Load Forecasting Using Artificial Neural Networks and Support Vector Regression

SILVIO MICHEL DE ROCCO[1,3] ALEXANDRE RASI AOKI[1,2] MARCUS VINICIUS LAMAR[3]

[1]Electrical Engineering Department
Parana Federal University
UFPR Polytechnic Center – PO Box 19011
BRAZIL
http://www.ufpr.br

[2]Institute of Technology for Development
LACTEC
UFPR Polytechnic Center – PO Box 19067
CEP 81531-980 – BRAZIL
http://www.lactec.org.br

[3]Paraná Energy Company
Rua José Izidoro Biazetto, 158 – Blobo C
CEP 81200-240 – BRAZIL
http://www.copel.com

[1]Computer Science Department
Brasilia Federal University – PO Box 4466
CEP 70919-970 – BRAZIL
http://www.unb.br

Abstract: - This paper describes a short term load forecasting system using two different techniques of Artificial Intelligence: Recurrent Artificial Neural Networks and Support Vector Regression. A brief analysis of the load over the distribution systems connection points in Brazilian Parana States is also done.

Key-Words: - Load Forecasting, Artificial Neural Networks, Support Vector Regression.

1 Introduction

With the privatization of the sector of electric energy and the commercialization of the energy between the generators, the load forecasting becomes important for these companies, as much long the how much short-term one.

The long term load forecasting is necessary for buying energy and short term load forecasting is important to guarantee that the distribution system works within the contracted limits, avoiding heavy penalties from Brazilian Regulatory National Agency.

An important requirement of this type of forecasting is that the error is less than the tolerance foreseen in these contracts that are generally of about 5%. The exchanges of electric energy are carried through the connection points of the distribution system that are the distribution substations directly connected to the rest of the national electrical system.

In this work it is analyzed the short-term load forecasting in connection points of the distribution system using two popular forecasting methods of time series: Artificial Neural Network (ANN) and Support Vector Regression (SVR).

Firstly it is done brief analysis of the typical load profile of these connection points. Following, the two methodologies used are described, the results obtained are analyzed and conclusions on this work are presented.

2 Methodology

2.1 Load Data Analysis

The load data used in this study consist of the integrated consumed active power in 15 minutes of the connection points in the state of Paraná, between 2004 to 2006. Examples of these data can be seen in Figure 1 and Figure 2.

Analyzing the data it can be observed daily repetition standards, related to the differences between the consumption during the week and in weekends.

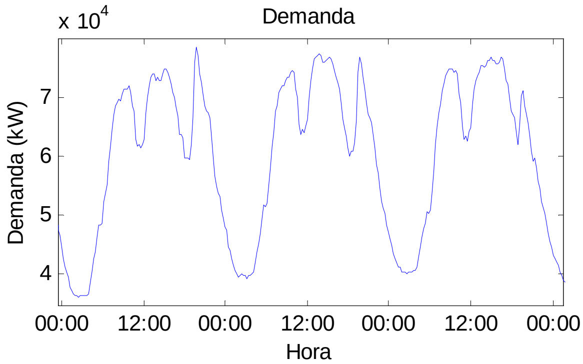

Figure 1 - Load in three consecutive days

Through Figure 1 it is noted daily seasonality, with peaks of consumption nearing to 09:00, 14:00 and 20:00 hours, and low consumption during the dawn, between 00:00 and 06:00 hours.

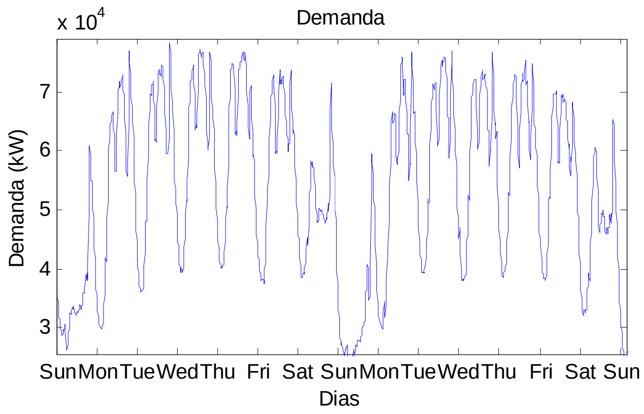

Figure 2 - Load in two consecutive weeks

In Figure 2 the weekly seasonality can be observed, with the Sundays presenting a differentiated profile to other days.
The annual seasonality could not be observed, once the observed period corresponds about one year and half. To allow one better understanding of the involved time series, it was done an analysis of the function of self-correlation of the data.

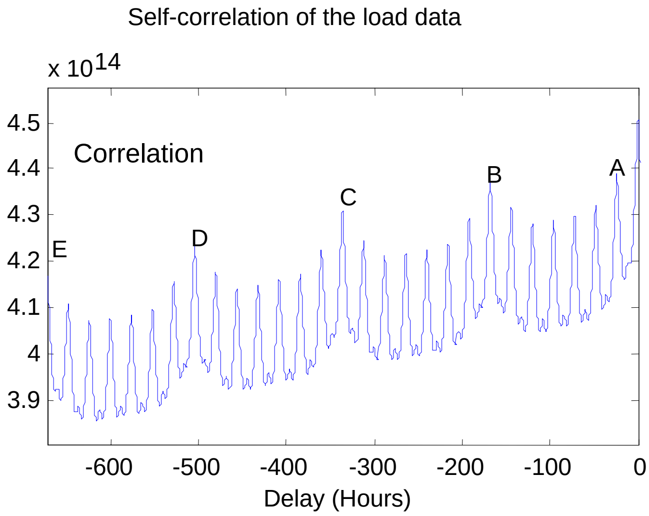

Figure 1 - Self-correlation of the demand data, up to 4 weeks of delay

Analyzing Figure 3, it is possible to conclude that the data correlation diminishes with distance with time between. Figure 3 shows that the higher correlation peaks in certain delays. Table 1 indicates the corresponding delay to the points nominated in Figure 3.
As it can be viewed the correlation of the load is visibly bigger in the same time with one day and one week of delay, besides the immediate previous measures (15, 30 and 45 minutes of delay).

Table 1- Point of interest in the self-correlation

Point	Delay
A	24 h (1 day)
B	168 h (1 week)
C	336 h (2 weeks)
D	504 h (3 weeks)
E	672 h (4 weeks)

As the objective of this work is the forecast of the daily peaks of demand, one possibility is to use only these points to achieve through the forecast. Figure 4 presents the characteristic of the demand when only the daily peaks are considered.

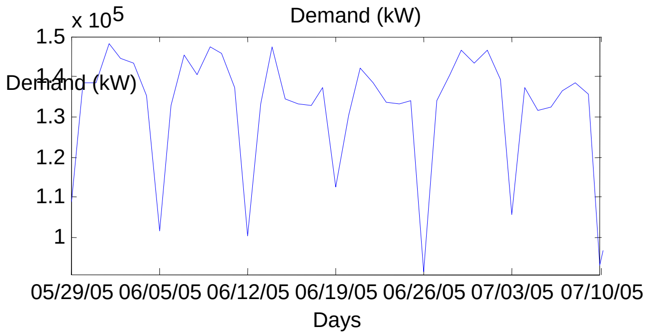

Figure 4 - Daily peaks of six weeks of load

Again, a weekly data seasonality can be noted in Figure 4. The self-correlation of these data can be seen in Figure 5.

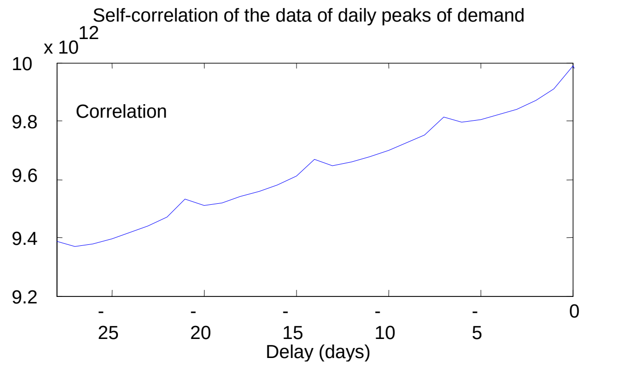

Figure 5 - Self-correlation of the daily peaks of demand

In contrast of what it occurred with the integrated data in 15 minutes, in Figure 5 it can be seen that the daily peak of demand doesn't present a significant correlation with no delay in special, occurring only one small peak for delay multiples of one week.

2.2 Load Forecasting Using Artificial Neural Networks

Initially, the use of Artificial Neural Networks was evaluated (ANN), a well consolidated technique, to do the prediction.

Artificial Neural Networks are computational systems based in an approach of the human brain modeling [1]. Simple units called neurons are linked to form a net.

The Neural Networks, opposite to the traditional techniques of computation, are not linearly programmed to solve problems. They learn the solution of a problem, and therefore must be trained through the presentation of examples of this solution.

For analysis of time series correlation, it is possible to re-feed the net in various different ways, creating nets whose reply not only with the current data in the entrance, as well as of the previously data presented. Amongst the diverse structures capable to deal with time series, the chosen configuration was the partially recurrent net of Elman [2].

The structure proposal for Elman is sufficiently simple, however its capacity was demonstrated to learn any time series, since that it has an adjusted structure, i.e. number of neurons, and either properly trained with an adequate learning algorithm [3]. This structure, presented in Figure 6, consists basically of a context layer, that stores the exits of the hidden intermediate layer, for then re-feeding these values in the net in the next time iteration.

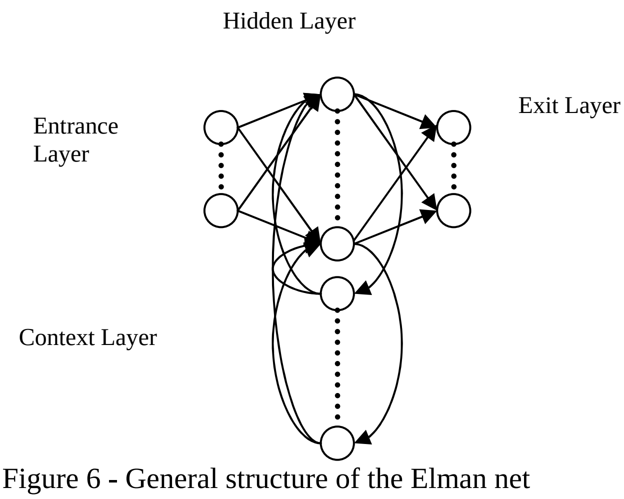

Figure 6 - General structure of the Elman net

The structure presented in Figure 6, given one time series, allows the forecast of a step of time ahead. So that it is possible to do the forecast some points ahead using the foreseen values already, that is, the exit of the net, as the entry of the same.

Firstly it was used a net with only one neuron of entrance and a neuron in the exit. As an analytical methodology does not exist to define the number of neurons in the hidden intermediate layer, 10, 25 and 50 had been tested neurons in this layer.

Another tested configuration was using some previous measures of load in the entrance of the net, in order to foresee its future behavior. In the intermediary layer again 10, 25 and 50 neurons were tested. The previous points to be used in the entrance were chosen in accordance with the analysis of the self-correlation function of the measured load data.

As it will be shown in section 3, the forecast with the complete data presents a considerable error, due to amount of points that must be foreseen. To solve this problem, were tested configurations using only the load daily peaks. The structures with a neuron in the entrance had been tested and with various relative entrance to previous measures.

It is possible to obtain better results with ANN, through the method of attempt and error for the definition of the structure, iterations during the training, and adjustments of the parameters of the trainings algorithm used, being that this task requires much time of simulation, to reach one structure ANN that presents excellent resulted for the load forecasting for one determined connection point of the distribution system. However due to the great number of connection points and its particularities, such methodology shows impracticable of being applied in a time of adequate training.

Given the difficulty to get satisfactory results with neural networks, new methodologies had been looked to solve the problem of load forecasting.

2.3 Load Forecasting Using Support Vector Regression

The Support Vector Regression (SVR), it is one technique of nonlinear regression based in Support Vector Machines (SVM) proposed for [4]. Both the techniques firmly are based in theory of statistical learning, or theory VC, that it has been developed in last the three decades buy Vapnik, Chervonenkis, among others.

Theory VC characterizes the properties of the learning machines in order to allow the excellent generalization of data not seen.

It was observed during the tests with neural networks that it is unnecessary to work with the integrated complete data of 15 in 15 minutes, then, in these experiments had been used only the daily peaks. The data had been normalized between -1 and 1, to prevent to work with the relatively great values of load. The function kernel chosen was the radial base function (RBF), that it is most complex among the functions used in SVM, but also the most common. The chosen edge was of 0.05, what, after the normalization, represents an error of 2.5% above or below the training data. To simplify the forecast, information on the day of the week in the entrances had been also codified, in accordance with Table 2 [5].

Table 2 - Codification of the days of the week

Week days	Code
Sunday	000000
Monday	100000
Tuesday	010000
Wednesday	001000
Thursday	000100
Friday	000010
Saturday	000001

The parameter cost was increased gradually until not having a significant reduction of the error

Observing that the data correspond to a seasonality of long delay (annual, monthly) and one of low delay (weekly), it was proposed to divide the data for filtering in two parts, one of low frequency, being represented the monthly, annual seasonality and the trend, and another high-frequency one, representing the weekly seasonality.

A digital filter FIR of order 100, was projected by the frame technique with a window of Hamming, for a frequency of normalized cut of around 1/14, what it would physically represent a complete periodic cycle to each two weeks. The high frequency was obtained through the subtraction between the original signal and the component of low frequency. So that it didn’t have delay and loss of data during the filtering, it was necessary to mirror the 50 initial points and the 50 final points of data, before the filtering process.

Figure 7 presents real data, the component of low frequency obtained by filtering and the high-frequency component. The addition of the signal of low frequency as high-frequency signal, perfectly reconstructs the original data, this technique based on analysis of signals for sub-bands.

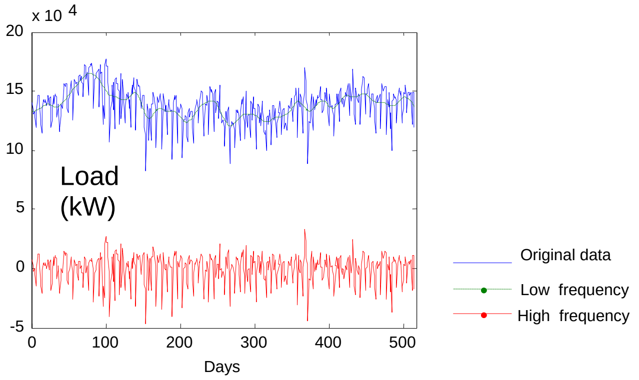

Figure 7 - Separation of data in sub-bands

3 Tests and Results

As measure of forecasting performance, it was is adopted Mean Absolute Percent Error (MAPE), defined by:

$$E_{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{V_{previsto} - V_{real}}{V_{real}}\right| \times 100\%$$

where Vprevisto corresponds to the foreseen value and Vreal the measured value of actual load, in N instants of time.

The period of time in which the error is measured, in all the cases, corresponds to one week of forecast.

The neural Network with only one entrance wasn’t capable to forecast satisfactorily the integrated data in 15 minutes. Using the 26 previous entrances with the better correlation with the current point (delays of 15 minutes 2:30 hours, 24 hours and between 6 days and 22 hours and 7 days and 1:30 hour), with 25 neurons in the hidden layer, a MAPE of 5.76% was obtained, with a standard deviation of 7%.

Using only one entrance to foresee the load peaks daily, with 10 neurons in the intermediate layer, an error of 2.49% was obtained, with standard deviation of 2%, however, it is possible to notice in Figure 8 a trend of delay of one day in the foreseen data. Doing the forecast for the following week, the error starts to be of 7.10%, showing that the net was not well adapted to the data.

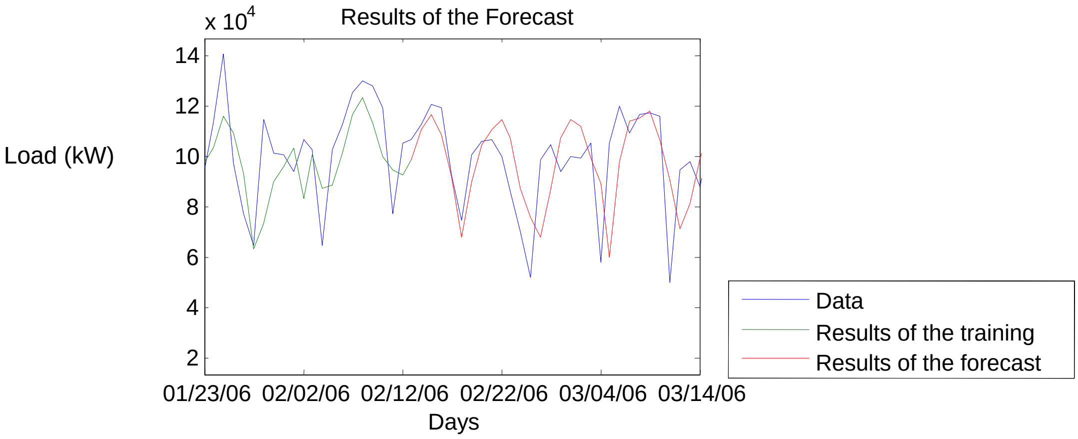

Figure 8 - Result of the Load forecasting peaks with 10 neurons in the intermediate layer

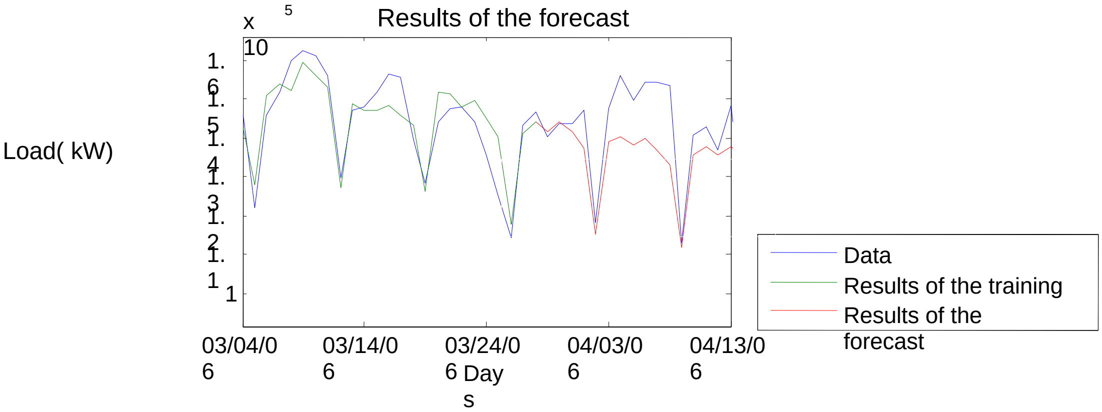

Figure 9 - Result of the prediction of the SVR with 7 entry points

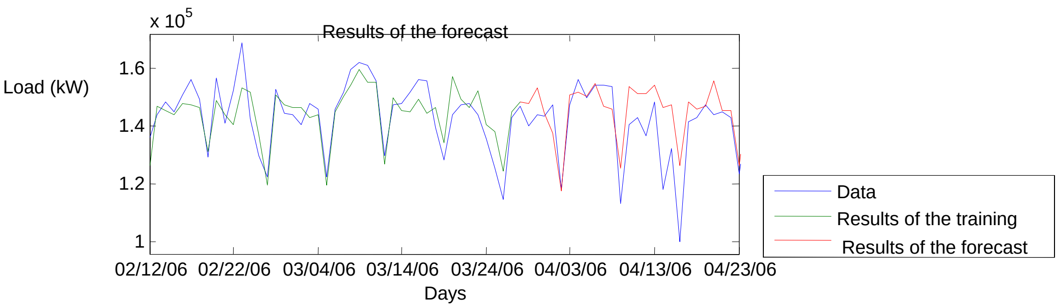

Figure 10 - Result of the forecast using 84 points

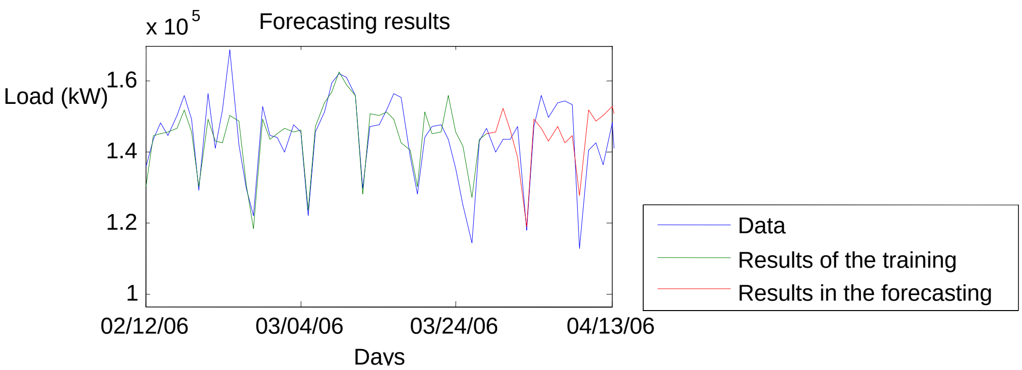

Figure 11 - Result of the forecast using the division of the data in sub-bands

Using as entered the 28 previous days to the day to be foreseen, the average forecast error is of 5.67% with standard deviation of 5%.
Using the SVR, with the corresponding entrances to the 7 previous days, it got a MAPE of 4.59% with standard deviation of 3%. However, observing the Figure 9 can be noticed that the forecast is only one repeated standard, not being therefore trustworthy.
Using 28 points as entered, the error was of 8.72%, with standard deviation of 6%. With 84 days previous as entered the error in the first week of forecast it is of 3.32%, with standard deviation of 3%. Doing the forecast one week later, the average error is of 3.02%, and two weeks later the error is of 3.17%. It can be seen by the Figure 10 that the SVR was not adapted completely to the peaks of the training data.
Doing the division in sub-bands, using 84 previous points in each SVR, the MAPE in the first week of forecast is of 3%, with 2% standard deviation. Doing the forecast in the following week, the error is of 2.42%, and in the next week the error is of 3.27%. It can be seen by the Figure the 11 that in this in case that the SVR were adapted better to the data of trainings, reaching the peak.
Table 3 presents the obtained results. Column MAPE indicates the error done in the forecast of the week after the end of the training data, and column "MAPE in the following week" after indicates the results of the forecast one week the end of the training data.

Table 3 - Results for the forecast of the daily peaks of demand

Structure	MAPE	MAPE in the following week
Neural Network, 1 entrance	2.49%	7.10%
Neural Network, 28 entrances	5.67%	-
SVR, 7 days of entrance	4.59%	-
SVR, 28 days of entrance	8.72%	-
SVR, 84 days of entrance	3.32%	3.02%
SVR, 84 days of entrance, with sub-bands	3.00%	2.42%

4 Conclusion

It was shown in this work two load forecasting methodologies using Artificial Neural Networks and Support Vector Regression.
The ANN obtained satisfactory results comparable with the published ones in literature, however its great disadvantage is the automatic definition of the net structure, that involves a great number of tests, and much experience on the part of the designer.
One form to alleviate these requirements, is the adoption of SVR technique, in which the only parameter to be adjusted is cost C, that it can be done of automatic form of complete search in a real universe, facilitating much its use in several connection points.
The division in sub-bands of the demand data, on the basis of the seasonalties of long term and seasonality of short term, was presented efficient, improving still more the performance of the system based on SVR. Such alternative could also be tested with ANN, however due to its inherent increased training time; this was not done in this work.

References:

[1] Haykin, S. *Neural Networks: a Comprehensive Foundation*, Prentice Hall PTR Upper Saddle River, NJ, USA. 1998.

[2] Elman, J.L. *Finding Structure in Time*, In: Cognitive Science, vol. 14, pp. 179-221. 1990.

[3] Kremer, S.C. *On the Computational Power of Elman-Style Recurrent Networks*, IEEE Transactions on Neural Networks, vol. 6, pp. 1000-1004. 1995.

[4] Smola, A.J., and Schölkopf, B. *A Tutorial on Support Vector Regression*, NeuroCOLT2 Technical Report Series TR NC-TR-98-030, Royal Holloway College, University of London, UK. 1998.

[5] Chen, B.J., Chang, M.W. and Lin, C.J. *Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001*, IEEE Transactions on Power Systems, Vol. 19, No. 4, Nov. 2004, pp. 1821-1830. 2004.

[8] LEE, C.C. *Fuzzy Logic in Control Systems: Fuzzy Logic Controller – Part I and II.* IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, 1990. p. 404-418.