J Control Autom Electr Syst (2013) 24:854–862
DOI 10.1007/s40313-013-0067-5

# A Hybrid Approach to the Load Forecasting Based on Decision Trees

**Ricardo Menezes Salgado · Rodrigo Renno Lemes**

Received: 28 September 2012 / Revised: 30 March 2013 / Accepted: 13 May 2013 / Published online: 22 October 2013


**Abstract** This paper proposes a hybrid approach based on ID3 decision tree and expert systems to the load forecasting. The proposed model considers the typification of weekdays and uses the influence of climate information to estimate the load. The proposed methodology was tested on real data of electrical load measured in southeast of Brazil. The results show that the proposed model was promising with competitive results in the load forecasting area.

**Keywords** Load forecasting · Decision trees · Expert systems · Power system planning · Power system operation

## 1 Introduction

Currently, in an electric power system the load control is a basic requirement in the planning and operation. To supply electrical power with good quality, safely and economically, a company needs to have mechanisms that allow solving several problems of technical and operational level. Specifically, in the context of short-term planning, load forecasting is important in preparing the program operation the next day in the analysis of security and stability, because the forecast errors can affect the efficiency and security of the system (Gross and Galiana 1987; Salgado et al. 2006, 2009b). Thus, the knowledge of the behavior of future load is the first prerequisite for planning of electric power system (Salgado et al. 2004). Increase the safety and economy of the systems operation is the motivation for performing more accurate load forecasting in power systems.

The problem of load forecasting has four horizons: long, medium, short, and very short-term (Parlos et al. 1986; Gonzalez-Romera et al. 2006; Drezga and Rahman 1999; Salgado et al. 2009a; Yang et al. 2006). In long-term, the prediction is performed with the purpose of assisting the power planning supply on the horizon 10–20 years ahead. In medium-term, the period involved is few weeks, months, or to 2 years. This type of forecasting is important for fuel planning supply, maintenance on electrical networks, and serve as market research to companies negotiate contracts with other companies reducing financial risks. The short-term load forecasting is held in 30-min intervals up to 1 week ahead is used as a basic input in the planning and operation of the power systems. The last load forecasting horizon is very short-term, its period varies from a few seconds to 15 min. It is used to control the production of electricity in real time (Salgado et al. 2006) and, according to Papalexopoulos and Hesterberg (1989), a forecasting with accuracy level suitable can avoid undesirable disturbances in operations involving electrical grid.

In literature, there are several models for load time series forecasting in which we highlight the statistical models and models based on computational intelligence. Statistical models have as main objective to understand the generating mechanism of temporal data. This understanding enables to effectively describe the behavior of the series, finding periodicities, and obtaining equations describing the historical behavior with the use of the explanatory variables. Among the statistical methods we can mention: models based on exponential smoothing (Rahman and Hazim 1993), autoregressive models (Gross and Galiana 1987; Huang 1997), Box–Jenkins models, ARMA and ARIMA models (Chen et al. 1991; Zagrajek and Weron 2002), kalman filters (Infield and Hill

R. M. Salgado (✉) · R. R. Lemes
Computer Science Center, Institute of Exact Sciences (ICEx), Federal University of Alfenas (Unifal-MG), Gabriel Monteiro da Silva St, 700, Alfenas, MG 37130-000, Brazil
e-mail: ricardo@bcc.unifal-mg.edu.br

R. R. Lemes
e-mail: rodrigo@bcc.unifal-mg.edu.br

1998), as well as models based on the spectral expansion (Moghram and Rahman 1989).

Recently, models based on computational intelligence started to be used for load forecasting. These models have the advantage of a better response to the nonlinearity of the time series in addition to not depend on complex mathematical models. There are also hybrid models that are a combination of various methods of artificial intelligence so as statistical models. Among the techniques based on computational intelligence we can mention: artificial neural networks (ANN) (Saleh and Hoyle 2008; Gonzalez-Romera et al. 2008), fuzzy logic (Mamlook et al. 2009), expert systems (ES) (Chen et al. 1991), support vector machines (SVMs) (Ma et al. 2008), genetic programming (Huo et al. 2007), and others.

To perform a load forecast accurately it is important to know weather information, define the days of the week, as well as strategies that enable the treatment of atypical days. Currently, it is common to develop forecasting models involving various techniques. This approach is interesting because it enables the creation of models more robust and able to perform more accurate prediction if compared to models that involve the use of only one technique. Aiming to develop a reliable method to treat different types of days (holidays, Sundays, weekdays) as well as climatic variations, this paper proposes a hybrid approach to load forecasting that combines a decision tree type ID3 (Quinlan 1986), a model of ANN (Haykin 1999) in a knowledge base using an ES (Hayes-Roth and Lenat 1983). Therefore, we developed an ES for prediction based on decision trees that capture relevant information in the time series to estimate the load prediction. The model uses weather information to explain the behavior of the load.

In specialized literature, there are several papers proposing models for load forecasting. The paper proposed by Rahman and Hazim (1993), focuses on load forecasting in the short-term, the authors used statistical techniques and a knowledge base weather. The results show the mean absolute percentage error (MAPE) between 1.22 and 2.7 % for every month of the year. The paper of Mamlook published in 2009 (Mamlook et al. 2009) used fuzzy logic to forecast short-term load, using weather data. This model obtained a MAPE between 1.2 and 3.2 % during the tests described by the author.

The paper proposed by Ding (2006), used decision trees for load forecasting. The results obtained by the forecast model based on decision trees ID3 were compared with other algorithms, such as multiple linear regression, cubic polynomial model, exponential smoothing and, according to the author, the approach based on ID3 tree obtained the best forecasting results. Decision tree model provides a greater ease of interpretation of events. Each event is registered with their respective variables. A simple example is the decision problem in which we want to know what the load profile of the next day. Knowing that the day will be Friday, in daylight saving time (DST) and holiday, these characteristics can be recorded as variables. Variable Friday leads to an answer, because when it is Friday the load profile is slightly different from all other days of the week due to the start of the weekend. As we know that the day is in the DST the decision tree already eliminates the profile of the other Fridays that are not DST. Finally, it happens over a decision because of the holiday, which presents a load profile different from the days that are not holidays. In this case, the decision tree should find an approximate load profile according to the known attributes of the day to be analyzed. We can see then that the decision tree enables to classify the different types of days of week in the load forecasting problem.

In paper published in 2009 by Hong (Hong 2009a), was proposed algorithm chaotic ant swarm (CAS) optimization to search for suitable combinations of parameters in the model of load forecasting support vector regression (SVR). The results indicated that the proposed model had better results than other methods called SVR with chaotic particle swarm optimization (PSO) (SVRCPSO) (Wei et al. 2010), SVR-CGA (SVR with chaotic GA) (Hong 2009a,b), the regression model and ANN models.

In paper published by Liu (2010), was made daily forecast of electricity based on fuzzy rules using the Takagi–Sugeno inference model. The authors proposed a model based on genetic algorithm that automatically produces fuzzy rules for prediction. The authors emphasize that the proposed method is better ANNs for three reasons: slow learning of ANNs, difficulties in determine the structures of networks, problems with local minimum in training. In the results, the forecast for the day October 20, 2006, the proposed method was better than the ANNs by 0.17 % on average day with a MAPE of 1.08 % for the ANN and 0.91 % for proposed model.

In paper published by Hong et al. (2010) was proposed hourly load forecasting model using multiple linear regression with interactions. The authors made a study of seasonality in temperature. In the case study, the authors showed that the more variables used in the estimates were the best MAPE results for the year tested. The numerical values of MAPE for the best and worst strategy tested were 3.5 and 23.13 %, respectively.

As described above, the proposed hybrid model combines a model of type ID3 decision tree with an ES. The proposed technique is presented according to the following organization: Sect. 2 presents some concepts about load forecasting, in Sect. 3 presents the theoretical basis of the techniques used in the model. Section 4 describes the proposed methodology detailing their operating characteristics. Section 5 shows the results and final considerations are made in Sect. 6.

## 2 The Load Forecasting Problem

The load forecast is a problem of time series forecasting, in which the main goal is to find a model that accurately represents the behavior of the series over time. Figure 1 shows the classic diagram of a load forecast model, with multiple inputs in previous moments, and one output, which represents the prediction at time $t$.

Based on the historical values of the load time series, it is possible to observe several characteristics in relation to the consumption profile and the type of day. For example, in Brazil, we can observe that the loads on Tuesday, Wednesday, and Thursday, when they have no holidays, are very approximate. On the other hand, when the day is a holiday, the load consumed dropped significantly (Fig. 2) causing the consumption profile of the holiday be similar on a Sunday for example. In addition, there is a shift in the peak load in the DST, as well as a specific characteristic in the days that are followed by holidays (pre- and post-holiday).

In addition to the properties that can be observed with the load history, there is also the history of climatic data. Climatic information influence the behavior of the load curve and are important in forecasting. According to Wood and Wollenberg (1984) and Salgado et al. (2006), this type of external information is represented by the following variables: temperature, wind speed, fog, relative humidity, rain, and lightning. These papers also mentions that of all these variables the more relevant is the temperature, because it determines the greater or lesser use of air conditioners, heaters, refrigerators, and similar devices, which cause significant impacts on the load consumption profile.

There are also isolated events that influence the load consumption behavior. Such events occur randomly and cause disruption in the load. According to the paper (Salgado et al. 2006), these factors can be sporting events, television programs, stoppage of industries, and start or stop events where the load is high. The difficulty in treating these factors are difficult to detect in its history. We know that it is not common to have history of the industry stoppage, moreover industries have different energy consumption, which leads to the problem of treating different load curves profiles. A complex load forecasting is one that is made with all parameters as external factors and variables that affect the curve. For this to be possible, is necessary a detailed study of the consequences of each factor in relation to the load profile of the curve. When the history is insufficient the predicting model is unable to handle all the situations that may happen in the future.

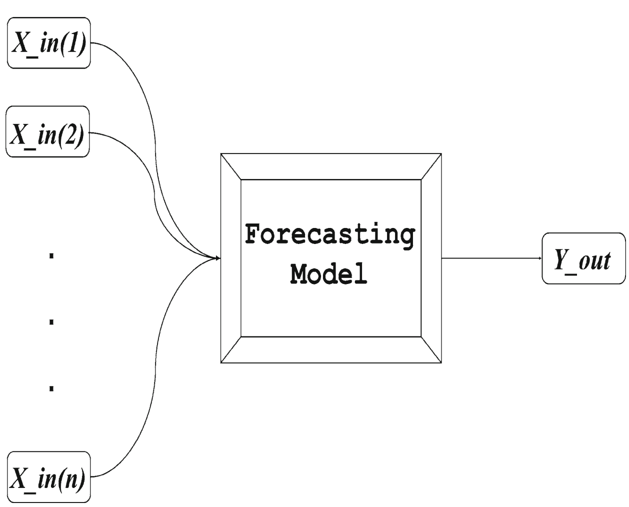


**Fig. 1** Diagram: forecast model

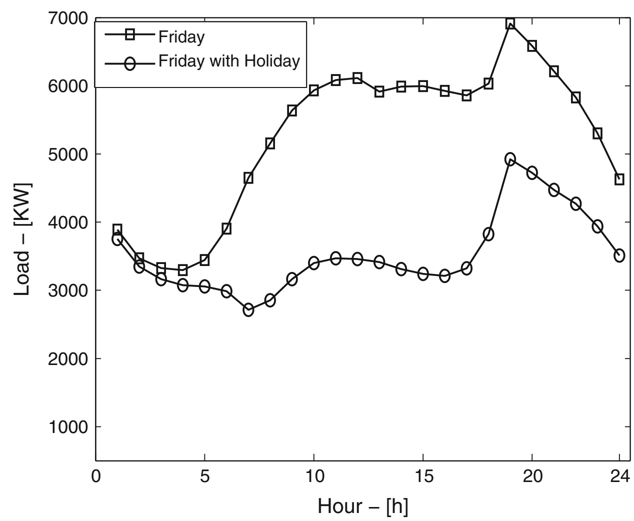


**Fig. 2** Load behavior: Friday (21/04/2000—holiday) and Friday (14/04/2000—typical workday)

As previously noted, there are several variables that affect significantly the load profile. With this viewpoint, the forecasting model proposed in this paper will be used as input variables the historical load (KW) combined with temperature information merged into a model that recognizes and treats different types of days (holidays, pre/post-holiday, working days, and weekends). The choice of these parameters was based on the influence they have on the explanation of the load as well as the availability of these variables in our historical data.

## 3 Theoretical Basis

The forecasting model proposed in this paper was designed using a hybrid approach based on three computational models: decision trees (Quinlan 1986), ANN (Haykin 1999), and a rule base represented by an ES (Hayes-Roth and Lenat 1983). Before treating the proposed modeling, the theoretical concepts of each of the three approaches used to design the hybrid model will be presented briefly.

### 3.1 Expert System

An ES can be defined as a knowledge-based intelligent model specially designed to act as a human expert in a specific field. They are typically used in situations where there is need speed in processing as well as actions in specific areas (Hayes-Roth and Lenat 1983).

An ES is composed of a knowledge base, consisting of facts, rules, and heuristics managed through an inference engine that acts as a input/output signal processor (Hayes-Roth and Lenat 1983). An ES is usually designed to act on a specific domain and should be able to provide information to assist decision making (Voelker and Ratica 1986). A type of ES is the model based on conditional rules. The rules are logical sentences (**IF ... THEN ... ELSE**), which form the knowledge base of the system. These sentences are called production rules that are used to represent the expert's knowledge (Reitman 1984).

### 3.2 Decision Trees

Decision tree is a simple and efficient algorithm for data classification. For this reason they are widely used as classification algorithms to construct classifiers that predict classes based on attribute values. The decision tree consists of a hierarchy of vertices (also called nodes) that can be internal or external. Usually decision trees use a strategy of divide and conquer in its construction, so a complex problem is decomposed into simpler subproblems and thus recursively the same strategy is applied to each subproblem (Quinlan 1986). In general, the external node will identify the final result for each existing class, related with the attributes to the classification. The internal node, also known as decision node or intermediate node, is the unit of decision making that evaluates by calculation what is the next node.

ID3 decision tree (Quinlan 1986) is a learning algorithm based on a family of algorithms top down induction of decision trees. ID3 uses the technique of induction of decision trees from a set of instances. The construction of a tree ID3 is an iterative process that part of the root to the leaves in which every step is necessary to determine the most relevant attribute (the attribute most relevant is that groups the instances according to the value of their respective classes in the best way) of the training set based on an evaluation of the function.

The algorithm used to build the decision tree ID3 has a heuristic of feature selection based on information gain during the construction process. That is, the algorithm aims to gain as much information as possible when new nodes are inserted into the tree. Furthermore, this algorithm has the property of build minimal trees.

ID3 decision tree is based on concepts of information theory such as entropy and information gain. To induce the decision tree, the ID3 need a set of training objects (instances of training). An instance can be represented by a set of attributes and their values. The ID3 purpose is to split the training set and iteratively to choose the term that results in greater information gain, thereby building the decision tree. The idea is that the higher the information gain, the lower the entropy of the nodes generated in the new branches. A interpretation of entropy measures in a data set is that the lower the entropy, the more homogeneous is the dataset and its distribution by categories. Entropy is a measure used in information theory and is interpreted as the amount of information contained in an attribute (or set of variables), the greater the entropy value, the higher the uncertainty regarding the content (Shannon 1948; Shannon and Weaver 1998). Thus, the entropy of a set of instances is the amount of information needed for the classification of an instance of any of this set with regard to its relevance to a particular class. It reaches the maximum value 1 when the class have the same probability of occurrence and reaches the minimum value 0 when all the instances in the same class. Thus, the higher the entropy of a given set of instances greater the uncertainty in the classification of an instance of this assembly with respect to its membership in a particular class.

To build a decision tree the ID3 algorithm search between the attributes of the training set, which provide the lowest entropy value. To calculate the entropy for a given attribute is necessary to divide the training set into subsets, grouping the instances that have the same value for the attribute in question. To build a decision tree is necessary to add nodes, to represent the attributes in the tree according to the information gain. When there are only objects of the same class, a leaf node with the name of that class is inserted. Note that maximize the gain means minimizing entropy. The entropy calculation is performed according to Eq. 1

$$\text{entropy}(x) = \sum_{i=1}^{n} p_i * \log_2(p_i), \tag{1}$$

where $p_i$ is the probability of response for each attribute of the node, $n$ is number of attribute of the each node.

The gain is a variable that is also part of the calculations and measuring the reduction in entropy caused by the partitioning of the training set according to the attribute values. The gain calculation is performed according to Eq. 2

$$\text{gain}(Vp, At) = \text{entropy}(Vp) - \sum \frac{Vp_v}{Vp} * \text{entropy}(Vp_v), \tag{2}$$

where $Vp$ is the set of all values partitioned, $At$ is selected attribute, $Vp_v$ is the entropy associated to a given attribute.

In general, the purpose of the tree is to reduce the entropy, i.e., to reduce the randomness of the variable object, since the greater randomness greater the difficulty of classification. As the algorithm recursively splits the dataset original training, the divisions are evaluated with smaller and smaller samples.

In order to minimize this problem and avoid over adjustment of training data with trees very complex there is a strategy used in this paper. The mechanism used in this paper was an alternative procedure to stop in tree. Further details about this model of decision tree can be found in the papers of Quinlan (1986) and Mitchell (1997).

### 3.3 Artificial Neural Network

ANN are parallel distributed processing systems that are capable of storing experiential knowledge (Haykin 1999). The ANNs can be classified as static neural networks (non-recurring) or as recurrent neural networks (dynamic). The main difference between both structures is the presence or absence of connections to resend signals for neurons present in the network architecture. When there is no dependence of output current with past values, the ANN are static.

This paper uses the multilayer perceptron (MLP) neural network to be ANN class most used to the time series forecasting. The MLP network with a single hidden layer is usually formed by $n$ inputs, $h$ neurons in the hidden layer, and $m$ neurons in the output layer. The neurons are connected through synaptic weights that should be adjusted during the training process. In this paper, we used the backpropagation training algorithm to the adjusting the weights (Rumelhart et al. 1986).

## 4 Proposed Methodology

The basic hypothesis for this methodology is the use of similar patterns sets to calibrate the prediction model. Based on this idea, for each time interval a table formed with similar patterns that will serve as input to the hybrid forecasting algorithm proposed in this paper is generated. The table of similar patterns, called decision table, is formed by the following information: holiday, post-holiday, weekday, month, DST, maximum, average and minimum temperature of the day, and the corresponding load measured for each time interval. In Table 1 we can observe an example of a decision table for a given hour of day.

The hybrid forecasting model proposed in this paper is built through two stages: the first stage is the forecasting of the temperature and the creation of the decision table. The second phase performs the analysis and data classification of decision table, via decision tree, and from this result makes the calculation of the load forecasting based on predetermined rules. Figure 3 shows the general diagram of the hybrid forecasting model. Both phases are described and discussed below.

Algorithm 1 show sequentially the steps of the proposed model. Sections 4.1 and 4.2 details the procedures necessary for the implementation of Phases I and II, respectively, with illustrative examples of the application of the model steps.

**Table 1** Decision table

| | 02/08/2000 | 01/08/2000 | 27/07/2000 |
|---|---|---|---|
| DDS | Wed | Tue | Thu |
| Month | Aug | Aug | Jul |
| TMD | 18.2 | 15.1 | 13.1 |
| TMX | 26.7 | 23.7 | 17.3 |
| TMI | 12.5 | 12.3 | 10.6 |
| FRD | No | No | No |
| DST | No | No | No |
| PF | No | No | No |
| CA | 4478.20 | 4351.95 | 4390.75 |

*DDS* weekday, *TMD* average temperature (°C), *TMX* maximum temperature (°C), *TMI* minimum temperature (°C), *FRD* holiday, *DST* daylight saving time, *PF* post-holiday, *CA* load (KW)

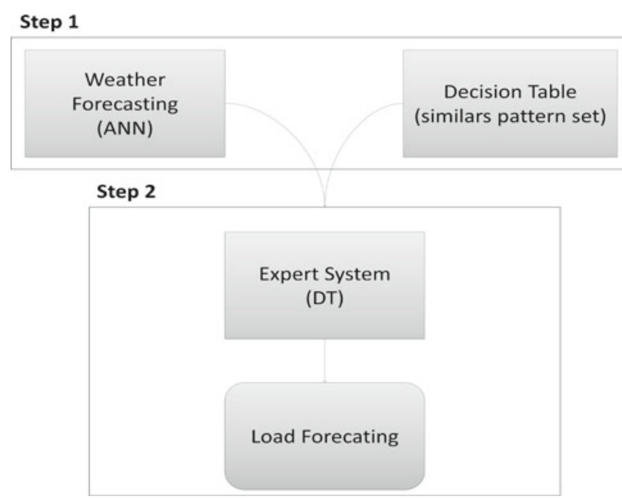


**Fig. 3** Diagram: proposed methodology

**Algorithm 1** Steps of proposed model

```
1. Create Decision Table (phase I);
2. Perform the temperature forecasting (phase I);
3. Use the ID3 decision tree to select the data (in decision table) that will be used to perform forecasting (phase II);
4. Apply the expert system for estimating the load forecasting (phase II).
```

### 4.1 Description: Phase I

In this paper, Phase I is responsible for obtaining the temperature forecasting and create the decision table for the load forecasting in Phase II. The decision table illustrated in Table 1 is built for each time interval (day/h). The base rules used for construction of the decision table is shown in Algorithm 2. Based on these rules, the ES can create a decision table for each time interval using the historical data.

**Algorithm 2** Rules for construction of decision table

```
1. If the day of the week is Tuesday,
Wednesday or Thursday is allowed to put in
the table only data measured in one of these
three kinds of days of the week;
2. If the day of the week Monday, will be
inserted in the table only loads measured in
Monday, the same rule goes for Friday, post-
holiday, Saturday and Sunday;
3. If the day is a holiday, will be included
in the decision table Sundays and holidays;
4. Are included in the table the following
attributes: maximum, minimum and average
daily temperature, holiday, post-holiday,
DST, month, day of week and load;
5. For each hour of the day is calculated a
new decision table.
```

Another goal of Phase I is to estimate the forecast of the temperature (maximum, average, and minimum) for the day to be forecasted. As described earlier, this forecasting can be performed with any technique for time series forecasting in the literature [ANN, fuzzy logic, ESs, support vector machines (SVMs), genetic programming, etc.]. In this paper, the prediction of the maximum, average, and minimum temperature were obtained through an MLP ANN model. The MLP neural network was configured to perform the forecasting of the three series of temperature (max, average, and min). The number of neurons in the hidden layer and the value of the term momentun were determined through an exhaustive search in the following sets [1, 15] and [0:01, 0.99], respectively. The learning rate was started with a value of 0.5 and in each epoch, an one-dimensional is performed to find the next value.

### 4.2 Description: Phase II

Phase II is responsible for obtaining the load forecasting based on data obtained in Phase I. After the construction of the decision table (Phase I), the forecasting algorithm needs to select which data will be used as the basis for performing load forecasting. To choose this data the model uses an estimated temperature (maximum, average, and minimum) calculated in Phase I. With the forecast of the temperature is possible to execute the decision tree and select, among the data of the decision table, the data will be used as the basis for estimating the load forecasting.

The ID3 algorithm search in decision table the data according to the attributes of the day to be forecasted, taking into consideration the minimum temperature, average and maximum forecasted in phase I. After selecting the data in the decision table the load forecasting is performed using an ES which performs a comparison with the average load data in decision table. The forecasted load value will be estimated according to the rules described in Algorithm 3:

**Algorithm 3** Expert system: calculation of load forecast

```
1. If in the chosen data the value of the
load is greater than the average load of the
decision table the forecasted load will be
the average of the records that have higher
value than the average load of the decision
table;
2. Otherwise, ie, if in the chosen data the
value of the load is below average loads of
decision table the forecasted load will be
the average of the loads that are below the
average load value of decision table;
```

Table 2 shows an example of decision table with 5 (five) data. This table was constructed to foresee the 23 h of the day 11/02/2000 (holiday in Brazil). We can observed that the selected records to the training table comprise data on Sunday as well as Thursday, 10/12/2000 which is a national holiday in Brazil. We can note that the average loads of data of decision table (Table 2) is 4001.34 KW.

The forecasted load is calculated based on the data selected by ID3 algorithm. Thus if the second-line is chosen by the decision tree, we have a load of 4094.54 KW which is higher than the average load of decision table. In this sense, the forecasted load is calculated from the average of the loads that are greater than the average table, in this case, the predicted load is 4100.96 KW.

**Table 2** Decision tables in load forecasting

| | 10/12/2000 | 10/22/2000 | 10/08/2000 | 10/15/2000 | 10/29/2000 |
|---|---|---|---|---|---|
| DDS | Thu | Sun | Sun | Sun | Sun |
| Month | Out | Out | Out | Out | Out |
| TMD | 21.0 | 20.4 | 18.0 | 19.8 | 16.9 |
| TMX | 38.5 | 27.5 | 24.0 | 23.7 | 21.6 |
| TMI | 17.5 | 17.2 | 14.7 | 17.6 | 15.1 |
| FRD | Yes | No | No | No | No |
| DST | Yes | Yes | Yes | Yes | Yes |
| PF | No | No | No | No | No |
| CA | 4107.38 | 4094.54 | 3985.84 | 3884.15 | 3834.80 |

On the other hand, if the record selected by the decision tree is the day 10/15/2000 that has load equal to 3884.15 KW which is lower than the average load of the decision table then the forecasted load will be the average of loads that are lower than the average load in decision table that in this case is equal to 3901.60 KW.

## 5 Case Study

To test the efficiency of the proposed model the simulations were performed involving different periods, various days, and various conditions. The use of the model in different scenarios aims to test its robustness in day-to-day of the system operator. The historical data used to adjust and test the model is formed by actual electrical load data measured at southeastern Brazil in the period 04/01/2000 to 04/30/2001. We also use a historical data of temperature (maximum, average, and minimum) measured in the same period of load.

The load forecasting model has been tested in four different periods: June/2000, August/2000, November/2000, and January/2001. In each period, the goal was to foresee the daily load of the month. Tests were performed considering two approaches to forecasting: one-step ahead and $n$-steps ahead. The forecast one-step ahead takes into consideration the recent history with loads known until the day before the day to be forecasted. The forecast $n$-steps ahead using a known load history up to a certain fixed day and also uses load values predicted by the model as inputs to foresee the load.

The numerical results will be presented sequentially considering approaches to forecast one-step ahead and $n$-steps ahead. Model performance is evaluated by the value of MAPE, shown in Eq. 3:

$$\text{MAPE}(\%) = \frac{100}{24}\sum_{i=1}^{24}\frac{|x_i - \widehat{x}_i|}{x_i} \tag{3}$$

where $x_i$ is the actual load and $\widehat{x}_i$ is the forescasted load.

### 5.1 Analysis of the Results

In the forecast of the temperature (max, average, and min), the neural network model found predictions with MAPE ranging from 3.5 to 9.2 % depending on the month selected. As the parameters of the model were determined by an exhaustive search, we can note that several configurations with different learning rates, term momentun, and number of neurons in the hidden layer showed promising results. The results of the temperature forecast make possible find good choices in the ID3 decision tree, providing good results in load forecasting.

Figures 4, 5, and 6 show the forecasted and observed daily curves for one-step and $n$-steps ahead approaches for some

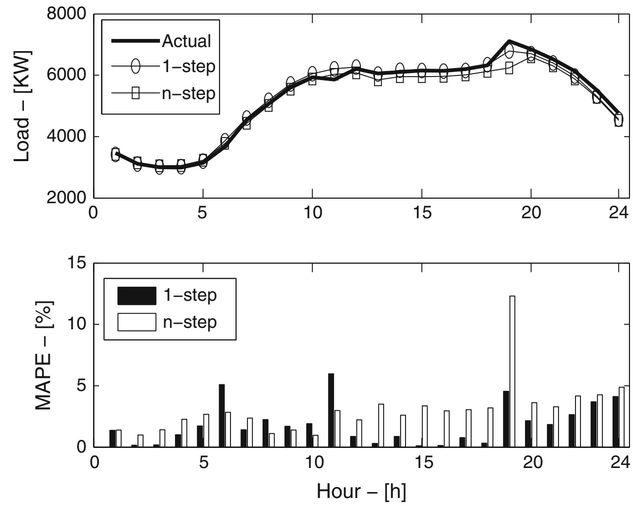


**Fig. 4** Results: Monday—09/04/2001

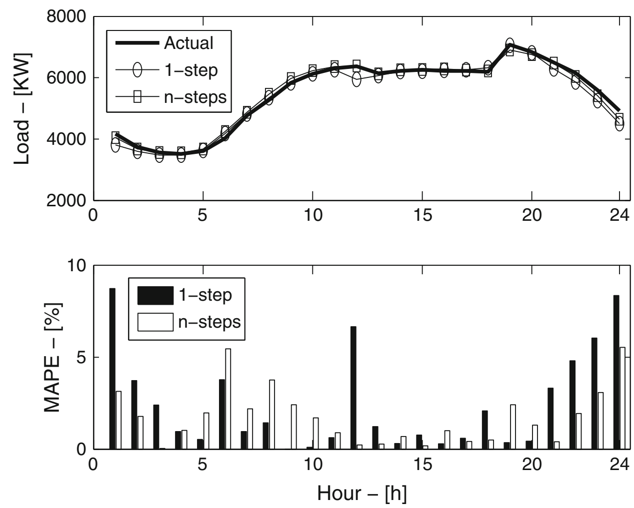


**Fig. 5** Results: Wednesday—11/04/2001

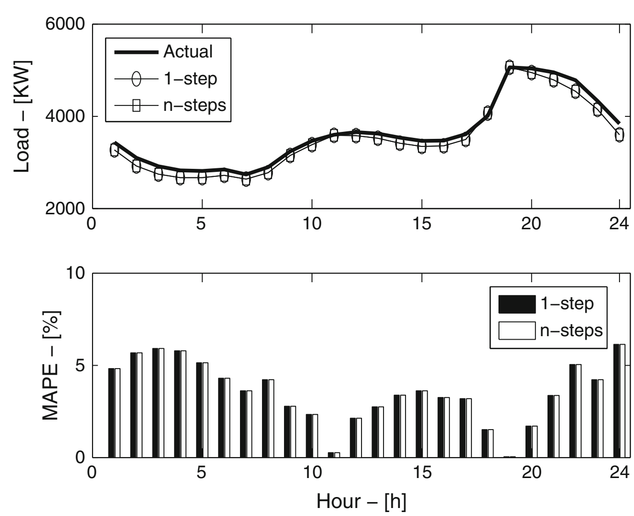


**Fig. 6** Results: Thursday 02/11/2000 (holiday)

specific types of days: Monday (09/04/2001), Wednesday (11/04/2001), and day holiday 02/11/2000. Observing the results shown in Fig. 4 we can note that the proposed model was efficient to foresee the daily curve in the two approaches adopted. Also in this figure we can see that the MAPE in one-step approach was lower if compared with the MAPE of $n$-steps approach. However, note that the model presented similar profiles suitable for the two approaches used in the simulations.

A different behavior occurs for the Wednesday day 04/11/2011, shown in Figure 5. In this case, the forecasting of the approach $n$-steps presented more accurate results in comparison with the one-step approach. This behavior, although unusual, can occur because the data available to the model $n$-steps may contain information that fitting more accurately the day that was selected for prediction. This does not mean that the $n$-step approach is superior, but in some isolated cases it can outperform the approach one step. However, we can observe that both approaches were able to correctly trace the curve profile provided daily.

Another behavior that may occur in the model is shown in Figure 6. In this simulation, the goal was to foresee a holiday and as the data selected for the decision table have a reduced number of elements, the model showed the same result for the one-step and $n$-steps approaches. This fact happens due to the limited number of registers of this type of days in database. And thus, both approaches are selecting the same data to perform the load prediction.

Figure 7 shows the average load forecast for the month of June/2000. Note that the proposed model showed satisfactory results for both approaches used in simulation, with most days having MAPE in accordance with acceptable limits for electric load forecasting short-term (Ghiassi and Zimbra 2006).

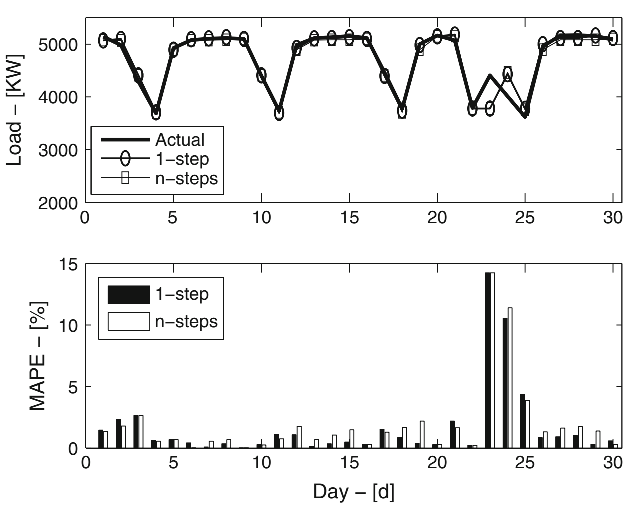


**Fig. 7** Results: average load—June 2000

**Table 3** MAPE in each month

| | Jun/2000 | Aug/2000 | Nov/2000 | Jan/2001 | AVG |
|---|---|---|---|---|---|
| 1-step (%) | 1.67 | 2.46 | 3.20 | 4.67 | 3.00 |
| $n$-steps (%) | 1.92 | 6.19 | 3.76 | 11.65 | 7.20 |

Table 3 shows the forecast errors for the average load in the months of June/2000, August/2000, November/2000, and January/2001. Note that, when the model runs in one-step approach has a better performance than the $n$-step approach. This behavior is expected, since the one-step approach has recent historical information making the forecasting more accurate. In general, we can say that the proposed model was effective in solving the problem of load forecasting obtaining competitive results compared to literature. The fact that the model take into account explanatory variables (temperature, typifying days, etc.) makes the model be robust making it possible to obtain more accurate results.

## 6 Conclusion

This paper presented a hybrid approach to load forecasting based on the combination of a model based in ID3 decision tree and a rule base represented by an ES. The model was developed to treat different types of days (holidays, Sundays, weekdays, etc.) as well as to consider weather data as explanatory variables in the prediction.

The results showed that the model was able to obtain results consistent with the literature in the area of load forecasting as well as providing results with errors in accordance with acceptable limits for electric load forecasting short-term in Brazil. It was also found that the model presented interesting results to approaches to one-step and $n$-step and this fact showing that even with the lack of recent history the model is able to estimate the load with an adequate accuracy and within safe parameters of load forecasting. It is noteworthy that for operation of the model the availability of data temperature minimum, average, and maximum is necessary. The lack of this variable prevents the use of the proposed model.

As the data used to obtain the results were real data (measured in southeastern Brazil), we can conclude that the proposed model can be applied in practical situations to obtain load forecasting to the regulators and electric utilities in Brazil.

**Acknowledgments** The authors would like thank the National Council for Scientific and Technological Development (CNPq) for financial support, the Federal University of Alfenas (Unifal-MG), and Computational Intelligence Laboratory (LINC) for providing the necessary infrastructure for the development and execution of this research.

## References


Chen, D., Chen, B., & Li, T. (1991). An expert system for short-term load forecasting. In *International Conference on Advances in Power System Control, Operation and Management—APSCOM-91* (Vol. 1, pp. 330–334).

Ding, Q. (2006). Long-term load forecast using decision tree method. In *IEEE PES Power Systems Conference and Exposition* (Vol. 9, pp. 1345–1355).

Drezga, I., & Rahman, S. (1999). Short-term load forecasting with local ann predictors. *IEEE Transactions on Power Systems*, *14*(3), 844–850.

Ghiassi, M., & Zimbra, D. K. (2006). Medium term system load forecasting with a dynamic artificial neural network model. *Electric Power Systems Research*, *76*(5), 302–316.

Gonzalez-Romera, E., Jaramillo-Moran, M., & Carmona-Fernandez, D. (2006). Monthly electric energy demand forecasting based on trend extraction. *IEEE Transactions on Power Systems*, *21*(4), 1946–1953.

Gonzalez-Romera, E., Jaramillo-Moran, M. A., & Carmona-Fernandez, D. (2008). Monthly electric energy demand forecasting with neural networks and fourier series. *Energy Conversion and Management*, *49*(11), 3135–3142.

Gross, G., & Galiana, F. D. (1987). Short-term load forecasting. *Proceedings of the IEEE*, *75*(12), 1558–1573.

Hayes-Roth, F., & Lenat, D. B. (1983). *Building expert systems* (Vol. 1). Boston, MA: Addison-Wesley Longman Publishing Co.

Haykin, S. (1999). *Neural networks: A comprehensive foundation*. New York: Macmillan Publishing Company.

Hong, W. C. (2009a). Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. *Energy Conversion and Management*, *50*(1), 105–117.

Hong, W. C. (2009b). Hybrid evolutionary algorithms in a svr-based electric load forecasting model. *International Journal of Electrical Power and Energy Systems*, *31*(1), 409–417.

Hong, T., Gui, M., Baran, M. E., & Willis, H. L. (2010). Modeling and forecasting hourly electric load by multiple linear regression with interactions. In *IEEE Power and Energy Society General Meeting* (Vol. 1, pp. 1–8).

Huang, S. (1997). Short-term load forecasting using threshold autoregressive models. *IEE Proceedings—Generation, Transmission and Distribution*, *144*(5), 477–481.

Huo, L., Fan, X., Xie, Y., & Yin, J. (2007). Short-term load forecasting based on the method of genetic programming. In *International Conference on Mechatronics and Automation* (Vol. 1, pp. 839–843).

Infield, D. G., & Hill, D. G. (1998). Optimal smoothing for trend removal short term electricity demand forecasting. *IEEE Transactions on Power Systems*, *13*(3), 1115–1120.

Liu, T. (2010). The daily electric load forecasting based on fuzzy rules. In *Second International Conference on Computer Modeling and Simulation* (Vol. 2, pp. 8–10).

Ma, L., Zhou, S., & Lin, M. (2008). Support vector machine optimized with genetic algorithm for short-term load forecasting. In *International Symposium on Knowledge Acquisition and Modeling—KAM'08* (Vol. 1, pp. 654–657).

Mamlook, R., Badran, O., & Abdulhadi, E. (2009). A fuzzy inference model for short-term load forecasting. *Energy Policy*, *37*(4), 1239–1248.

Mitchell, T. (1997). *Machine learning* (1st ed.). New York: The McGraw-Hill Companies, Inc.

Moghram, I., & Rahman, S. (1989). Analysis and evaluation of five short-term load forecasting techniques. *IEEE Transactions on Power Systems*, *4*(4), 1484–1491.

Papalexopoulos, A. D., & Hesterberg, T. C. (1989). A regression based approach to short-term system load forecasting. *IEEE Transactions on Power Systems*, *5*(4), 414–423.

Parlos, A. G., Oufi, E., Muthusami, J., Patton, A., & Atiya, A. (1986). Development of an intelligent long-term electric load forecasting system. In *Intelligent Systems Applications to Power Systems* (Vol. 1, pp. 288–292).

Quinlan, J. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106.

Rahman, S., & Hazim, O. (1993). A generalized knowledge-based short-term load-forecasting technique. *IEEE Transactions on Power Systems*, *8*(2), 508–514.

Reitman, W. (1984). *Artificial intelligence applications for business* (1st ed.). Norwood, NJ: Ablex.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Saleh, J. M., & Hoyle, B. S. (2008). Improved neural network performance using principal component analysis on matlab. *Journal of the Computer, the Internet and Management*, *16*(2), 1–8.

Salgado, R., Ballini, R., & Ohishi, T. (2009a). An aggregate model applied to the short-term bus load forecasting problem. In *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES* (Vol. 1, pp. 1–8).

Salgado, R., Ballini, R., & Ohishi, T. (2009b). An hybrid aggregate model applied to the short-term bus load forecasting problem. In *PowerTech, 2009 IEEE Bucharest* (Vol. 1, pp. 1–8).

Salgado, R., Ohishi, T., & Ballini, R. (2004). Clustering bus load curves. In *Power Systems Conference and Exposition, 2004. IEEE PES* (Vol. 3, pp. 1251–1256).

Salgado, R. M., Pereira, J. J. F., Ohishi, T., Ballini, R., Lima, C. A. M., & Zuben, F. J. V. (2006). A hybrid ensemble model applied to the short-term load forecasting problem. In *Proceedings of the International Joint Conference on Neural Networks—IJCNN* (Vol. 1, pp. 2627–2634).

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423.

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication* (1st ed.). Champaign: The University of Illinois Press.

Voelker, J. A., & Ratica, G. B. (1986). The genesis of a knowledge-based expert system. *IBM Systems Journal*, *25*(2), 181–189.

Wei, C., Hong, C., Lai, W. M., & Hung, Y. D. (2010). Electric load forecasting by svr with chaotic ant swarm optimization. In *IEEE Conference on Cybernetics and Intelligent Systems (CIS)* (Vol. 27, pp. 102–107).

Wood, A., & Wollenberg, B. (1984). *Power generation, operation, and control* (1st ed.). New York: Wiley.

Yang, H. Y., Ye, H., Wang, G., Khan, J., & Hu, T. (2006). Fuzzy neural very-short-term load forecasting based on chaotic dynamics reconstruction. *Chaos Solutions & Fractals*, *29*(2), 462–469.

Zagrajek, J. N., & Weron, R. (2002). Modeling electricity loads in california: Arma models with hyperbolic noise. *Signal Processing*, *82*(12), 1903–1915.