \hat{h}(X', S_n^{\Theta_l}) \tag{6}$$

The main advantage of bootstrap aggregation is immunity to noise, since it generates uncorrelated trees through different training sets. A weak predictor (regression tree) may be sensitive to noise, while the average of many uncorrelated trees is not. The selection of a random subset of features at each split has the same purpose of de-correlating trees.

Two main features characterize the random forest: the out-of-bag error (*OOBE*) and the measure of variable importance. The *OOBE* , or generalization error, is a kind of internal cross validation; it is the mean prediction error of first-seen observations, i.e. using only the trees that did not see them while training. More explicitly, for each observation (X_i, Y_i) of S_n, an estimation $\widehat{Y}_i$ of Y_i is made by aggregating only the trees constructed over bootstrap samples not containing (X_i, Y_i). The *OOBE* is very useful to estimate the generalization ability of the constructed model.

$$OOBE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2 \tag{7}$$

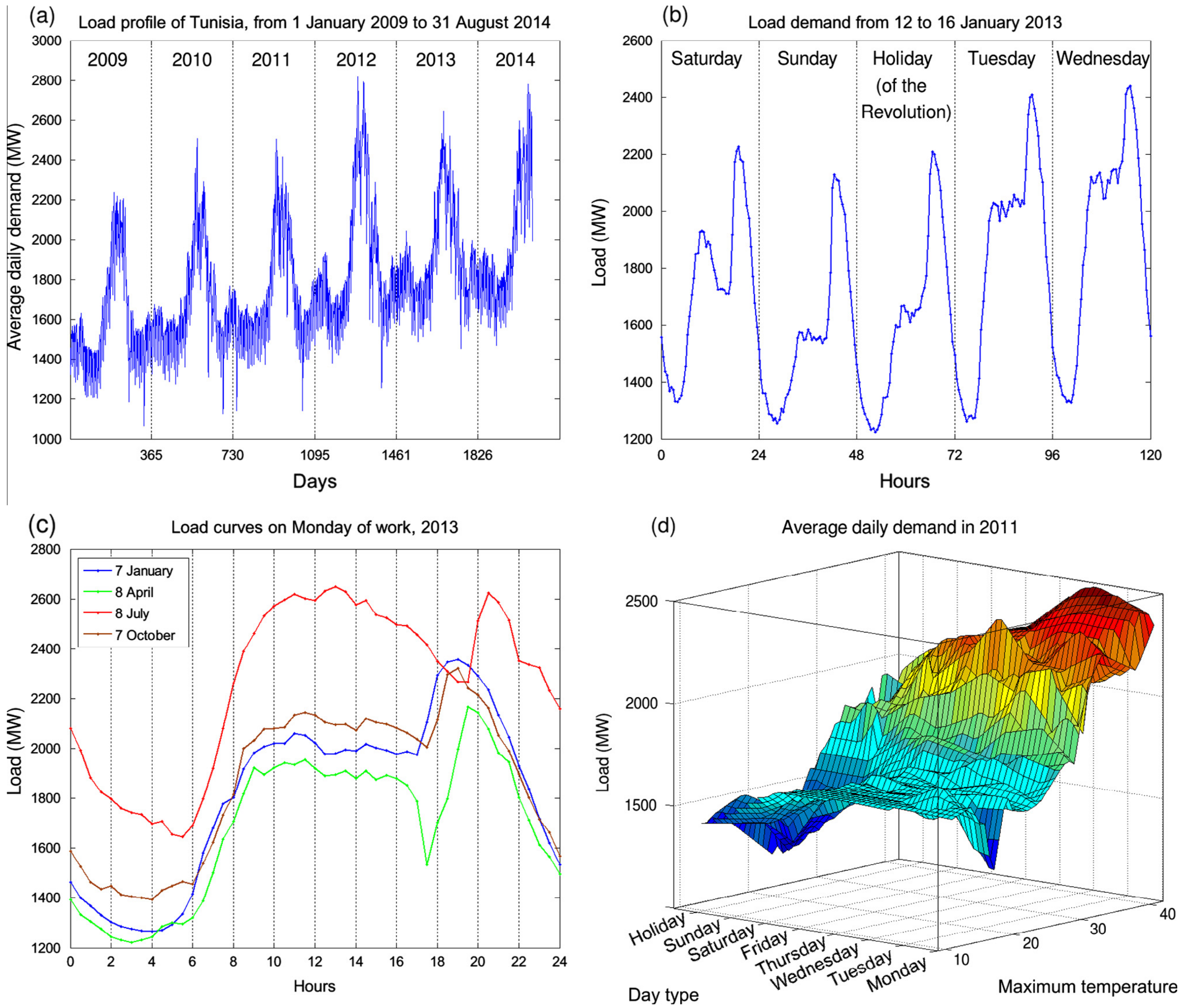

Fig. 2. Load profile.

The measure of variable importance is computed by permuting a feature and averaging the difference in *OOBE* before and after permutation over all trees. Let's consider a bootstrap sample $S_n^{\Theta_l}$ and its associated OBB_l, i.e. the set of observations not included in $S_n^{\Theta_l}$. Let's then compute the $OBBE_l$ of the tree constructed from $S_n^{\Theta_l}$. For a fixed j among the p features, the values of the jth variable are permuted randomly over OBB_l to get a disturbed sample called $\widetilde{OOB}_l$. The new $\widetilde{OOBE}_l$ of the disturbed sample is then computed. These operations are repeated for each bootstrap sample $S_n^{\Theta_l}$. The importance of the jth feature, denoted $VI(X^j)$ is defined by the difference between the mean errors of disturbed and original OBB_l.

$$VI(X^j) = \frac{1}{q}\sum_{l=1}^{q}(\widetilde{OOBE}_l - OBBE_l) \tag{8}$$

Hence, if the random permutations of the jth feature generate an increase of error, the feature is important. The greater is the score $VI(X^j)$, the more important is the feature X^j. In the remainder of the paper, the number of trees q is denoted *ntree*.

4. Methodology

4.1. Load profile analysis

The load signal is a complex time series driven by many factors. It presents some periodic behavior with a general trend. The general trend is observed through the yearly demand increase, obvious in Fig. 2(a). This increase is due to urbanization, population growth, trade of electrical household appliances and so on. Periodic behavior, or seasonality, appears at different levels. At the season level, there is a spectacular rise every summer because of heat waves. At daily basis, the curves are similar for working days, from Monday to Friday. Saturday and Sunday have their own behavior, as shown in Fig. 2(b). Holidays are a bit particular; they are similar to Sundays except for some religious holidays especially in summer. In fact, the curve shape is closely related to Tunisian culture.

By analyzing Fig. 2(c), it is possible to distinguish the rise between 6 and 8 am, which corresponds to the beginning of work. Then, there is a stabilization of demand except for summer, where a flat peak appears due to excessive air conditioning. Finally comes the evening peak, which is lagged according to seasons and day

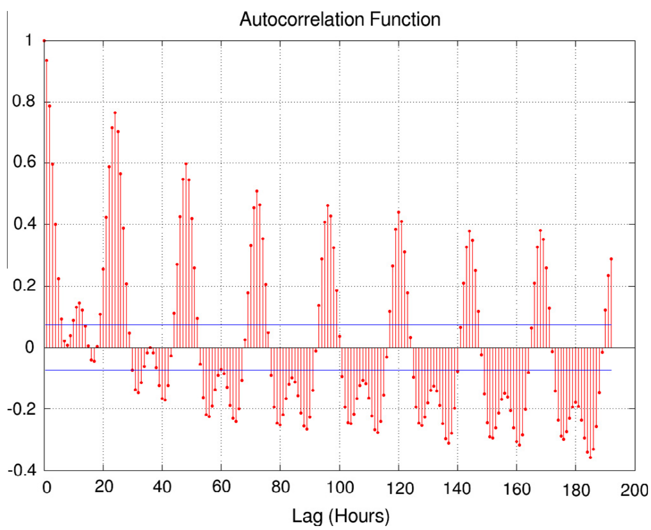

Fig. 3. Autocorrelation plot of the load signal.

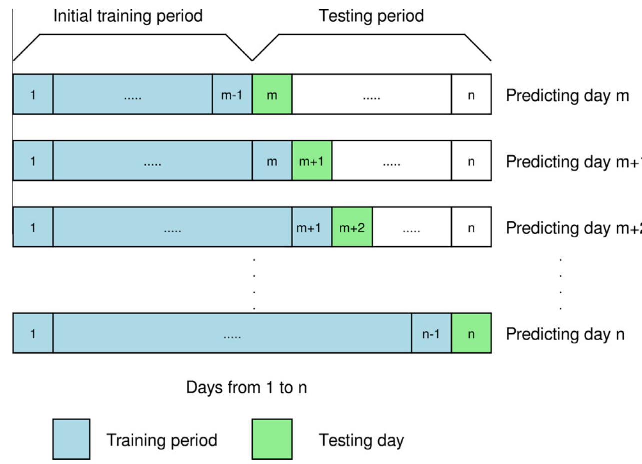

Fig. 4. Learning process.

length. This is the regular behavior of working days for many years. Fig. 2(d) shows the evolution of average daily load in terms of maximum temperature and day type, through extrapolation surface fitting. Clearly, in the case of Tunisia and hot countries, high temperature entails high demand. Also, the day type is of great importance in determining the load behavior.

Apart from mentioned factors, rare events may occur and disturb load, such as political troubles, blackouts or very hot days. In this case, the proposed model may misbehave, and an expert adjust may be needed.

4.2. Choice of model inputs

The autocorrelation plot of Fig. 3 is used to detect the similarity between the signal and its lagged versions, in order to detect the most influencing past samples on the future load. Obviously, the load values at the same hour of the previous days (lag of 24 h) have the greatest similarities. This plot pushes towards constructing a model for each hour of the day, i.e. 24 models, instead of using one model called recursively 24 times, which may lead to growing error. Thus, the load values of the two previous days at the same hour are chosen as inputs for the predictor, as well as morning and evening peaks of the previous day. Concerning external factors, month number, day type, maximum and minimum temperature of the predicted day are chosen. Table 1 summaries all the chosen features, their types, ranges, correlations with the output and importance for the year 2013. The variable importance is computed according to Eq. (8), as explained in Section 3.2.

Some inputs are poorly correlated with the output, or have weak variable importance, such as input 1, so they can be safely removed. However, they are maintained to prove the machine immunity to irrelevant inputs. Also, it is important to mention that temperatures of the predicted day are measured in this work, since both training and testing sets are from historical data. But in real cases, temperatures are also forecasted, and a bad forecast entails misbehavior of the predictor. Some researchers did temperature forecast along with load [1]. But in general, the models that include temperature rely on the weather forecast service.

4.3. Adopted strategy and expert selection

The learning process of the machine is online, as shown in Fig. 4. This means that in each new day of the testing period, the training period is growing, so that the real load values of the previous day will be considered in the learning for the next forecast. This is useful to overcome brutal changes in load especially between seasons. Contrariwise, an offline learning means that training is achieved only one time before testing, thus the training period is fixed.

The flowchart of Fig. 5 details the adopted forecast strategy. Two embedded loops represent hourly and daily forecast routines. Train and test phases are achieved using the random forest original code [54]. Selection of train/test sets is of critical importance, since it allows modifying the endogenous inputs of Table 1 according to the case, through a set of linguistic if–then rules. For example, for

Table 1
Chosen inputs of the model.

Type	Input	Range	Correlation with output	Variable importance ($\times 10^6$)
Exogenous and relative to the predicted day	1. Month number	$1 \cdots 12$	0.1671	0.2186
	2. Day type	1000 (Monday) $\cdots$ 7000 (Sunday) > 8000 (Holiday)	−0.4663	6.7462
	3. Minimum temperature	$2 \cdots 29$ °C	0.5289	0.5306
	4. Maximum temperature	$7 \cdots 43$ °C	0.5453	1.2864
Endogenous and relative to previous days	5. Morning peak of the previous day	$0 \cdots 4000$ MW	0.7858	7.1279
	6. Evening peak of the previous day	$0 \cdots 4000$ MW	0.5446	2.4659
	7. Load before 24 h	$0 \cdots 4000$ MW	0.7741 (at 10 am)	7.8397 (at 10 am)
	8. Load before 48 h	$0 \cdots 4000$ MW	0.6364 (at 10 am)	1.8203 (at 10 am)

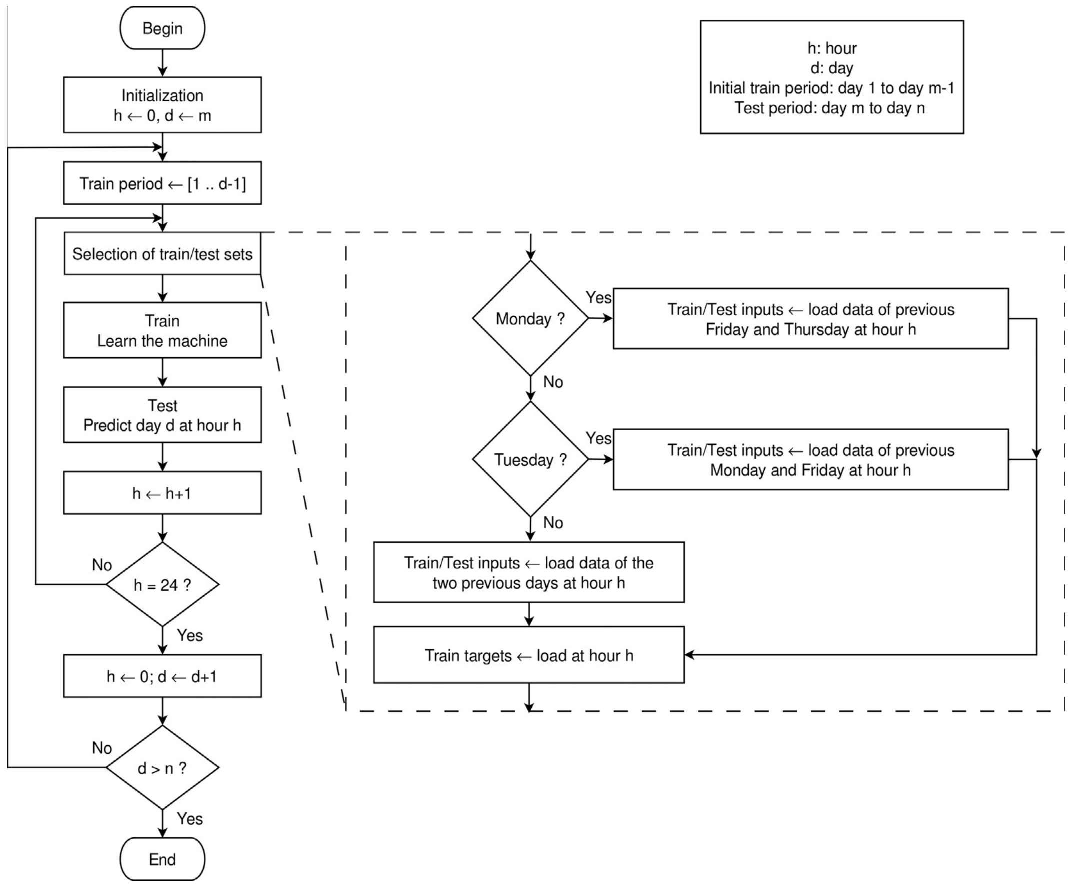

Fig. 5. Flowchart of the prediction strategy.

Mondays, inputs 7 and 8 are from Friday and Thursday, for both train and test sets. Likewise, inputs 5 and 6 are from Friday. Without this adjust, Monday load will be always underestimated, since it will be influenced by Sunday low demand. Same thing is done for Tuesday. However, linking Saturday and Sunday with previous Saturday and Sunday does not result in any improvement, since the load transition from Friday to Sunday is a bit smooth, not like Sunday to Monday transition, which is brutal. These transitions are obvious in Fig. 6(d). In addition, the huge jump in days indexes from Sunday (7000) to Monday (1000) justifies the need for correcting Mondays and Tuesdays, while Saturdays and Sundays do not have this problem. This adjust is like research for similar days or feature selection elaborated by many researchers, but it is not automated, thus there is no distance to minimize. However, it needs an expert opinion. The advantage here is that the user of the model can select its own rules, relative to his country or whatever. It is possible for example to add this rule for warm countries: if temperature >35 °C, then inputs 7 and 8 are taken from a historical very hot day. This custom adjustment allows more flexibility and improves the generalization ability of the proposed model.

5. Case study

5.1. Results

The data used in this research are provided by the Tunisian Company of Electricity and Gas, as half-hourly load demand of the whole country from 1 January 2009 to 31 August 2014. The data preprocessing includes hourly sampling and replacing missing or incorrect values by previous ones. No normalization is applied, so the data remain in raw format and real ranges. Special days such as holidays are intentionally maintained so that the machine can learn their behavior. Some contingencies such as blackouts and political troubles are also maintained for testing, to determine when it is possible to improve prediction and when it is not. Concerning the RF machine configuration, the number of trees *ntree* is the default 500, while the number of variables *mtry* to split at each node is equal to 4. Generally, these parameters do not have remarkable impact on accuracy, and this is the strong point of random forest. It is recommended though that *mtry* follows Eq. (5), while 500 trees are enough to get satisfactory results.

To predict the week from 21 to 27 October 2013, an initial training period is fixed from 1 January to 20 October 2013, and then begins expansion following the strategy explained in Section 4.3. To predict 20–26 January 2014, the training period is extended to 19 January and so on. So the initial training set is always from 1 January 2013 to the day before the beginning of the test, and the testing set is one week. The training set, following the online learning process, is growing every day to take into account the real load of the previous day. The predictor did not see the testing data before test, because the expansion is done after prediction, so the test is performed for first-seen data. Both training and testing sets contain the eight features of Table 1. There is no rule for choosing learning period, but it should be long enough to cover all seasons. However, very long period may degrade the results, not to mention excessive simulation time. In fact, very old load samples are quite different from actual ones, and this is due to long term effects not included in the model. The simulation results show very accurate prediction from Monday to Sunday for the four seasons, given in

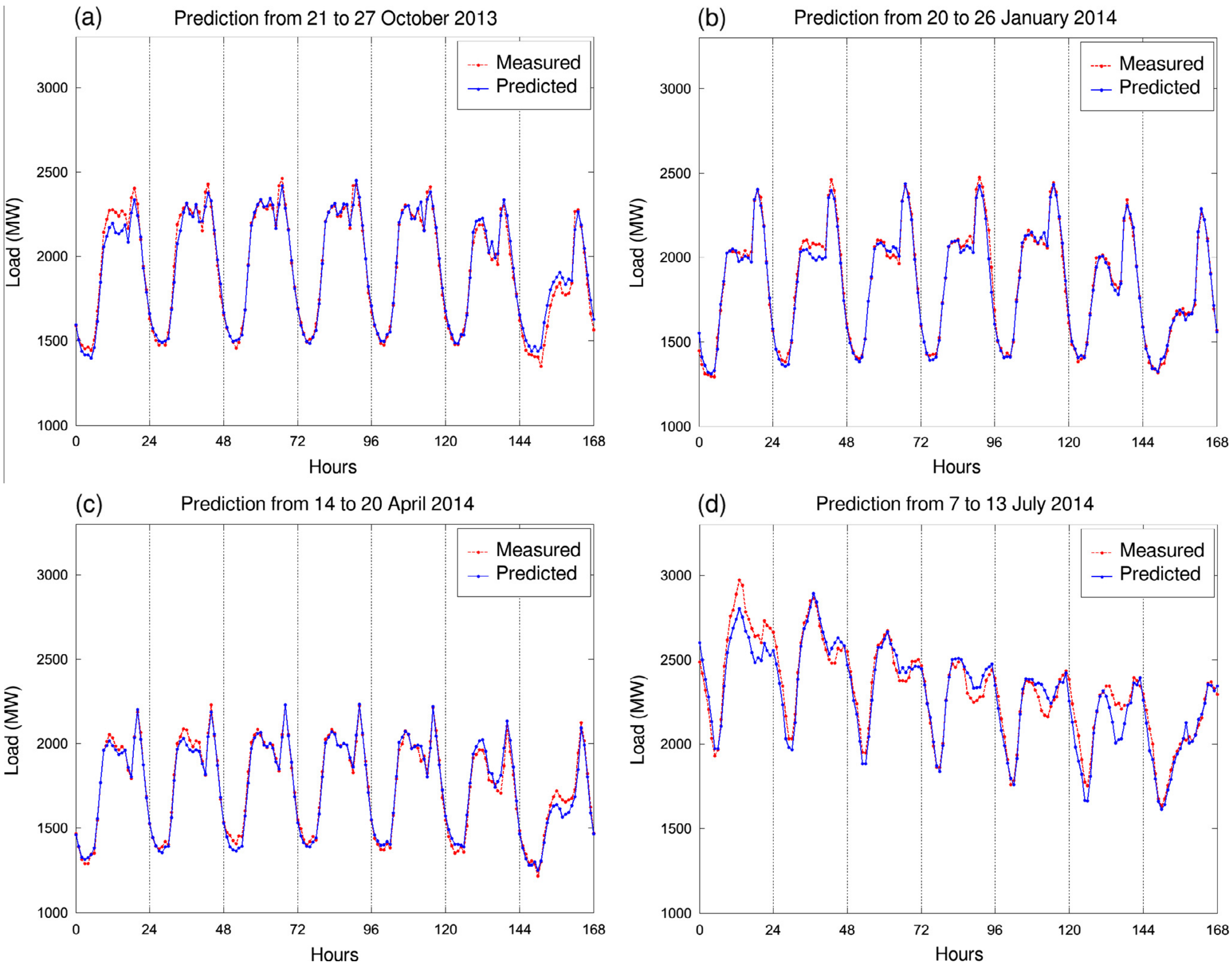

Fig. 6. Forecast results.

Table 2
Forecast errors.

	Fall 2013 21–27 October	Winter 2014 20–26 January	Spring 2014 14–20 April	Summer 2014 7–13 July
Monday	2.9573	1.7651	1.0019	4.2302
Tuesday	2.0358	2.3948	1.5630	2.3781
Wednesday	1.0339	1.1314	1.7281	2.2637
Thursday	1.0388	1.8062	1.1029	1.8342
Friday	1.1931	1.3294	1.3353	2.1097
Saturday	1.9190	1.2111	2.7520	4.2180
Sunday	3.7984	1.2074	2.5791	1.7955
Average	1.9966	1.5493	1.7232	2.6899
Holiday	5 November 2.3962	13/14 January 4.2686/3.7361	1 May 1.9130	25 July 3.6064

Fig. 6. Nevertheless, summer prediction has some difficulties compared to others.

To assess further the quality of prediction, the test is performed for a half-year, from 1 January to 30 June 2014. The average obtained MAPE is 2.24%, excepting holidays which are studied apart. A special forecast is performed for holidays in the four seasons. In this case, prediction is of lower quality, especially when there is succession of holidays such as the case of 13 and 14 January. Religious holidays like El-Moulid (13 January) are moving throughout the year following the lunar calendar, and this fact makes the prediction harder. To quantify the results numerically, the well-known mean absolute percentage error (*MAPE*) is chosen as evaluation criterion.

$$MAPE = \left(\frac{1}{24} \sum_{h=0}^{23} \frac{|L(h) - \widehat{L}(h)|}{L(h)} \right) \times 100 \tag{9}$$

$L(h)$ stands for the actual load at hour h, while $\widehat{L}(h)$ represents the predicted load value. The *MAPE* of the predicted days as well as the average *MAPE* are given in Table 2.

5.2. Improvements

Even if results seem to be accurate, some unexpected events may occur and cause unusual behavior, as illustrated in Fig. 7. For example, the holiday of Aid Al-Fitr (religious holiday) in 2014 was in the summer, and people do not behave the same way if this holiday was in winter, not to mention the spectacular temperature increase in the second day. Other problem, the day before this holiday was Sunday, and its own load demand was much lower than expected, due to people preparations to celebrate; shopping, buying new clothes, gifts, etc. These preparations would not be intense if it was not Sunday. In 31 August 2014, there was a complete blackout in the whole country due to technical problems worsened by the butterfly effect, an event which occurs only one time in several years. In January 2011, the Tunisian Revolution arose, and huge political troubles led to load demand much lower than expected. The major problem encountered in all these cases is scarcity, and the machine is unable to learn rare events. For example, the distribution of daily maximum temperature shows that values superior to 41 or inferior to 10 are very rare. Thus when temperature reaches these values, the predictor does not know how to react.

A possible solution is to correct again the selection of train/test sets in the flowchart. In the case of Aid Al-Fitr of 2014, this rule is added: if it is Aid Al-Fitr and temperature is superior to 35 °C, then input 7 is taken from Aid Al-Fitr of the previous year, i.e. 8 August 2013, as shown in Fig. 8. The conditions are similar; same holiday and it was a very hot day (42 °C). This adjustment results in the improvement visible in Fig. 7(d), where the *MAPE* dropped from 8% to 5%. However, the Sunday prediction remains the same even when linked to the previous year, since the day before the holiday of the previous year was not Sunday. In general, it possible to add these rules: when it is very hot, refer to a historical very hot day, and when it is holiday, refer to the previous holiday of the same type.

5.3. Comparison and discussion

The aim of this section is to compare the proposed model with concurrent machine learning techniques, namely ANN and SVM, and to assess the generalization ability trough tests on other markets. The persistence model (PER), where the prediction is literally equal to the measure of the previous day, is added as a reference. To ensure fair comparison, no optimization process is applied to adjust the parameters of the three machines. But since ANN and SVM are very sensitive to their parameters, their configuration is derived from the default one recommended by the constructor of the code, with slight manual adjustments to get the best of them. The final goal is to compare these three methods without resorting to optimization algorithms. The strategy of Fig. 5 is applied for each of them; where they have to predict the load at hour h while given

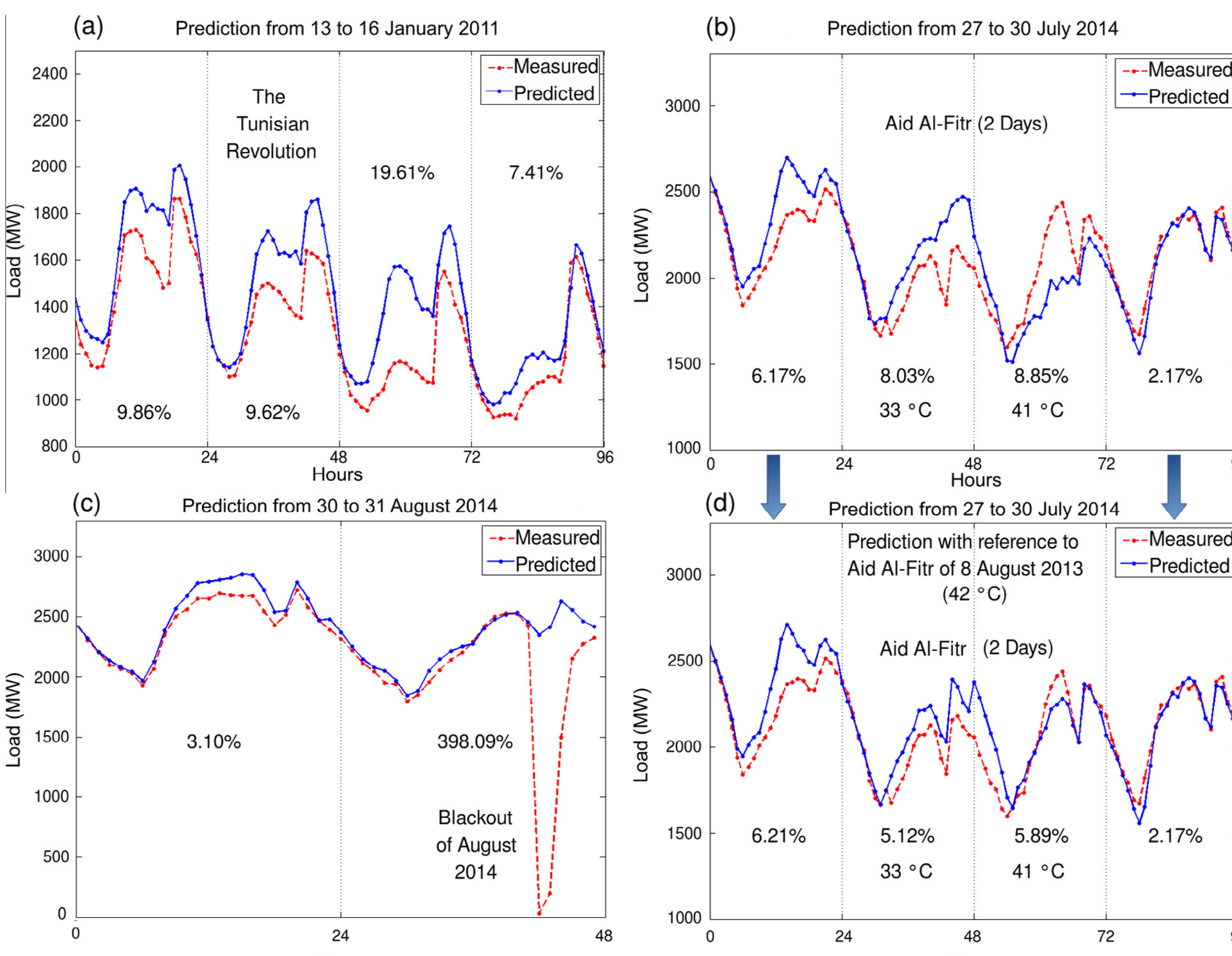

Fig. 7. Bad predictions with daily *MAPE* and possible improvements.

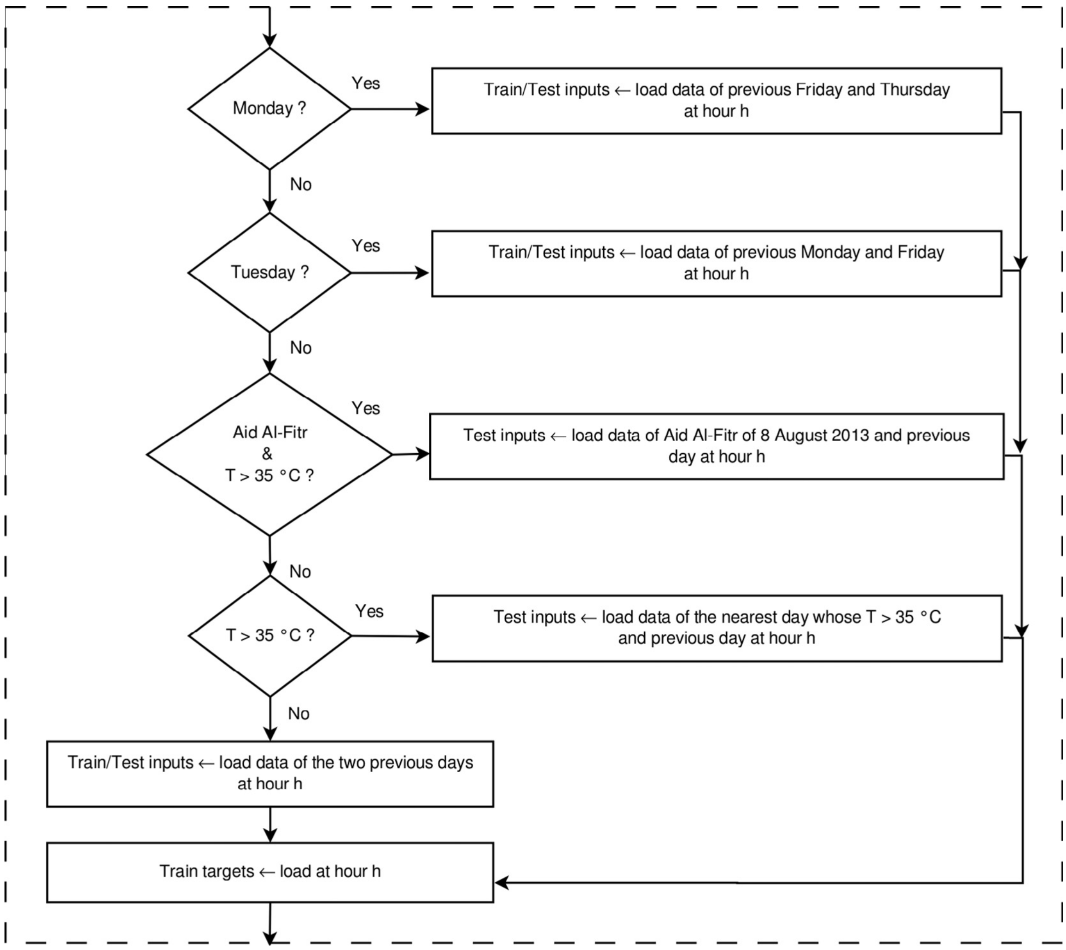

Fig. 8. Refined selection of train/test sets.

the eight features of Table 1 as inputs. The proposed configurations are as follows:

- ANN: 3 layers feed-forward network; input layer with 8 neurons for the eight features, one hidden layer with 20 neurons and one output layer with one neuron. The training algorithm is the Levenberg–Marquardt back-propagation, with the mean squared error as performance function. The activation function of each neuron is the hyperbolic tangent sigmoid given in Eq. (10).

$$tansig(x) = \frac{2}{1+e^{-2x}} - 1 \tag{10}$$

- SVM: It is in fact SVR (support vector regression), type epsilon. The kernel is polynomial of degree 1 as shown in Eq. (11). The kernel parameter γ is equal to 0.0003, r is 0 while the regularization parameter C is 10 using grid search [55]. The termination criterion is 0.001.

$$K(u, v) = \gamma u^T v + r \tag{11}$$

- RF: The same as before, $ntree = 500$ and $mtry = 4$.

The three machines are tested using data from the Tunisian power system and the PJM market (Pennsylvania-New Jersey-Maryland Interconnection) in the USA as view of comparison. The prediction results for 4 weeks are shown in Table 3. The test periods are chosen so that holidays are avoided. The results for the Tunisian system are more accurate when using the RF predictor in most cases. The accuracy of ANN and SVM may be improved by optimization, but this leads to other problems of choosing the appropriate algorithm and its parameters.

The results of PJM market forecasting are less accurate due to three main reasons. First, the PJM is deregulated and presents high volatility. Second, there is no refinement applied to particular days so the rules are generic (only Monday and Tuesday correction). Third, the temperature is measured only in New York, which is supposed to be the most consuming city. To improve the results, the user of the model has to adjust the model according to the JPM market characteristics; by taking the temperature information from many locations according the coverage area, and by adding his own rules to take into account particular days. The proposed model in this paper shows its effectiveness when it is refined according the citizens culture and behavior.

In general, the *MAPE* of day-ahead prediction is between 1% and 3% in the majority of recent researches [56,57,42,1,58,43,59]. Similar results are found in this paper, but using generic RF parameters without any optimization algorithm. The proposed scheme is also suited to RF prediction since more attention is paid to organizing train/test sets with expert selection, instead of tuning the machine parameters. The main advantage of RF over other methods is the few parameters to set. In general, choosing the default values for the two parameters, *ntree* and *mtry* is enough to get the best from this method. However, ANN and SVM accuracy depends enormously on their parameters. Besides, the internal cross validation procedure gives good generalization ability. Compared to ANFIS and expert systems, the main advantage of

Table 3
Comparison between the different methods in terms of MAPE.

	Tunisia power system				PJM market			
	PER	RF	ANN	SVM	PER	RF	ANN	SVM
Fall	*21–27 October 2013*				*21–27 October 2013*			
Monday	10.2473	2.9573	5.9970	1.1842	11.6758	2.8623	3.9795	1.8143
Tuesday	1.4937	2.0358	3.4987	1.0261	1.2088	0.9653	2.9234	2.7147
Wednesday	1.1342	1.0339	2.2240	1.0531	3.8949	4.2642	5.0863	2.4308
Thursday	1.3563	1.0388	2.5128	0.9776	2.4319	2.9714	3.5950	3.0299
Friday	1.0737	1.1931	1.8585	1.7933	2.0794	1.9389	4.6577	2.6091
Saturday	4.5218	1.9190	3.5893	2.6264	7.7393	2.4416	4.4413	2.6219
Sunday	10.8898	3.7984	3.8425	6.8002	5.5892	3.0055	5.5879	2.1995
Average	4.3881	1.9966	3.3604	2.2087	4.9456	2.6356	4.3244	2.4886
Winter	*20–26 January 2014*				*13–19 January 2014*			
Monday	12.7193	1.7651	1.8818	5.1333	8.9871	1.1432	2.2401	3.6559
Tuesday	3.6586	2.3948	2.4101	1.9979	2.5049	2.6942	5.2133	1.8779
Wednesday	1.7609	1.1314	1.4194	1.5722	5.3825	5.1880	3.0915	5.7661
Thursday	2.1357	1.8062	1.8745	3.0358	2.9150	5.8280	5.3839	4.3063
Friday	2.0410	1.3294	1.9987	2.8130	1.4735	3.4029	4.6590	4.8541
Saturday	5.4608	1.2111	1.8029	2.7589	2.8463	3.4065	5.0160	3.7134
Sunday	9.9884	1.2074	2.4129	5.2611	5.8810	1.5673	2.4824	2.0096
Average	5.3949	1.5493	1.9715	3.2246	4.2843	3.3186	4.0123	3.7405
Spring	*14–20 April 2014*				*7–13 April 2014*			
Monday	10.7557	1.0019	1.9814	4.5069	10.5069	1.9821	2.8370	1.4882
Tuesday	2.7122	1.5630	2.1754	1.1834	2.6453	2.5960	4.3617	3.0996
Wednesday	1.3405	1.7281	2.4537	1.3258	2.1431	3.2794	3.7305	2.7160
Thursday	0.8667	1.1029	1.3012	1.7798	1.9861	2.8144	2.3001	1.6258
Friday	1.8373	1.3353	1.9889	1.9818	2.8596	2.1364	5.1039	3.1789
Saturday	4.6041	2.7520	4.9910	2.6412	10.6370	1.4290	2.3294	6.2736
Sunday	8.0959	2.5791	2.5934	4.3904	4.1369	4.3520	4.6895	3.9764
Average	4.3161	1.7232	2.4979	2.5442	4.9879	2.6556	3.6217	3.1941
Summer	*7–13 July 2014*				*21–27 July 2014*			
Monday	13.6542	4.2302	3.3521	2.2994	15.5319	6.7983	4.1693	5.1931
Tuesday	4.1347	2.3781	2.4733	3.6409	7.3031	5.8869	3.8159	4.5301
Wednesday	5.1276	2.2637	2.6060	2.8871	3.3315	1.7904	1.5886	2.1203
Thursday	4.9159	1.8342	3.0465	2.1154	12.2722	7.5953	5.4861	9.3693
Friday	3.5551	2.1097	2.5774	1.4218	4.8534	1.3925	4.4891	4.6822
Saturday	1.7505	4.2180	3.8909	3.9019	3.6804	2.6274	4.0350	4.0649
Sunday	8.386	1.7955	2.4515	3.6495	1.8542	2.0841	2.3800	2.6695
Average	5.9320	2.6899	2.9140	2.8452	6.9752	4.0250	3.7091	4.6614

RF is the ability to handle huge training sets, which is complicated when using fuzzy if–then rules. The expert feature selection is the main contribution over previous RF forecast works, where the refined train and test sets make the prediction more flexible. Other novelties are also added, such as choosing inputs by analyzing the autocorrelation function and the measure of importance (Table 1) in addition to the online learning process.

6. Conclusion

This paper proposes a model to predict the electrical load demand of day-ahead, by a step of one hour. The random forest, which is among machine learning techniques, is chosen to construct the model. Characterized by immunity to parameters change, internal cross validation and proper variable importance measure, the random forest predictor has 24 versions for the 24 h, and an expert selection is performed to refine inputs. An online process drives the forecast procedure over the entire test period, in order to overcome sudden load variations. After analyzing the market conditions and the load profile, results are given in the four seasons, for regular working days and for holidays. The error, exceeding rarely an average of 2.3%, reflects the accuracy of the approach. The procedure of selecting train and test inputs through if–then rules also proves its efficiency and ability to be adjusted according to the country culture or market specificity. The RF coupled with expert selection is able to capture complex load behavior and solve some special cases that are specific to culture, high temperature, religious events and moving holidays thanks to appropriate choice of inputs.

Acknowledgment

The authors would like to thank the Tunisian Company of Electricity and Gas (STEG) for providing half-hourly load data which were used in this work.

References

[1] Kulkarni S, Simon SP, Sundareswaran K. A spiking neural network (SNN) forecast engine for short-term electrical load forecasting. Appl Soft Comput 2013;13(8):3628–35. http://dx.doi.org/10.1016/j.asoc.2013.04.007. <http://www.sciencedirect.com/science/article/pii/S1568494613001245>.
[2] Javed F, Arshad N, Wallin F, Vassileva I, Dahlquist E. Forecasting for demand response in smart grids: an analysis on use of anthropologic and structural data and short term multiple loads forecasting. Appl Energy 2012;96(0):150–60. http://dx.doi.org/10.1016/j.apenergy.2012.02.027. <http://www.sciencedirect.com/science/article/pii/S0306261912001213> [smart Grids].
[3] 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(0):371–84. http://dx.doi.org/10.1016/j.enconman.2015.02.023. <http://www.sciencedirect.com/science/article/pii/S0196890415001363>.
[4] Kaur A, Pedro HT, Coimbra CF. Impact of onsite solar generation on system load demand forecast. Energy Convers Manage 2013;75(0):701–9. http://dx.doi.org/10.1016/j.enconman.2013.08.026. <http://www.sciencedirect.com/science/article/pii/S0196890413004858>.
[5] Lahouar A, Ben Hadj Slama J. Wind speed and direction prediction for wind farms using support vector regression. In: 2014 5th international renewable energy congress (IREC), 2014. p. 1–6. http://dx.doi.org/10.1109/IREC.2014.6826932; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6826932>.
[6] Fan S, Wu Y-K, Lee W-J, Lee C-Y. Comparative study on load forecasting technologies for different geographical distributed loads. In: Power and energy society general meeting, 2011 IEEE, 2011. p. 1–8. http://dx.doi.org/10.1109/

PES.2011.6039364; <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6039364>.

[7] Fan S, Methaprayoon K, Lee W. Multi-area load forecasting for system with large geographical area. In: Industrial and commercial power systems technical conference, 2008. ICPS 2008. IEEE/IAS, 2008. p. 1–8. http://dx.doi.org/10.1109/ICPS.2008.4606287; <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4606287>.

[8] Goia A, May C, Fusai G. Functional clustering and linear regression for peak load forecasting. Int J Forecast 2010;26(4):700–11. http://dx.doi.org/10.1016/j.ijforecast.2009.05.015. <http://www.sciencedirect.com/science/article/pii/S0169207009000879>.

[9] Ramanathan R, Engle R, Granger CW, Vahid-Araghi F, Brace C. Short-run forecasts of electricity loads and peaks. Int J Forecast 1997;13(2):161–74. http://dx.doi.org/10.1016/S0169-2070(97)00015-0. <http://www.sciencedirect.com/science/article/pii/S0169207097000150>.

[10] Amral N, Ozveren C, King D. Short term load forecasting using multiple linear regression. In: UPEC 2007. 42nd international universities power engineering conference, 2007, 2007. p. 1192–8. http://dx.doi.org/10.1109/UPEC.2007.4469121; <http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4469121>.

[11] Pappas S, Ekonomou L, Karamousantas D, Chatzarakis G, Katsikas S, Liatsis P. Electricity demand loads modeling using autoregressive moving average (ARMA) models. Energy 2008;33(9):1353–60. http://dx.doi.org/10.1016/j.energy.2008.05.008. <http://www.sciencedirect.com/science/article/pii/S0360544208001357>.

[12] Lee C-M, Ko C-N. Short-term load forecasting using lifting scheme and {ARIMA} models. Expert Syst Appl 2011;38(5):5902–11. http://dx.doi.org/10.1016/j.eswa.2010.11.033. <http://www.sciencedirect.com/science/article/pii/S0957417410012704>.

[13] Cho M, Hwang J, Chen C-S. Customer short term load forecasting by using ARIMA transfer function model. In: Proceedings of EMPD '95., 1995 international conference on energy management and power delivery, 1995. vol. 1; 1995. p. 317–22. http://dx.doi.org/10.1109/EMPD.1995.500746; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=500746>.

[14] Taylor JW. An evaluation of methods for very short-term load forecasting using minute-by-minute british data. Int J Forecast 2008;24(4):645–58. http://dx.doi.org/10.1016/j.ijforecast.2008.07.007. <http://www.sciencedirect.com/science/article/pii/S0169207008000708> [energy Forecasting].

[15] Al-Hamadi H, Soliman S. Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model. Electr Pow Syst Res 2004;68(1):47–59. http://dx.doi.org/10.1016/S0378-7796(03)00150-0. <http://www.sciencedirect.com/science/article/pii/S0378779603001500>.

[16] Kiartzis S, Bakirtzis A, Theocharis J, Tsagas G. A fuzzy expert system for peak load forecasting. application to the greek power system. In: MELECON 2000. 10th mediterranean electrotechnical conference, 2000. vol. 3; 2000. p. 1097–100. http://dx.doi.org/10.1109/MELCON.2000.876726; <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=879726>.

[17] Mastorocostas P, Theocharis J, Bakirtzis A. Fuzzy modeling for short term load forecasting using the orthogonal least squares method. IEEE Trans Pow Syst 1999;14(1):29–36. http://dx.doi.org/10.1109/59.744480. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=744480>.

[18] Bakirtzis A, Petridis V, Kiartzis S, Alexiadis M, Maissis A. A neural network short term load forecasting model for the greek power system. IEEE Trans Pow Syst 1996;11(2):858–63. http://dx.doi.org/10.1109/59.496166. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=496166>.

[19] 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(15–16):2128–42. http://dx.doi.org/10.1016/j.enconman.2005.12.008. <http://www.sciencedirect.com/science/article/pii/S0196890405003419>.

[20] Senjyu T, Takara H, Uezato K, Funabashi T. One-hour-ahead load forecasting using neural network. IEEE Trans Pow Syst 2002;17(1):113–8. http://dx.doi.org/10.1109/59.982201. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=982201>.

[21] Kiartzis S, Zoumas C, Theocharis J, Bakirtzis A, Petridis V. Short-term load forecasting in an autonomous power system using artificial neural networks. IEEE Trans Pow Syst 1997;12(4):1591–6. http://dx.doi.org/10.1109/59.627863. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=627863>.

[22] Lin C-T, Chou L-D. A novel economy reflecting short-term load forecasting approach. Energy Convers Manage 2013;65(0):331–42. http://dx.doi.org/10.1016/j.enconman.2012.08.001. <http://www.sciencedirect.com/science/article/pii/S0196890412003160> [global Conference on Renewable energy and Energy Efficiency for Desert Regions 2011].

[23] Wu C-H, Tzeng G-H, Lin R-H. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. Expert Syst Appl 2009;36(3, Part 1):4725–35. http://dx.doi.org/10.1016/j.eswa.2008.06.046. <http://www.sciencedirect.com/science/article/pii/S095741740800300X>.

[24] Dudek G. Short-term load forecasting using random forests. In: Filev D, Jablkowski J, Kacprzyk J, Krawczak M, Popchev I, Rutkowski L, et al., editors. Intelligent Systems'2014. Advances in intelligent systems and computing, vol. 323. Springer International Publishing; 2015. p. 821–8. http://dx.doi.org/10.1007/978-3-319-11310-4_71.

[25] Cheng Y-Y, Chan P, Qiu Z-W. Random forest based ensemble system for short term load forecasting. In: 2012 international conference on machine learning and cybernetics (ICMLC). vol. 1; 2012. p. 52–6. http://dx.doi.org/10.1109/ICMLC.2012.6358885; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6358885>.

[26] Wang J, Li L, Niu D, Tan Z. An annual load forecasting model based on support vector regression with differential evolution algorithm. Appl Energy 2012;94(0):65–70. http://dx.doi.org/10.1016/j.apenergy.2012.01.010. <http://www.sciencedirect.com/science/article/pii/S0306261912000165>.

[27] 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(8):4773–88. http://dx.doi.org/10.1016/j.asoc.2011.07.005. <http://www.sciencedirect.com/science/article/pii/S1568494611002663>.

[28] Che J, Wang J, Wang G. An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting. Energy 2012;37(1):657–64. http://dx.doi.org/10.1016/j.energy.2011.10.034. <http://www.sciencedirect.com/science/article/pii/S0360544211006931> [7th Biennial International Workshop Advances in Energy Studies].

[29] Abdel-Aal R. Improving electric load forecasts using network committees. Electr Pow Syst Res 2005;74(1):83–94. http://dx.doi.org/10.1016/j.epsr.2004.09.007. <http://www.sciencedirect.com/science/article/pii/S0378779604002238>.

[30] Lu Y. Short-term load forecasting method based on structural neural network. In: WCICA 2008. 7th world congress on intelligent control and automation, 2008. 2008. p. 4434–8. http://dx.doi.org/10.1109/WCICA.2008.4593637; <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4593637>.

[31] Zhao F, Su H. Short-term load forecasting using Kalman filter and Elman neural network. In: ICIEA 2007. 2nd IEEE conference onindustrial electronics and applications, 2007, 2007. p. 1043–7. http://dx.doi.org/10.1109/ICIEA.2007.4318567; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=4318567>.

[32] Chaturvedi D, Sinha A, Malik O. Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network. Int J Electr Pow Energy Syst 2015;67(0):230–7. http://dx.doi.org/10.1016/j.ijepes.2014.11.027. <http://www.sciencedirect.com/science/article/pii/S0142061514007091>.

[33] Ramos S, Soares J, Vale Z, Ramos S. Short-term load forecasting based on load profiling. In: 2013 IEEE power and energy society general meeting (PES). 2013. p. 1–5. http://dx.doi.org/10.1109/PESMG.2013.6672439; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6672439>.

[34] Tang N, Zhang D-J. Application of a load forecasting model based on improved grey neural network in the smart grid. Energy Proc 2011;12(0):180–4. http://dx.doi.org/10.1016/j.egypro.2011.10.025. <http://www.sciencedirect.com/science/article/pii/S1876610211018510> [the Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE 2011)].

[35] Kouhi S, Keynia F. A new cascade {NN} based method to short-term load forecast in deregulated electricity market. Energy Convers Manage 2013;71(0):76–83. http://dx.doi.org/10.1016/j.enconman.2013.03.014. <http://www.sciencedirect.com/science/article/pii/S019689041300157X>.

[36] Ghofrani M, Ghayekhloo M, Arabali A, Ghayekhloo A. A hybrid short-term load forecasting with a new input selection framework. Energy 2015;81(0):777–86. http://dx.doi.org/10.1016/j.energy.2015.01.028. <http://www.sciencedirect.com/science/article/pii/S036054421500047X>.

[37] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009;34(1):46–57. http://dx.doi.org/10.1016/j.energy.2008.09.020. <http://www.sciencedirect.com/science/article/pii/S0360544208002454>.

[38] Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. Expert Syst Appl 2010;37(3):2531–9. http://dx.doi.org/10.1016/j.eswa.2009.08.019. <http://www.sciencedirect.com/science/article/pii/S0957417409007799>.

[39] Hong W-C. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. Energy Convers Manage 2009;50(1):105–17. http://dx.doi.org/10.1016/j.enconman.2008.08.031. <http://www.sciencedirect.com/science/article/pii/S019689040800318X>.

[40] Escobar A, Perez L. Application of support vector machines and ANFIS to the short-term load forecasting. In: Transmission and distribution conference and exposition: Latin America, 2008 IEEE/PES. 2008. p. 1–5. http://dx.doi.org/10.1109/TDC-LA.2008.4641839; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=4641839>.

[41] Wu Q. Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing gaussian noises for power load forecasts. Expert Syst Appl 2011;38(1):379–85. http://dx.doi.org/10.1016/j.eswa.2010.06.075. <http://www.sciencedirect.com/science/article/pii/S0957417410005798>.

[42] Yang X, Yuan J, Yuan J, Mao H. An improved {WM} method based on {PSO} for electric load forecasting. Expert Syst Appl 2010;37(12):8036–41. http://dx.doi.org/10.1016/j.eswa.2010.05.085. <http://www.sciencedirect.com/science/article/pii/S0957417410004963>.

[43] Enayatifar R, Sadaei HJ, Abdullah AH, Gani A. Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting. Energy Convers Manage 2013;76(0):1104–16. http://dx.doi.org/10.1016/j.enconman.2013.08.039. <http://www.sciencedirect.com/science/article/pii/S019689041300513X>.

[44] Haque AU, Mandal P, Meng J, Pineda RL. Performance evaluation of different optimization algorithms for power demand forecasting applications in a smart grid environment. Proc Comput Sci 2012;12(0):320–5. http://dx.doi.org/10.1016/j.procs.2012.09.078. <http://www.sciencedirect.com/science/article/pii/S1877050912006692> [complex Adaptive Systems 2012].

[45] Sudheer G, Suseelatha A. Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models. Int J Electr Pow Energy Syst 2015;64(0):340–6. http://dx.doi.org/10.1016/j.ijepes.2014.07.043. <http://www.sciencedirect.com/science/article/pii/S0142061514004785>.

[46] Amjady N, Keynia F. Mid-term load forecasting of power systems by a new prediction method. Energy Convers Manage 2008;49(10):2678–87. http://dx.doi.org/10.1016/j.enconman.2008.04.008. <http://www.sciencedirect.com/science/article/pii/S0196890408001441>.

[47] Gonzlez-Romera E, Jaramillo-Morn M, Carmona-Fernndez D. Monthly electric energy demand forecasting with neural networks and fourier series. Energy Convers Manage 2008;49(11):3135–42. http://dx.doi.org/10.1016/j.enconman.2008.06.004. <http://www.sciencedirect.com/science/article/pii/S0196890408002288> [special Issue 3rd International Conference on Thermal Engineering: Theory and Applications].

[48] Chang P-C, Fan C-Y, Lin J-J. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. Int J Electr Pow Energy Syst 2011;33(1):17–27. http://dx.doi.org/10.1016/j.ijepes.2010.08.008. <http://www.sciencedirect.com/science/article/pii/S0142061510001444>.

[49] Azadeh A, Ghaderi S, Sohrabkhani S. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. Energy Convers Manage 2008;49(8):2272–8. http://dx.doi.org/10.1016/j.enconman.2008.01.035. <http://www.sciencedirect.com/science/article/pii/S0196890408000526>.

[50] Yu S, Wei Y-M, Wang K. Chinas primary energy demands in 2020: predictions from an MPSO–RBF estimation model. Energy Convers Manage 2012;61(0):59–66. http://dx.doi.org/10.1016/j.enconman.2012.03.016. <http://www.sciencedirect.com/science/article/pii/S0196890412001628>.

[51] Piltan M, Shiri H, Ghaderi S. Energy demand forecasting in iranian metal industry using linear and nonlinear models based on evolutionary algorithms. Energy Convers Manage 2012;58(0):1–9. http://dx.doi.org/10.1016/j.enconman.2011.12.022. <http://www.sciencedirect.com/science/article/pii/S0196890412000052>.

[52] Breiman L et al. Classification and regression trees. New York: Chapman & Hall; 1984.

[53] Breiman L. Random forests. Mach Learn 2001;45(1):5–32. http://dx.doi.org/10.1023/A:1010933404324.

[54] Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2(3):18–22. <http://cran.r-project.org/web/packages/randomForest/index.html>.

[55] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011;2:27:1–27:27. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

[56] Mori H, Kurata E. An efficient kernel machine technique for short-term load forecasting under smart grid environment. In: Power and energy society general meeting, 2012 IEEE. 2012. p. 1–4. http://dx.doi.org/10.1109/PESGM.2012.6345687; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6345687>.

[57] Kouhi S, Keynia F, Ravadanegh SN. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection. Int J Electr Pow Energy Syst 2014;62(0):862–7. http://dx.doi.org/10.1016/j.ijepes.2014.05.036. <http://www.sciencedirect.com/science/article/pii/S014206151400307X>.

[58] Kaur A, Pedro HT, Coimbra CF. Ensemble re-forecasting methods for enhanced power load prediction. Energy Convers Manage 2014;80(0):582–90. http://dx.doi.org/10.1016/j.enconman.2014.02.004. <http://www.sciencedirect.com/science/article/pii/S0196890414001125>.

[59] Fan S, Chen L, Lee W-J. Machine learning based switching model for electricity load forecasting. Energy Convers Manage 2008;49(6):1331–44. http://dx.doi.org/10.1016/j.enconman.2008.01.008. <http://www.sciencedirect.com/science/article/pii/S0196890408000216>.