Comparative Analysis of Self Organizing Maps vs. Multilayer Perceptron Neural Networks for Short - Term Load Forecasting

S. Valero *IEEE Member*(1), J. Aparicio(2), C. Senabre(1), M. Ortiz, *IEEE Student Member*(1), and J. Sancho
Department of Industrial Systems Engineering (1) and ORC(2)
Universidad Miguel Hernández de Elche
Edif. Quorum-V, Elche, Spain
svalero@umh.es

A. Gabaldon *IEEE Member*
Department of Electrical Engineering
Universidad Politécnica de Cartagena
Cartagena, Spain

***Abstract*— The objective of this research is to analyze the capacity of the Multilayer Perceptron Neural Network (MLP) versus Self-Organizing Map Neural Network (SOM) for Short-Term Load Forecasting. The MLP is one of the most commonly used networks. It can be used for classification problems, model construction, series forecasting and discrete control. For the forecasting problems, a backpropagation (BP) algorithm is normally used to train the MLP Neural Network. On the other hand, the SOM is a type of artificial neural network that is trained using unsupervised data to produce a low-dimensional, discretized representation of an input space of training samples in a cell map. Historical data of real global load demand were used for the research. They were presented to the two neural networks as a daily load value, in hours, from the workdays of certain years. The load curves were obtained from the Spanish Electrical System Operator. For testing purposes, new year data were also collected. The main objective of the research is to use the capacity of MLP and SOM to classify and memorize historical data, followed by taking advantage of this memorization to identify similarities between the historical data and the demand of a new day corresponding to new year data. Certain pre-processing of the input data is needed in order to obtain good results as the demand evolves over the years. It is important to establish the measurement index used to check the accuracy of the tool. This index is the Mean Absolute Percentage Error (MAPE), which measures the accuracy of fitted time series and forecasts. Different parameter configurations for the SOM and MLP training (training periods, algorithms, etc...) are also being tested to improve the behaviour of the forecast. The results of this analysis have proved the suitability of both networks to make short-term predictions, as good results are obtained when different methodologies are applied. These tools could assist Companies, Utilities and Independent System Operators (ISO) in predicting the short-term energy demand.**

Keywords: Short-term Load Forecasting, Self-Organizing Map (SOM), Multilayer Perceptron Neural Network.

I. INTRODUCTION

Industrialized countries have experienced this decade a global electricity demand growth. With the power system demand and their complexity increase, many factors have become influential to the electric power generation and consumption. Therefore, the forecasting process has become even more complex and accurate forecasts are needed. The supply industry requires forecasts with lead times that range from the short term (a few minutes, hours, or days ahead) to the long term. But the relationship between the load and its exogenous factors is complex and nonlinear, making it quite difficult to demonstrate through conventional techniques, such as time series and linear regression analysis. Short-term forecasting techniques are useful tools in the decision making to keep the generation/consumption balance. Previous researches on electrical demand and inherent factors in different countries [1] and [2] allow to identify a group of basic variables explaining the evolution of the electrical demand. Normally, the variables are common to all the industrialized countries. These variables are weather conditions, economic factors, demographical factors, calendar and random factors. Weather conditions, mainly temperature, are the most influential factors. Economic and demographical factors have an influence in a long-term forecasting but not in a short-term one. Calendar, i.e. year, month, day and hour, has a great effect on the model, although it is related to weather conditions too. Random factors such as strikes or sport events should be treated as outliers or abnormal data. Based on various time intervals, load forecasting can be divided into three main categories: short-term forecasting, medium-term forecasting and long-term forecasting. Short-term forecasting usually makes forecasts from one hour to one week, medium-term forecasting concerns the future electric load from a week to a month, and long-term forecasting often predicts the load of one year or even longer. The short-term forecasting is used for controlling and planning power generation, and also as input to study load-flow or carry out a contingency analysis [3].

II. SHORT-TERM LOAD FORECASTING WITH NEURAL NETWORKS

Neural Networks (NNs) have succeeded in several power system problems such as: planning, control, analysis, protection, design, load forecasting, security analysis and fault diagnosis [4], [5], [6]. In this sense, the NNs´ability in mapping complex nonlinear relationships is responsible for the growing number of their applications to short-term load forecasting [7]. Two different approaches are possible in the model construction of short-term load forecasting using neural networks. In [3] and [8] forecasting is done simultaneously for the whole 24 hours. These models have an output layer of 24

The work described in this paper is being supported by the Valencian Government, Department of Education under Research Project GV/2010/080.

neurons, one for each day hour that, depending on the input data shown to the trained network, generate a forecasting for the next 24 hours. Therefore, this kind of network does the forecasting process in a unique step. The inconvenience of this method is that the hours already known for the forecasted day are not used at all by the neural network, as the 24 hours are all generated simultaneously. This means that if, for instance, we want to know the load value for the 12 p.m. hour, the known load value for the hours from 1 a.m. to 11 a.m. can not be used. The networks used by this kind of methods can be taken as static forecasting models. There are other models that avoid the problem of static forecasting [6] and [9]. These models do an hour forecasting using the previous hours known. The method used for the research of this paper uses an hour to hour forecasting as the results obtained for dynamic models are better than the static ones.

III. MULTILAYER PERCEPTRON NEURAL NETWORK FOR PREDICTION

The Multilayer Perceptron (MLP) is one of the most common networks. It can be used for classification problems, model construction, series forecasting and discrete control [10]. This neural network model is developed adding hidden layers to a simple perceptron. Each layer employs several neurons and each neuron in a layer is connected to the neurons in the adjacent layer with different weights. Signals flow into the input layer, pass through the hidden layers, and arrive at the output layer. With the exception of the input layer, each neuron receives signals from the neurons of the previous layer weighted by the interconnected values between neurons. For the forecasting problems, a backpropagation (BP) algorithm is normally used to train the MLP Neural Network [11]. There are a lot of different variations of the backpropagation algorithm but the Levenberg-Marquardt BP algorithm (LMA) is perhaps the fastest and most efficient MLP training algorithm. This algorithm [12] provides a numerical solution to the problem of minimizing a function, generally nonlinear, over a space of parameters of the function. These minimization problems arise especially in least squares curve fitting and nonlinear programming. The LMA used in MLP Neural Network is much more efficient than any of the techniques when the network contains no more than a few hundred weights. Therefore, this method was used as it can assure a fast convergence when there are only a few parameters available for the network. Although the method needs more processing time, it achieves convergence only in a few iterations reaching convergence in a small total computation time.

IV. SELF-ORGANIZING MAPS FOR PREDICTION

SOM deals with the most popular artificial neural network algorithm in the unsupervised-learning category. The aim is to learn to map similar input vectors to similar regions of the output array of nodes. The Kohonen neural network is representative of a self-organizing mapping technique that allows the projection of multidimensional points to a two dimensional network. There are two important key concepts in understanding the Kohonen network; competitive learning and self organization. Competitive learning is simply finding a neuron that is the most similar to the input pattern, and the network modifies this neuron and its neighboring neurons to become even more similar to it. The advantage of SOM is that the relationship between the original vectors is to some extent preserved in the output space, providing a visual format where a human operator can “easily” discover clusters, relations, and structures in the usually complex input space database. The map consists of a regular processing unit, neuron grid. A multidimensional observation model, possibly a vector consisting of features, is associated with each unit. Previous researches demonstrate the capacity of SOM to predict time series [13] and how it is applied to load forecasting [14] and [15]. The SOM can perform efficiently this task and compete equally with well-known neural architectures, such as MLP and Radial Basis Function networks (RBF), which are more commonly used. The main advantages of SOM-based models over MLP or RBF-based models are the inherent local modeling properties, which favor the interpretability of the results.

V. CASE OF STUDY AND INPUT DATA

For the research, only weather, the date and hour of the previous days were used for training the MLP model. In the case of the SOM only historical load data on weekdays were used. Economic and population growth rates were not considered as they have an effect in the long term and the research aims for a short term one. The other data used as input for the model include the load value from previous months, days and even hours (24 hours). They were probably the most important input of the model to forecast the future load curve. They were obtained from the Spanish Electrical System Operator [16]. For the training step of both models, only 2004 data were used. Fig. 1 shows the Spanish global load demand curve of a week in September 2004.

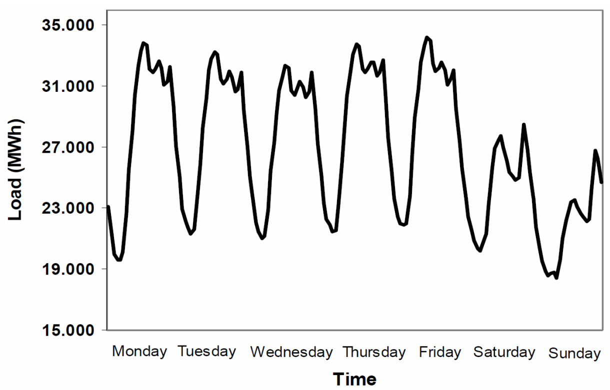

Figure 1. Global Load Demand from September 6th to 12th in 2004.

For the MLP model weather conditions were also translated into a series of global average temperatures in Spain. They were obtained from a world weather information webpage [17], where several year data are available. Fig. 2 shows the average temperature in Spain for the year 2004.

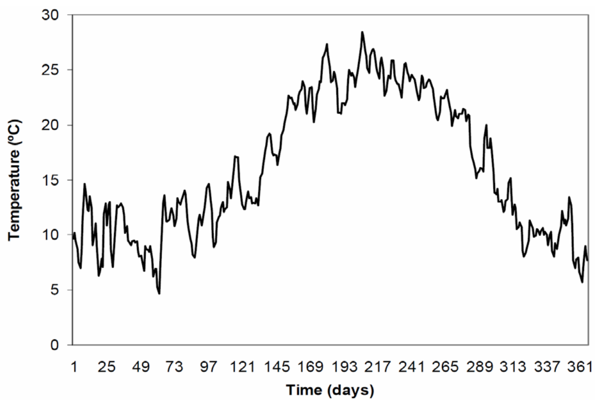

Figure 2. Global Average Temperature in Spain for the year 2004.

The research is intended to forecast one hundred days, from September 1st to December 20th in 2004, and to compare the results of both neural models. The input data used for training the MLP network are based on previous months. The SOM is trained with the daily global load curves (only working days) of the years 2002 and 2003.

VI. Error Measurement of the Forecasted Curves

Before showing the simulations of the model, it is important to establish the measurement indexes used to check the accuracy of the tool. As the model makes forecasts from hour by hour, the index to choose was the forecasting error of each hour. An error index, called MAPE, was calculated. Mean absolute percentage error (MAPE error) measures the accuracy in a fitted time series, and provides accurate forecasts. The advantage of this is that it is easy to understand and is widely used by the other referenced authors.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{|L_{Ri} - L_{Pi}|}{L_{Ri}}\right]x100\% \qquad (1)$$

Where: N is the number of forecasted hours (in this case is 16 if we use the first 8 hours of the forecasting day), L_{Ri} is the real load value of the i hour and L_{Pi} is the forecasted load of the i hour.

VII. MLP Model

A. Procedure with the MLP Model

The network used for the model was a MLP with Levenberg-Marquardt BP algorimth and with only a neuron in its output layer. This neuron gives the L(i) output value which contains the load value of the forecasted hour. As it was stated previously, the forecasting was done for one hundred days in 2004 from September 1st to December 20th. This period contains some holidays having a different load curve. In order to avoid an odd result, these days must be filtered, not only excluding them but also the previous and posterior days. Along the research, different input data and parameters setups were tested. Following points are about the three configurations that achieved the best results. In order to have a successful forecasting it is very important to have quality data for the training phase. It is necessary to pre-process the data to improve the learning ability of the network. The pre-processing process consists of the following steps:

-Filtering: This step reviews all the data looking for outliers or odd data. The outliers are found not only in load data but in temperature series too. The erroneous load data days are deleted from the input base whereas the odd temperature days are replaced with the average value of the previous and posterior data.

-Coding: In section 4 it is explained how to correctly code the training data.

-Normalizing: Either load and temperature data were normalized. The scale of the variables is changed dividing the value by the maximum. This produces more homogenous data, being 1 the maximum value. Once the network is trained, the output data obtained must be post-processed in order to get coherent values. Therefore, the post-processing process basically reverses the normalization for the output data of the training. Table I shows that in order to improve its performance, the neural network needs to know what hour of the day and day of the week is going to be forecasted. This information is given to the network by a binary code that can be seen in Tables II and III:

TABLE I. The Input Data Set for the MLP Model

MLP Training Historical Data Set	Previous 3 months to the forecasting
MLP Parameters	Hidden neurons =18, Neurons on input layer=19 and Training periods=20
One Hour Forecasting Input Load Data	- Previous 3 load data hours to the hour to be forecasted (L_{i-1}, L_{i-2} and L_{i-3}). - Load data of the hour to be forecasted of the previous day (L_{i-24}). - Load data of the hour to be forecasted of the day of the previous week (L_{i-168}).
One Hour Forecasting Input Temperature Data	- Average temperature forecasted for the day to be predicted. - Average temperature of the previous day. - The difference between the average temperature of the previous day and the average temperature forecasted for the day to be predicted. - Average temperature of the previous week. - The difference between the average temperature of the previous week and the average temperature forecasted for the day to be predicted.
One Hour Forecasting Input Information	- Five binary digits that mean the hour to be forecasted. - Four binary digits that mean the day to be forecasted.
Output data	Forecasted load for the i hour (L_i)

TABLE II. Binary Code for Days

Days	Binary Code
Monday	0001
Tuesday – Friday	0010
Saturday	0100
Sunday	1000

TABLE III. BINARY CODE FOR HOURS

Day Hour	Codification
1	0001
2	0010
. .	. .
24	1000

B. *Results with the MLP Model*

Fig. 3 demonstrates predicted and real load curves of 10/29/2004. The medium error of the total 100 days is shown in Table IV.

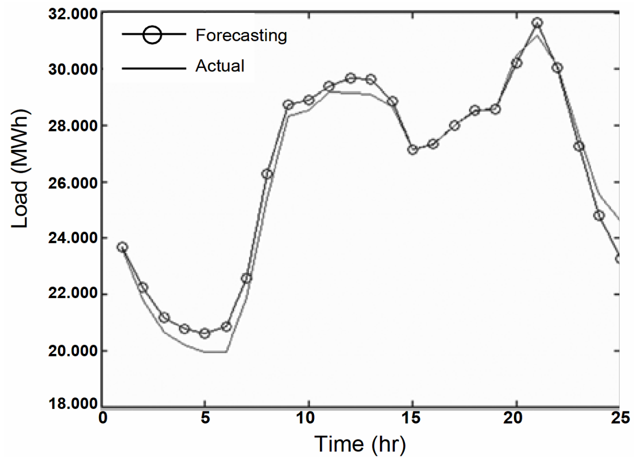

Figure 3. Load Forecasting of October 29th with the MLP Model.

TABLE IV.

MAPE (Average of 100 days)	Best Value
1.97 %	1.11%

VIII. SOM MODEL

A. *Procedure with the SOM Model*

A Self-Organizing Map (SOM) is a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional, discretized representation of an input space of training samples in a cell map. This methodology was introduced by Kohonen two decades ago. The advantage of SOM is that the relationship between the original vectors is to some extent preserved in the output space, providing a visual format where a human operator can “easily” discover clusters, relations, and structures in the usually complex input space database. The map consists of a regular processing unit, neuron grid. A multidimensional observation model, possibly a vector consisting of features, is associated with each unit. The software used as a processing tool for the map creation was Matlab. It is based on training a SOM with the daily global load curves (only working days) of the years 2002 and 2003. After the training, the map is tested with the load curves of some days of 2004. In this second step only the first 8 hours of the day are considered to identify as a result of the most similar days to the tested day. The input data are labelled with the date of the load curve in the format -*mmddyy*-, month, day and year. For each 8 data vector the neural network gives a winner cell. This cell contains one or more days of the original trained map, which are indeed the most similar days to the tested vector and therefore, can be used to define the rest of the studied day. Fig. 4 shows a simple diagram of the methodology.

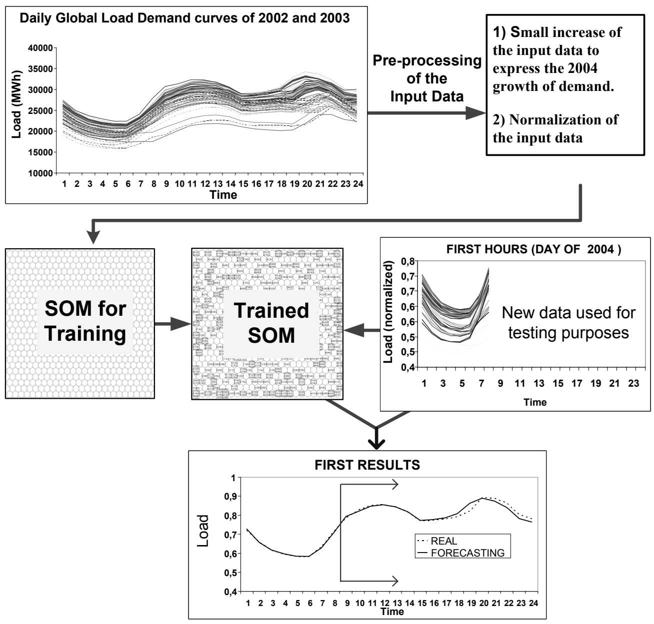

Figure 4. SOM Methodology of the research.

B. *Results with the SOM Model*

Fig. 5 identifies the first 8 hours of the 23rd of September 2004, which is the used testing day, and the curves of the winner cell containing the labels of the 25th and 26th of September 2003. Fig. 6 shows the average demand load curve of both days vs. the real curve of the 23rd of September 2004.

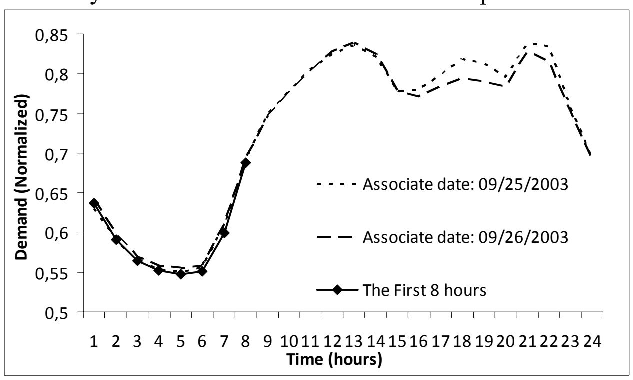

Figure 5. Results of the load curves identified by the SOM with the first 8 hours of the 23rd of September 2004.

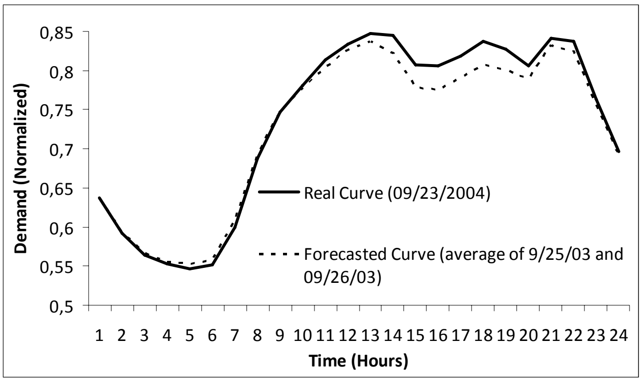

Figure 6. Real load demand vs Estimated load curve (Average of the winner cell).

The medium error of the total 100 days is shown in Table V.

TABLE V.

MAPE (Average of 100 days)	**Best Value**
1.81 %	0.87%

IX. CONCLUSIONS

With the MLP model and after several months of trainings and simulations with different training and input patterns, and matching results with their real value, it can be concluded that it is possible to accurately forecast the global load curve of a day from the load and temperature data of the previous months. This methodology obtains a small error when forecasting one hundred days. It is important to clarify that the selection of the number of hidden neurons, training epochs and an adequate training algorithm is critical. Levenberg-Marquardt Back Propagation algorithm helped to obtain a fast convergence. The great inconvenience of the model is the risk of falling into a local minimum which would totally spoil the output load curve. Finally, it is important to use the model depending on the year season, as weather is an influential variable. The results obtained through the SOM model show a low error index after the comparison of the estimated days with the real data of 2004 simply using historical data of the years 2002 and 2003. Several parameter setups were tested which finally achieved an optimum combination that reduced the error to a 1.81 % average. Nevertheless, the research objective from now onwards is to test new training configurations with gradually more complex input data, for instance adding weather conditions such as temperature, in order to improve the error indexes. Both tools are useful as they help the electrical system operator on the decision making and allow to schedule the short-term power generation to be able to cover the incoming demand. Future researches will be based on a comparative study of the results with other artificial intelligence techniques such as Logical Fuzzy and other neural networks types.

ACKNOWLEDGMENT

This research was financed by the Research Project with reference GV/2010/080-Valencian Government, Department of Education, and directed by Dr. Sergio Valero Verdú.

REFERENCES

[1] P. Murto, “Neural Network Models for short-term load forecasting”, (Department of Engineering, Physics and Mathematics. Helsinki University of Technology), 1998.

[2] J. Bao, “Short-Term Load Forecasting bases on neural network and moving average”, (Artificial Intelligence Lab, Dept. of Computer Science Iowa State University), 2000.

[3] K. Y. Lee, Y. T. Cha, and J. H. Park, “Short-Term Load Forecasting Using An Artificial Neural Network,” IEEE Trans. on Power Systems, vol. 7, pp. 124-132, Feb. 1992.

[4] H. Mori, “State-of-the art overview on artificial neural networks in power systems”, (Eds: IEEE Catalog no. 96TP112-0, pp. 51-70, 1996.

[5] M. Asari, B. Kermanshahi, “Application of Neural Network on Winter Peak Load Forecasting of Distribution Feeders”, Proc. of IASTED International Conference on Power Systems & Engineering, Wakayama, Japan, pp. 11-14, September 1994.

[6] Tarik Rashid, M.-T. Kechadi and B.Q. Huang, “Short-Term Energy Load Forecasting Using Recurrent Neural 184 Network”, The 8th IASTED International Conference on artificial intelligence and soft computing, Marbella, Spain, vol. 1, pp. 451-150, September, 2004.

[7] Jesus Riquelme Santos, Jose Luis Martinez Ramos, Antonio Gomez Exposito, Daniel Cros, “Possibilities of Artificial Neural Networks in Short-Term Load Forecasting”. Proceedings of the Iasted International Conference: Power and Energy Systems. Iasted International Conference on Power and Energy Systems, pp. 165-170, 2000.

[8] C.-N. Lu, H.-T. Wu, S. Vemuri, “Neural network based on short-ter load forecasting” Power Systems, IEEE Transactions. vol. 8, issue: 1, pp. 336-342, 1993.

[9] D.C. Park, M.A. El-Sharkawi, R.J. Marks, L.E. Atlas, M.J. Damborg, “Electric load forecasting using an artificial neural network”, Power Systems, IEEE Transactions, vol. 6, issue 2, pp. 442-449, May 1991.

[10] D. Hush, C. Abdallah, B. Hore, “Model following using multilayer perceptrons Decision and Control”, Proceedings of the 29th IEEE Conference, pp. 1730-1731, Dec. 1990.

[11] M. Kazeminejad, M. Dehghan, M.B. Motamadinejad, H. Rastegar, “A New Short-Term Load Forecasting Using Multilayer Perceptron Information and Automation”. ICIA 2006. International Conference, pp. 284-288, Dec. 2006.

[12] Kenneth Levenberg (1944), “A Method for the Solution of Certain Non-Linear Problems in Least Squares”. The Quarterly of Applied Mathematics, 2, pp. 164–168.

[13] Guilherme A. Barreto, “Time Series Prediction with the Self-Organizing Map: A Review”, Studies in Computational Intelligence. Perspectives of Neural-Symbolic Integration. Springer Berlin-Heidelberg, vol. 77, pp. 135-158, 2007.

[14] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta, V.O. Cencellli, “A neural network based technique for short-term forecasting of anomalous load periods”, Power Systems, IEEE Transactions, vol. 11, issue: 4, pp. 1749-1756, 1996.

[15] M. Farhadi, S.M.M. Tafreshi, “Effective model for next day load curve forecasting based upon combination of perceptron and Kohonen ANNs applied to Iran power network”, Telecommunications Energy Conference, p.p. 267-273, 2007.

[16] REE, *Red Eléctrica de España*, www.ree.es.

[17] World Weather Information, www.tutiempo.net.