Peak Load Forecasting using Kohonen Classification and Intervention Analysis

Francisco J. Marín[1], Francisco García-Lagos[2], Gonzalo Joya[2], and Francisco Sandoval[2]

[1] Dpto. Electrónica, E.T.S.I. Informática, Universidad de Málaga,
29071 Málaga, SPAIN
marin@ctima.uma.es
[2] Dpto. Tecnología Electrónica, E.T.S.I. Telecomunicación, Universidad de Málaga,
29071 Málaga, SPAIN
{lagos,joya,sandoval}@dte.uma.es

Abstract. This paper presents a model for short-term load forecasting using artificial neural networks and intervention analysis. The development of this model consists of two phases: a prior one, in which, starting from historical data, each day of the year is classified according to its 48 half-hour load profile by means of self-organizing feature maps; a second one, consisting on modeling the load profile for each class by intervention analysis based on statistical methods and knowledge about electrical demand curves. Once that the predicted load profile is obtained, the maximum daily electrical load (peak load) is extracted. Experimental tests have been carried out from historical data correspond to the Eastern Slovakian Electricity Corporation centered in the prediction of the 31 days of January 1999.

1 Introduction

The research group at the University of Málaga is dedicated to the analysis, modelling and implementation of several processes that are involved in the Management of an Electric Energy System. For this task, novel computing methods -known, in general, as Computational Intelligence techniques- are used. The following processes are studied: Energy Demand Prediction, emphasizing the generation of new neural networks based predictors, which are adapted to the new conditions of the energy market, as well as neuro-fuzzy predictors to deal with holidays and special periods [1]; State Estimation [2], in particular Observability Analysis (determining whether the solution of the estimation problem from the available measures is solvable), and Topological Estimation (obtaining the current topology) [3], by means of synchronized measures from phasors; and, Security Analysis, focused on Contingency Analysis (evaluating the dangerousness of a particular fault in the network from the knowledge of its current state) [4].

The issue in short-term load forecasting is to extrapolate past load behaviour while taking into account the effect of the other influential factors , such as weather, season day of the week, etc. However, the relationships between load and these factors are complex and non-linear.

Artificial neural networks (ANNs) are being applied to forecasting problems since their distributed structure of weights and neurons permits to approach complex relationships between variables without specifying them explicitly in advanced [5]. However, load forecasting models using a single neural network can not offer good performance when dealing with the complete set of days of a year. These days present profiles being dissimilar because of seasonal, economic and cultural factors. Thus, in order to avoid these restrictions, it is necessary to carry out a prior classification stage of days according to the similarity of their load profiles. A self-organized map to obtain typical load daily curves has been used [6]. In a second stage, each class obtained must be independently treated. We have used recurrent neural networks for forecasting complete curve of the 24 hourly values for the next day of the whole year, for the central Spanish area with historical data from 1989 until 1999 [1]. In the present work, the objective is to predict the peak load of the 31 days of January 1999 of the Eastern Slovakian Electricity Corporation from the historical data corresponding to years 1997 and 1998. Thus, the main problem is the very reduced number of historical data (only two years), and, besides, with different weather conditions. Thus, the recurrent neural networks (and other paradigms) can not be used. We have applied an intervention analysis that include knowledge about electrical demand curves.

2 Approach Description

The approach used for the peak load forecasting is based on the following steps: from the historical data of load, every day of the year is classified according to the similarity of their load profiles by means of a two-dimensional Kohonen map networks; next, with the obtained classes, every day have associated a characteristic profile that is adjusted with basic information on the electrical demand.

2.1 Kohonen Classification

The Kohonen algorithm [7] is a powerful self-organization process, which has the special property of effectively creating spatially organized "internal representations" of various features of input patterns, and their abstractions. The Kohonen algorithm carries out a distribution of an input space V_I in another space of smaller dimension V_M, preserving the topological relationships among the input vectors. That is, similar input vectors are distributed to close points in the output space. The output space V_M is represented by a two-dimensional array of neurons. The topology conservation is carried out by means of a non-supervised competitive learning, in which each input vector x is compared with the weight vectors w_i of each neuron in the network. The neuron whose weight vector is the nearest to the vector x is selected, modifying its weights and those of its neighbors according to equation (1),

$$w_i(t+1) = w_i(t) + l_r h_{iv}(x - w_i(t)) \quad (1)$$

where $h_{iv}()$, neighborhood function, determines the weight increment of each neuron as a function of proximity to the winner neuron. In our case, the neighborhood area is determined by a square centered in the winner neuron whose side diminishes until zero along the training; l_r is the dynamic rate learning, which evolves along the learning according to equation (2),

$$l_r(t) = \frac{l_{r0}}{(1 + \frac{c \cdot t}{nn})} \quad (2)$$

being l_{r0} the initial rate learning (0.3), c constant (0.2), t the current iteration and nn the number of neurons. The Kohonen algorithm is used with a toroidal two-dimension network (10x10 size) where both top and bottom as the left and right sides are attached.

2.2 Intervention Analysis

The intervention analysis is based on characteristics about the electrical demand. The main characteristics of electrical demand are the following.

Weekly periodicity with seasonal patterns. The load curves repeat every seven days and are dissimilar for each season. This characteristic is captured in the classification phase, as we can see in the section 4.1.

Nonstationary behavior. The load changes with the national economy conditions. It grows if economy do, and decreases with economical recession. In the time series analysis methodology for to make stationary series a differencing operator is applied. In our methodology the predicted curves must be adjusted (see section 4.2.).

Meteorological influence. It is well know than meteorological variables are the ones that have more influence on the electrical demand in annual average. The influence of the meteorological variables is more important in hours of light and the afternoon than at the first hours of the day and the sleep hours. The peak load occurs when the meteorological variables take bigger influence. That is way, the temperatures (only one available) play a fundamental role in a correct prediction (see section 4.2.).

The optimal prediction must be adjusted according to theses characteristics. Because the number of load pattern is very little, the adjust can not be performed by a neuronal paradigms. Thus we use an intervention analysis based on statistical methods.

3 Data Description

The historical data have been supplied by world-wide competition within the EUNITE network. The problem to be solved is the forecasting of maximum daily electrical load (peak load) based on electrical load values and temperatures data. The historical data are half an hour loads and average daily temperatures of the time period 1997-1998, including the holidays for the same period of time. Besides, average daily temperatures data of the years 1995 and 1996 are available, too.

The actual task is to supply the prediction of maximum daily values of electrical loads for January 1999 (31 data values altogether).

3.1 Load Values

Figure 1 shows the maximum daily load, in MW, for the years 1997 and 1998. One can see that the maximum load for the first days of January 1997 is bigger than January 1998, while that, in general, the maximum load for the rest of months is contrary. This issue can be see more clear in the figure 2. This is due to the different temperatures (table 2). December presents similar characteristics.

Table 1 shows statistical results for the daily demand (the additionadd of the 48 half an hour load) between January, December, summer months (May to August), and annual. The increase from 1997 to 1998 for December and summer months is about 4.5%, but the increase for January has minus sign, which points out that there is a decrease about 4.5%. This is due to different temperatures between January 1997 and 1998 (table 2). The maximum load for 1997 is 876 MW at 16:30 hours and for 1998 is 839 MW at 20:00 hours, in the time period of maximum daily load is between 9:30 hours to 20:30 hours.

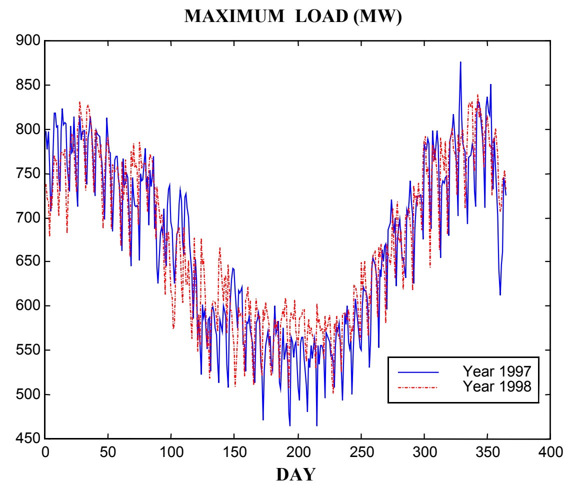

Fig.1. Maximum daily loads from time period 1997-1998.

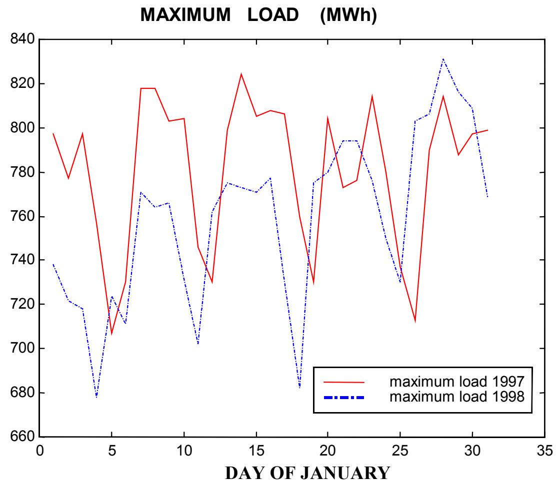

Fig. 2. Maximum load for years 1997 and 1998.

Table 1. Average daily load and its percentages of increase during several time periods. (*) minus sign mean decrease.

	Average daily 1997 (Mw)	Average daily 1998 (Mw)	Percentage increase (%)
Annual Load	28484	28766	1 %
January load	34491	32899	-4.61 % (*)
December Load	32919	34424	4.57 %
Period May-August	23453	24487	4.41 %

3.2 Temperature values

Figure 3 shows the average daily temperatures from time period 1995-1998. For the task of prediction of peak load for January 1999, we are specially interesting in the temperature data of December and January, which are shown in the table 2.

Table 2 show that January and December are the months with more low temperatures. For the whole time period the maximum is 6.7 ºC for December 1997, and the minimum is –14.2 ºC for December 1998. Figure 3 show that the evolution of the temperature is continua, without sudden changes, and that is corroborated by mean of the table 2. For this way, it is hoped that the temperatures of January 1999 should be lot like than January 1997 (with a perceptron multiplayer this issue is verified). January 1997 has more low temperatures, which explain the different curves for peak load of both years 1997 and 1998 (figure 1). That is, there is more high peak load values in 1997 than 1998. The temperatures for the rest of months are very similar for whole time period 1995-1998. This, too, can be seen in the classification phase (see section 4.1).

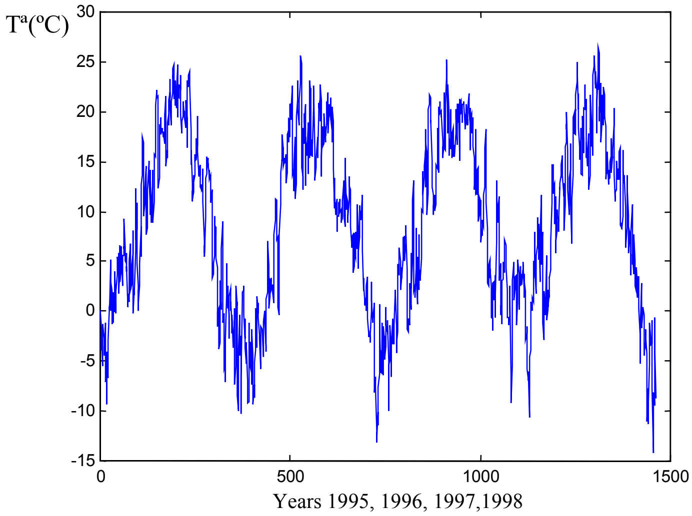

Fig. 2. Average daily temperatures from time period 1995-1998.

Table 2. Average maximum temperatures from time period 1995-1998 for January and December

	Maximum (ºC)	**Average (ºC)**	**Minimum (ºC)**
January 1995	5.20	-2.22	-9.3
December 1995	2.60	-1.71	-10.0
January 1996	2.60	-3.34	-10.3
December 1996	3.50	-3.64	-13.2
January 1997	0.7	-3.82	-10.0
December 1997	6.7	0.56	-9.2
January 1998	5.2	0.91	-7.0

December 1998	-0.6	-5.6	-14.2

4 Experimental results

o Classification

Figure 3 shows the classification by the Kohonen map of the load profiles, where the curves has been distributed by similarity between them.

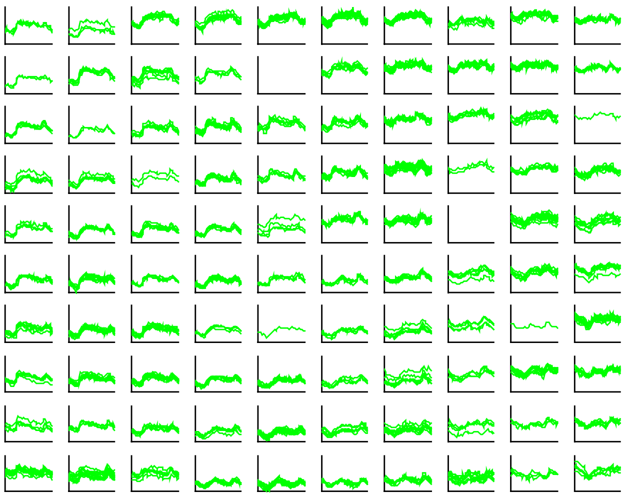

Fig. 3. Two-dimensional array of neurons. Classification of load profile.

From this classification, a second clustering process in superclasses have been carried out. We find that it is convenient to join the relations extracted by Kohonen map with the previous knowledge about the system. For example, we can see that the neuron 1 to 3 of the rows 6 to 9 in figure 3 correspond to working days of June to July. Thus, they make a natural superclass. Table 3 shows the final classification where the daily load patterns (including holidays and atypical days) were separated into seventeen different classes. Classification results incorporate, as it was expected, the interactions between weather variables and calendar variables (day of the week and seasons). In this way, the classes respond to characteristics of the different days: Sundays, Mondays, Saturdays, weekdays, and to meteorological characteristics (seasonality). It is necessary to remark that the classification is different from years 1997 and 1998, although the class for both years are adjacent (classes 1, 6, 8, 9). For summer period (classes 3, 12) and transition months (rest of classes) the classes are identical for both years. There is a specific class, class 17, for holidays in 25 and 26 December and 1 January. Thus, table 3 only shows the classes of year 1997.

Table 3. Day type identification with Kohonen network.

Class 1	Sunday from first fortnight of December, January and last week of February
Class 2	Sunday and Saturday from last week of February to May
Class 3	Sunday and Saturday in June, July and August
Class 4	Sunday and Saturday in September, October and first fortnight of November
Class 5	Sunday and Saturday of second fortnight of November to first fortnight of December
Class 6	Saturday from first fortnight of December, January and last week of February, and 6-January
Class 7	Sunday and Saturday from March to May, except Easter Week
Class 8	Monday from January to last week of February
Class 9	Tuesday to Thursday from January to last week of February
Class 10	Friday from January to last week of February
Class 11	Monday to Friday of last week of February, March and April, except Easter Week
Class 12	Monday to Friday from May to July
Class 13	Monday to Friday of August
Class 14	Monday to Friday from September to last week of October
Class 15	Monday to Friday from last week of October to last week of November
Class 16	Monday to Friday from last week of November to December
Class 17	Holidays in 25 and 26 December and 1 January

- **Intervention Analysis**

In the section 2.2. the main characteristics of the electrical demand have been numerated: weekly periodicity with seasonal patterns, nonstationary behavior and meteorological influence. The first of them have been incorporated to the prediction model, but the other ones must be contemplated by an intervention analysis. Respect to the nonstationary behavior, table 1 detects an annual increase of the electrical load about 4.5%. On the other hand, weather variables are abnormally dissimilar for our time period of study (January) regarding years 1997 to 1998. Thus, the temperatures of January 1999 should be similar to January 1997 (as January of 1995 and 1996) (section 3.2). Thus, the maximum daily loads for January 1999 should be similar to January 1997 with the increase of 9.0% (difference of two year), assuming that this time period presents equal economical conditions.

Consequently, the definitive values of the maximum daily load predicted for January of 1999 are the following:

- Friday 1 January 1999 and Saturday 2 January 1999 (class 17): 869 MW.
- Sundays 3, 10, 17, 14 and 31 January 1999 (class 1): 785 MW.
- Saturdays 9, 16, 23, 30 January (class 6): 796 MW.
- Wednesday 6 January (class 6): 796 MW.
- Mondays 4, 11, 18, 25 January 1999 (class 8): 871 MW.
- Tuesday to Thursdays 5, 7, 12, 13, 14, 19, 20, 21, 26, 27, 28 January (class 9): 881 MW.
- Fridays 8, 15, 22, 29 January 1999 (class 10): 879 MW.

To propose a more refined maximum daily load prediction for the class obtained is temerarious due to the reduced number of historical data and the very long prediction horizon (31 values), as well as the bigger errors always in peak load prediction front to load profile.

5 Conclusions

This paper describes the methodology, implementation, and results of a load forecast procedure. Using historical data of half an hour loads, every days are classified according to its load profile by means of a Kohonen map. Then, for each class obtained, an intervention analysis is performed considering the main characteristics of the electrical demand.

The performance of our proposed model was extensively assessed through many experimentations, and compared with other statistical and neural paradigms for the central Spanish area. Robustness and adaptability of our system to other regulation areas is based on the capability of Kohonen maps to extract non evident environmental, cultural and economic factors.

6 Acknowledgment

This work has been partially supported by Eliop S.A., contract No. 8.06/58.818, and the Spanish Comisión Interministerial de Ciencia y Tecnología (CICYT), Project No. TIC98-0562. Thanks are due to Red Eléctrica de España (REE).

References

1. Marin, F.J., García-Lagos, F., Joya, G., Sandoval, F.: Global Model for Short-Term Load Forecasting using Artificial Neural Networks. Accepted for publication in IEE Proc. on Generation, Transmission and Distribution
2. García-Lagos, F., Joya, Marin, F.J., G., Sandoval, F.: Hopfield neural networks for state estimation: parameters, efficient implementation and results. E&I, Elektrotechnik und Informationtechnik, Vol. 1, pp. 4-7, (2000), 4-7
3. García-Lagos, F., Joya, Marin, F.J., G., Sandoval, F.:. A Modular Power System Topology Assessment Based On Gaussian Functions ", Proceedings of the International Conference Power and Energy Systems (PES2000), (2000)
4. García-Lagos, F., Joya, Marin, F.J., G., Sandoval, F.: Neural Networks for Contingency Evaluation and Monitoring in Power Systems. LNCS 2085. Springer, (2001), 711-718
5. McMenamin, J.S., Monforte, F.A.: Short Term Energy Forecasting with Neural Networks. The Energy Journal, (1998), Vol. 19, No. 4, 43-61
6. Cottrell, M., Girard, B., Rousset, P.: Forecasting of Curves Using a Kohonen Classification. Journal of Forecasting, (1998), Vol. 17, 429-439
7. Kohonen, T.: Self-Organized Map. Proceedings of the IEEE, Vol. 78, No. 9, (1990), 1464-1480