Load Forecasting based on Neural Networks and Load Profiling

J. C. Sousa, L. P. Neves and H. M. Jorge, *Member, IEEE*

Abstract-- **This work presents a novel perspective of load forecasting based on neural networks and load profiling. In addition to the variables that are typically used to predict future load demand, such as past load values, meteorological variables, seasonal effects or macroeconomic indexes, it is expected that the use of load profiles and detailed information of individual consumers could favor the forecasting process. The methodology can be extended to different temporal horizons being predicted and the eventual threat of overparametrization is attenuated by the use of neural networks since the complexity of the model does not necessarily depends on the number of its weights and biases, as some of these parameters might be found irrelevant in the process. Another way to reduce the risk of overparametrization and overfitting is through the use of a considerable number of data points (whenever historical data is available) to train the network.**

Index Terms-- **Load Forecasting, Load Profiling and Neural Networks**

I. INTRODUCTION

The deregulation of electric power markets has effectively brought new and varied types of challenges, being actually well proven that special attention has being paid to issues such as load forecast and electricity consumers classification.

Load forecasting has ever been understood as a key issue in electricity sector, since it helps to provide a physical equilibrium between the supply and the demand, being essential in order to support an analysis of an eventual strengthening or expansion of the existing infrastructure and implementation of a maintenance scheduling (in a long-term management to guarantee a reliable system operation). It is also important to provide an optimized network configuration, being valuable to unit commitment or increasingly more helpful to plan the integration of dispersed generation (in a short-term perspective) knowing that electric energy must be always available at any period of time and at a minimum price.

Assisting to the worldwide trend of the sector liberalization, load forecasting gains even more importance, not only for system operators but also for any other participants, so that adequate energy transactions could be scheduled and forecasting errors reduced because it would result in increased operational costs and loss of revenue. Several types of methods has been applied to forecast future demand, such as time series models [1], [9], [11], [12] (using linear, polynomial or exponential regression), based on previous records of the variable being predicted and/or based on exogenous factors that can directly or indirectly affect the load behavior, in which the variable can be estimated as a function of meteorological conditions, economic activity, macroeconomic indexes (consumer price index or average salary earning), day type or electricity price. Recent load forecasting research adopts artificial intelligence and the neural networks models are undoubtedly the most popular strategy in this context [1], [2], [11], [12].

On the other hand, the consumer's classification is actually mandatory for many electricity markets, being crucial to enable the market participation of all consumers, even those who have not installed hourly-metering systems. The use of load profiles for each consumer's group characterizes different kinds of electricity usage and makes the settlement between distributors and suppliers possible, as the total energy record available in the traditional meters can be distributed by different interval periods. The information given by load profiles may be useful to retail companies as it helps to establish contracts with different kinds of consumers, with specific characteristics, practicing adjustable prices for classes and offering added-value services adjustable to each customer group. In the distributor's perspective, this detailed information may also have relevance in load management and also in electric network access tariffs definition.

This work intends to evaluate the eventual contribution of consumer classification in load forecasting thematic, adopting neural networks to capture relationships between the variable(s) being predicted and some specific information (inputs in the forecasting model) known to have strong correlation with the variable(s) *a priori* unknown. The main objective is to integrate load profiles information in the input variable space, evaluate the methodology and comment the results obtained.

This article is organized according to the following description: the next section will describe in more detail the

J. C. Sousa is with Polytechnic Institute of Leiria, School of Technology and Management, Campus 2, Morro do Lena, Alto do Vieiro, 2411-901 - Leiria (e-mail: jcsousa@estg.ipleiria.pt).

L. P. Neves is with Polytechnic Institute of Leiria, School of Technology and Management, Campus 2, Morro do Lena, Alto do Vieiro, 2411-901 – Leiria (e-mail: lneves@estg.ipleiria.pt).

H. M. Jorge is with the Department of Electrical Engineering and Computers, University of Coimbra, Pólo II, Pinhal de Marrocos, 3030-290 Coimbra (e-mail: hjorge@deec.uc.pt).

All authors are also with R&D Institute INESC Coimbra, Portugal

978-1-4244-2235-7/09/$25.00 ©2009 IEEE

load forecasting methods focusing on neural networks. Section 3 is dedicated to load profiling, advantages and motivation to this strategy and also to synthesize different ways used to segment electricity consumers based on the energy use. In Section 4 the case study will be presented with detailed description of different stages associated with the proposed methodology. Section 5 presents some results obtained and also includes comments and comparisons between different forecasting methods used. The last section points out the principal conclusions and topics for further work.

II. Load Forecasting

The use of artificial neural networks (ANN) in load forecasting is actually pointed out as an alternative or, in some cases, a complement to the traditional time series models. Falling into the group of artificial intelligence techniques, the ANNs have been explored since the middle 80's, with tendency to gain more adepts and proving its accuracy to some sceptic researchers [1]. Some of the principal advantages often commented are its ability to model nonlinear and complex relationships better than traditional linear models, providing an easy way to deal with multivariate models and exploring an automatic mapping of the relationships only with the presentation of the input(s) and output(s) (procedure known as the learning process). The skill to deal with noisy data is also commented as an advantage compared with traditional regressive methods [11], [12].

The disadvantages that can be described are related with the architecture chosen, normally too large to accommodate a large number of parameters that have to be predicted based on a small number of data points (when few historical data is available). Besides this shortcoming, it is often alleged that methods are not systematically tested and the results are not well commented and/or not presented in a clarified way.

A. Artificial Neural Networks – Architecture, Activation Functions and Training Algorithms

The multilayer feed-forward architecture is still the preferred model in load forecasting applications, so it was also used in this work. In order to avoid unnecessary heavy and complex structures, it was considered just one hidden layer of neurons between input and output layers. Different numbers of neurons in the hidden layer were tested to compare different architectures, as it will be explained later in Section 4.

Fig. 1 gives a simple example of a neural network with 2 neurons in the hidden layer to relate one output (y) with respect to three independent variables (x_1, x_2 and x_3). This example is illustrated to exemplify the connections (with associated weights) between the elements (inputs/neurons and neurons/output). Each connection is simply characterized by a weight, for example, the connection between input x_1 to the second neuron is defined by the weight w_{12}.

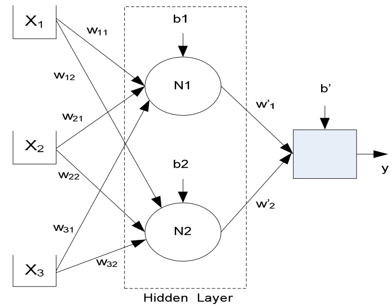

Fig. 1. Example of an ANN architecture with 2 neurons in the hidden layer

Each neuron as the basic element of the architecture processes the information captured in its inputs, such as outputs from the inputs or from the previous hidden layer already manipulated by the weight connections, and also includes a scalar bias [10], [18].

Giving the example of the first neuron, it would then produce the following output:

$$\mathrm{output}(n_1) = f_1(x_1 \cdot w_{11} + x_2 \cdot w_{21} + x_3 \cdot w_{31} + b_1) \qquad (1)$$

It must be noted that function f_1 is the activation function adopted. Typically the choice of activation functions depends on the forecasting case and the most popular functions are hyperbolic tangent function (*tansig*), sigmoid function and linear transfer function. The adoption of non-linear functions is essential in cases of non-linear relationships between outputs and inputs.

The output y is obtained applying the same methodology to the output layer:

$$y = f_2[\mathrm{output}(n_1) \cdot w'_1 + \mathrm{output}(n_2) \cdot w'_2 + b'] \qquad (2)$$

The activation function f_2 can be distinct from activation functions applied in previous layers.

The historical data assumes an important role in network training period, because it gives to the network the possibility to change its parameters (weights and biases) in order to minimize deviations between real values and estimations made by neural network. Several training methods are commonly used and, working as iterative optimizing methods, the objective is to minimize the performance function (typically based on mean square error). The backpropagation algorithm, typically the most common method, is used to calculate the gradient of the network error (E) with respect to the network's modifiable weights and biases. Each weight and bias is updated using an iterative process of gradient descent [10]:

$$\Delta w_{ij} = -\varepsilon \cdot \frac{\partial E}{\partial w_{ij}} \qquad (3)$$

$$\Delta b_j = -\varepsilon \cdot \frac{\partial E}{\partial b_j} \qquad (4)$$

where ε is the learning rate used to be multiplied with the negative of the gradient to determine the changes to weights and biases. The learning rate controls the training speed, being a faster algorithm with a large learning rate, however it could compromise convergence. On the other hand, small learning rates would make training considerably time-consuming. The training algorithm stops whenever one of the following conditions is verified [18]:

- a maximum number of epochs is attained;
- if the performance function is below a predetermined value;
- if the magnitude of the gradient is less than a predetermined value;
- a maximum training time was reached.

Another way to interrupt the training period is using the cross-validation technique with the complete data set available divided into training sample and validation sample. The training sample is used to adapt weights and biases while the validation set is used to avoid overfitting, or generalization inability. It is expected that during the training period, the error associated with training sample is always being reduced but whenever the error in validation set starts to degrade, it means that ANN continuously get more accuracy to characterize the training set although it doesn't represent an effective accuracy gain to characterize new data (validation set).

A more recent training method is the Levenberg-Marquardt algorithm and while backpropagation algorithm is a steepest descent algorithm, Levenberg-Marquardt is an approximation to Netwon's method.

For the Newton method it must be obtained a Jacobian matrix *J*, containing the first derivatives of the performance function in respect of all network weights and biases. The weights and biases update respects expression (5).

$$\Delta x = \left[J^T \cdot J \right]^{-1} \cdot J^T e(x) \qquad (5)$$

with:
Δx – vector with network weights and biases updates;
J – Jacobian matrix;
I – Identity matrix;
e(x) – vector with network errors (associated with all network weights and biases, already considering a backward propagation of the output(s) error(s)).

Levenberg-Marquardt algorithm incorporates an adaptive parameter to keep looking for good solutions in a faster way but also preserving convergence [3],[10].

$$\Delta x = \left[J^T \cdot J + \mu \cdot I \right]^{-1} \cdot J^T e(x) \qquad (6)$$

The adaptive parameter μ is multiplied by some factor β whenever a step results in an increased performance index. When a step reduces the performance function, the adaptive parameter μ is divided by β. With large adaptive parameter, the algorithm becomes similar to a gradient descent algorithm with a small step size. Considering a reduced adaptive parameter the algorithm tends to approximate to Newton's method. The algorithm therefore provides a compromise between the speed of Newton's method and the guaranteed convergence of steepest descent.

In this algorithm the specifications for the training period are the maximum number of epochs; a performance goal; a minimum performance gradient; maximum number of validation failures; maximum training time; the adaptive parameter μ; associated μ increase and decrease factors as well as a maximum adaptive parameter.

B. Input variables usually adopted in neural networks for load forecasting

In bibliography it is common to find out some particular interest in the selection of the input vector in ANN architectures [4], using past load values identified by a justified entropy analysis (to discriminate the number of contiguous data values) and/or by a trend concept (to discriminate homologous periods, e.g. similar days of previous weeks), or using exogenous factors such as macroeconomic variables, weather and seasonal variables. Recent approaches are also exploited to ascertain the most influent variables, e.g. using principal component analysis [13] to reduce the original input space to several characteristic variables or using support vector machines, an alternative to neural networks, that implements a structural risk minimization principle in regression thematic [14]. There are some causes for unexpected and sudden load fluctuation, as the presence of special days like holidays (and normal days surrounding holidays) or recording failures. These anomalous load periods and some forms to minimize associated load forecasts deviations are well described in some works [4], [5], [15], [16]. The holidays can be treated imposing *a priori* that input patterns used to estimate the demand in those days must be affected by a reduction factor. Other suggestion is a binary variable inclusion in input vector, in order to give to the ANN a clear distinction between normal and special days. This effect brings two direct advantages: the ANN notes that holidays' forecasts are attenuated (reflecting less demand in these special days) and forecasts on normal days after the holidays should not depend blindly on holidays' data. References [5] and [16] propose an unsupervised stage providing a classification of the historical data with Kohonen Self Organizing Maps, and afterwards this prior classification leads to separate learning processes of ANNs to forecast future demand. This supervised stage performs ANN training in data previously separated (typically putting in evidence the occurrence of special days) rather than on the entire data set.

The present work aims to explore how can be interpreted the effect of integrate load profiles for different consumers classes in load forecasting issues and how this assumption can also be helpful to deal with forecasting problematic triggered by special days.

III. Load Profiling

In competitive markets the detailed knowledge of how the electric energy is hourly distributed (quarter-hourly distributed under Portuguese legislation) is actually mandatory to different consumer groups. According to the new policies adopted by energy markets, all customers should have access to the retail market, even small customers who are not subjected to hourly based load measurements. To enable that participation, the load profiling is a way to obtain typical load shapes adjusted to specific groups of consumers. The load profiles elaboration presupposes a monitoring stage of a representative sample of consumers through a temporarily installation of hourly metering systems. The consequent classes' segmentation can be done following different strategies:

- The most accurate approach is the use of clustering algorithms, such as fuzzy c-means or competitive neural networks [6]. It results in load profiles for each class, and the clustering algorithm ensures homogeneous groups of consumers (in respect of energy distribution for a specific period) and also relevant differences between groups formed. In load profiling issues, it is assumed that the use of clustering algorithms is the ideal way to segment ways of electric usage, however afterwards it is essential to allocate those profiles to consumers out of sample following specific commercial or activity sector attributes, but in most cases this task is not straightforward [17];

- Use of profiles for each activity sector (residential, commercial, industrial, services, …) [7];

- Use of load profiles with a segmentation based on commercial information.

Following one of the previous ways of consumers' segmentation based on energy use, the settlement between distributors and suppliers (being mandatory) becomes possible as the total energy record available in the traditional meters can be distributed for different interval periods.

Load profiling in Portugal [19] uses a classification of Low Voltage (LV) consumers (with *rms* voltage between phases below or equal to 1 kV) based on contracted power and annual energy consumption (see Table I), respecting the principle of non-discrimination on the basis of energy use, imposed by the local regulatory authority.

TABLE I
Description of Low Voltage Classes
Adopted by the Portuguese Legislation

Class number	Class	Class Description
1	Special Low Voltage	Contracted Power > 41,4 kW
2	Normal Low Voltage – Class A	Contracted Power > 13,8 kVA but also Contracted Power ≤ 41,4 kW
3	Normal Low Voltage – Class B	Contracted Power ≤ 13,8 kVA and Annual Consumption > 7140 kWh
4	Normal Low Voltage – Class C	Contracted Power ≤ 13,8 kVA and Annual Consumption ≤ 7140 kWh

Low Voltage – RMS Voltage between phases ≤1 kV

The load profiles for these low voltage classes can be examined in more detail in Fig. 2.

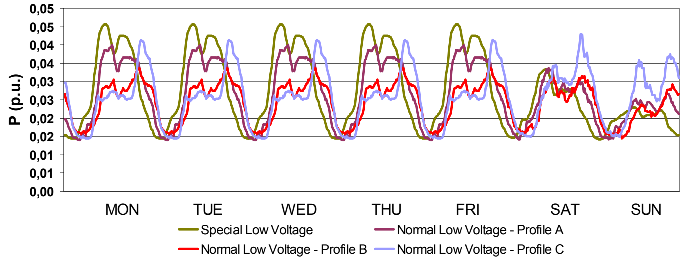

Fig. 2. Low Voltage Load Profiles approved by the Portuguese Legislation (based on one complete week of October 2005)

The profiles are prepared in a monthly basis, so in a simple way, it is considered that during each month:

- all workdays assume a similar load shape;

- saturdays and sundays have particular profiles and also different from each other;

- in the presence of a holiday, the associated profile assumes a profile similar to a Sunday.

A. Integrating Load Profiling information in Load Forecasting Thematic

It was already assumed that load forecasting and load profiling are two subjects that have been treated independently, however from a manager point of view this separation can be irrelevant and even unnecessary [8]. The load profiling is actually viewed as a way to enable the participation of the whole set of consumers in liberalized markets, and is particularly dedicated to small consumers that have not installed hourly metering systems. Load profiling simply used as billing purposes allow the settlement between distribution operator and suppliers in the market, based on the typical distribution of energy, given by load profiles, during the settlement period. Once the consumer's usage of electricity is estimated in load profiles, this knowledge could be transposed to load forecasting issues (based on a bottom-up approach).

To investigate load profile information effects in load forecasting it was considered the load profiles approved by Portuguese legislation, because this information is easily available, as it is made public by the independent regulator (before the inclusion of weather and seasonal effects – approved for the whole civil year) and by the transmission operator (after the inclusion of seasonal and weather effects).

Bearing in mind that simple assumptions are really assumed with this methodology, whereas consumer segmentation is only based on commercial information and profiles don't give an expressive difference between day types. Even so, being proved that the methodology is successful with this simple assumption, it is expected that the adoption of the alternative based on a more accurate consumer's segmentation and with different load profiles associated, would enhance the effectiveness of the strategy adopted.

IV. Case Study Description

This study uses consumption data collected from a sample of 1147 Low Voltage (LV) consumers with an integration period of 15 minutes. Besides this consumption data, a detailed commercial database is available for all consumers composing the sample. A random aggregation of a fixed number of 740 contributors' consumers has led to an accumulated load diagram composition that was used to simulate a public substation typical load diagram. The electricity consumers monitoring was implemented in scope of a national project aiming to establish load profiles and create all necessary and mandatory conditions to allow an open market. These audits were available from 23-Mar-2003 until 23-Apr-2007, but the period effectively considered was from 16-Mar-2004 until 16-Mar-2006 (two complete years) because in this temporal horizon the consumption data was quite stable for the considered consumers without significant data error detections.

It was considered the consumption distribution through different classes along the considered period. For each hour, a description of the contributors' consumers was analysed and confronted with the commercial database where historical annual consumptions are available and therefore, during the considered period, total consumption can be distributed through different classes as it can be analysed in Fig. 3.

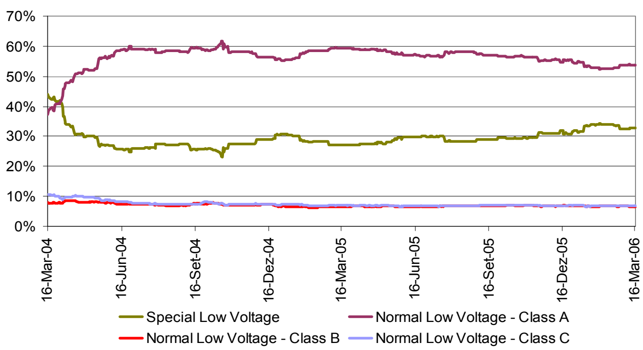

Fig. 3. Total consumption distributed through different classes for the considered period

The annual profiles available and used in Portuguese market were explored to obtain a reconstructed load diagram for each class according to (7).

$$LD_{r\_class\,i\_hour\,h} = W_t \cdot CP_{class\,i\_hour\,h} \cdot \frac{LP_{class\,i\_hour\,h} \cdot CP_{class\,i\_hour\,h}}{\sum_{h} \sum_{i=1}^{4} LP_{class\,i\_hour\,h} \cdot CP_{class\,i\_hour\,h}} \tag{7}$$

$$i = 1, \ldots, 4$$

where:

$LD_{r\_class\ i}$ is the reconstructed load diagram for class i;
W_t is the total consumption estimated for aggregation;
$LP_{class\ i}$ is the load profile for class i;
$CP_{class\ i}$ is the consumption percentage of class i;
h is each considered hour [between hour 0 of 30-Mar-2004 until hour 23 of 16-Mar-2006].

The profiles contribution for each reconstructed load during one week can be exemplified in Fig. 4.

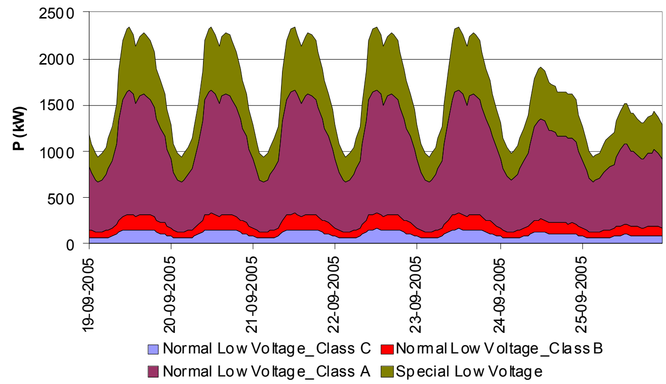

Fig. 4. Reconstructed load diagrams obtained for each class – aggregation effect during one week of September 2005

In Fig. 5 it is possible to confront the aggregated reconstructed load diagram with the real one during one week.

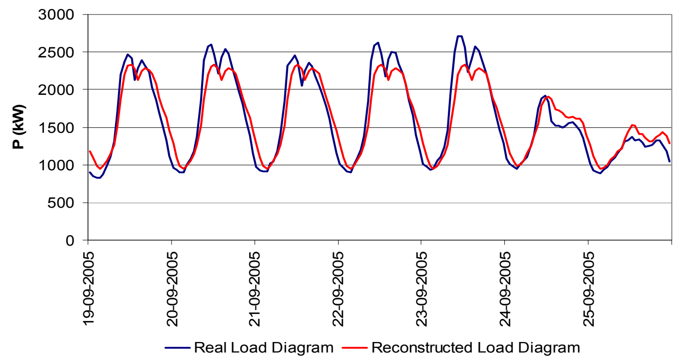

Fig. 5. Comparison between real and aggregated reconstructed load diagram during one week of September 2005

The assumption of similar workdays assumed by profiles leads to a considerable deviation from the real aggregated load diagram, as consumption patterns effectively vary according to day type, mainly in non-domestic consumers. Despite some eventual inaccuracy associated with the reconstructed and aggregated load diagram, this information was used as ANN inputs to forecast future load demand.

A. Models created to load forecast using Artificial Neural Networks

The present work intends to evaluate and compare different models to forecast electric load for the 24 hours of the following day. The models were trained from 30th March 2004 until 31st December 2005, while during the period from 1st January 2006 and 16th March 2006 the forecasting methodology was just tested and compared. During this testing period the network was just simulated.

The models were based on ANN structures following some specifications:

- all networks created do not have more than one hidden layer to avoid unnecessary complex models;

- all networks created use a *tansig* transfer function between input and hidden layers and a linear transfer function between hidden and output layers.

- all networks were trained using the Levenberg-Marquardt algorithm, knowing that it congregates reduced training time with demonstrated skills to guarantee convergence;

- for each model it was considered and evaluated a variable number of neurons in the hidden layer (from 3 to 15 neurons);

- for each model, and for different number of neurons in the hidden layer, twenty scenarios of network initialization were tested and compared, bearing in mind that each different scenario creates distinct initial random network parameters (weights and biases);

- for these scenarios, it was also used a fixed and a random validation set in order to interrupt the training period in different ways. In both cases data available from 30th March 2004 until 31st December 2005 was split in the following way: 20% for validation data and 80% for training data;

- network training parameters were defined to have a maximum of 1000 epochs, with a minimum gradient of 10^{-10}, with an adaptive parameter μ adjusted to 0.001, a increase/decrease factor β set to 10/0.1 and with no training time limit. All simulations were effectively stopped by the cross-validation criterion.

Different models incorporating distinct input patterns were evaluated. The inputs and outputs were normalized in order to have zero mean and a unitary variance. In that way the influence of different scales are attenuated. A brief description of models created is presented here:

- Model I: this model is based on studies identified in the bibliography and uses 72 inputs such as hourly load values of the last day available, of the same weekday of previous week and same weekday two weeks before, considering these ones the most correlated with variable(s) being predicted. For example, to forecast Thursday 15th September 2005 hourly load values (being estimated during Wednesday 14th September and with that day not entirely known), the model will use hourly load values of Tuesday 13th September 2005, Thursday 8th September 2005 and also Thursday 1st September 2005.

- Model II: this model with 75 inputs incorporates in Model I three more inputs: one binary input to distinguish normal days and special days (like holidays) and two cyclical inputs to characterize day types. Cyclical functions such as *sin* and *cos* were adopted, and the need to deal with these two functions is to guarantee that each day has a univocal representation.

- Model III: besides the 72 inputs of Model 1 it is integrated the 24 hourly load values obtained from the reconstructed load diagram of the day being predicted.

The comparison between several networks created for each model was possible using different error measures for training sample and also for testing sample. The error measures used in this work and normally used in similar studies are:

- Mean Percentage Error (MPE):

$$\text{MPE }(\%)=\frac{\sum_{h}\left(\frac{\text{Real Load Value}_h-\text{Forecasted Load Value}_h}{\text{Real Load Value}_h}\right)}{\text{Number of hours considered}} \qquad (8)$$

- Mean Absolute Percentage Error (MAPE):

$$\text{MAPE }(\%)=\frac{\sum_{h}\left(\frac{\left|\text{Real Load Value}_h-\text{Forecasted Load Value}_h\right|}{\text{Real Load Value}_h}\right)}{\text{Number of hours considered}} \qquad (9)$$

- Root Mean Squared Percentage Error (RMSE):

$$\text{RMSE }(\%)=\sqrt{\frac{\sum_{h}\left(\frac{\text{Real Load Value}_h-\text{Forecasted Load Value}_h}{\text{Real Load Value}_h}\right)^2}{\text{Number of hours considered}}} \qquad (10)$$

- Mean Absolute Deviation (MAD):

$$\text{MAD }(\%)=\frac{\left(\frac{\sum_{h}\left(\text{Real Load Value}_h-\text{Forecasted Load Value}_h\right)}{\text{Number of hours considered}}\right)}{\left(\frac{\sum_{h}\text{Real Load Value}_h}{\text{Number of hours considered}}\right)} \qquad (11)$$

V. Results

Some of the fundamental results to be presented are the error measures associated with different models analyzed. The errors were also detailed for training period and testing period, as appropriate neural networks must be able to generalize for data not used in the training stage. The results shown in Table II were chosen from several network simulations, being the best results found in each scenario.

TABLE II
COMPARISON BETWEEN DIFFERENT MODELS
DIFFERENT ERROR MEASURES FOR TRAINING AND TESTING DATA

	Error measures	MPE	MAPE	RMSE	MAD
Model I	Training Data	-0.76%	5.60%	9.00%	5.60%
	Testing Data	1.93%	7.09%	11.47%	7.25%
Model II	Training Data	-0.74%	4.94%	6.74%	4.96%
	Testing Data	0.71%	5.54%	7.26%	5.83%
Model III	Training Data	-0.36%	4.26%	5.69%	4.09%
	Testing Data	1.95%	4.96%	6.76%	5.08%

Comparing model II with model I, both training and testing data denote a considerable reducing in forecasting errors. It is also evident that the adoption of model III makes the forecasting procedure even more accurate. Afterwards a special attention was dedicated to forecasting problems caused by special days. Bearing in mind that the load profiles already include information about these special days with the subsequent consumption reduction effect (probably with some simple assumptions because holidays profiles are assumed similar as Sundays profiles), this information can thus be passed to the load forecasting process. Fig. 6 exposes the comparison of different error measures associated with special days for the models analyzed.

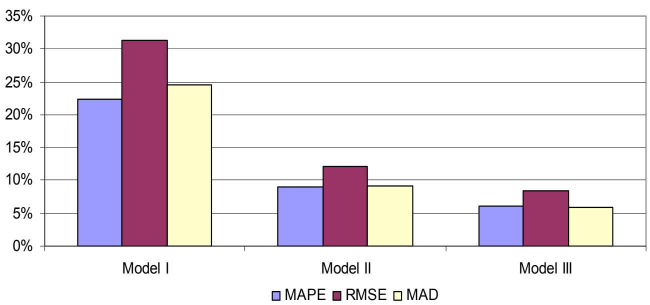

Fig. 6. Detailed analysis of Error Measures identified in different models for special days (holidays) of training set

Model I still reveals serious limitations to forecast future load demand, being particularly evident in special days. For instance, RMSE measure presents an unacceptable result, demonstrating that the model is strongly dependent of large individual errors normally verified in special days. Model II reduces these errors because it includes a binary variable to distinguish special days and normal days, so the neural network is more sensitive to these special days during the training process. Model III besides using historical data it also integrates future load demand prediction (available by the use of load profiles for different classes) and it demonstrates important benefits to the forecasting methodology. To enhance the model viability, a model comparison in forecasting a specific special day is available in Fig. 7.

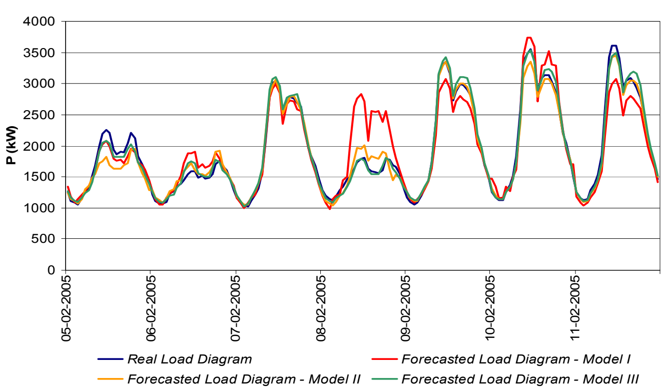

Fig. 7. Comparison between models to forecast one week with a special day - Tuesday, 8th February 2005 corresponds to the carnival day.

Although Model III is the one with more inputs, it doesn't mean that it should need more neurons in the hidden layer to model possible relationships between outputs and inputs. In fact, comparing the proposed models discussed in this Section, it can be said that Model III is the one with less neurons in the hidden layer. Model I uses 7 neurons, model II was trained with 9 neurons, while the model III analyzed only uses 3 neurons. In the hidden layer of an ANN architecture used in Model III, more than 5 neurons only serves to degrade forecasting errors in the testing period, which means that the ANN can be appropriate to the training sample but does not generalize this capability to the testing set. A brief description of the neural network used in Model III is shown in Fig. 8.

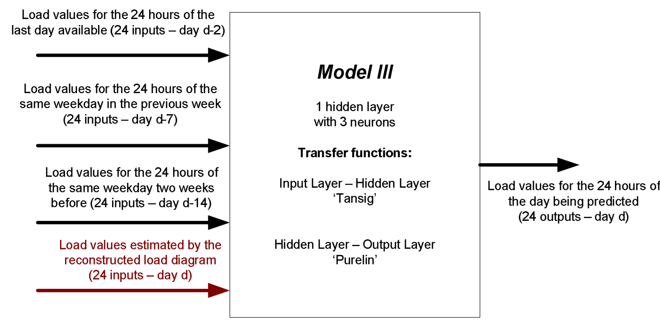

Fig. 8. ANN Architecture for Model III

VI. Conclusions and Further Work

This work intends to analyse the effect of integrate information obtained from available load profiles to load forecasting techniques based on neural networks. The detailed consumers' behaviour knowledge could be in that way used to forecast future global load demand. Once this case study has some particularities such as the global load diagram that was obtained from a LV consumers' aggregation, it is not intention to compare this case study with other load forecasting works already discussed and applied to real distribution or public substations. The load profiling methodology adopted is also a questionable approximation, knowing that classes' segmentation based on clustering methods could be more accurate. Anyway, even with those assumptions and the results presented, the method still represents a promising technique in load forecasting. It is recognized that the viability of the technique strongly depends on adequate consumer segmentation, load profiles accuracy representing each class, proper selection of neural network input vector, lead time to predict (one hour, one day, one week,…) and updated commercial information to characterize the individual contributions of a global load diagram.

In order to get better results in further work and also to make the methodology applicable in realistic cases, it can be described some aspects to be improved:

- adoption of a more accurate methodology to segment consumers;

- creation of distinct load profiles for different day types (evaluate the need to deal with more than one profile for workdays and dedicate more attention to holidays effects – distinction between fixed and mobile holidays);

- analysis of the weight vector between input and hidden layers, discriminating the importance of the vector associated with the reconstructed load diagram among the remaining inputs;

- adoption of an incremental training during the testing phase, making possible the network adaptation to seasonal or unexpected load changes.

VII. References

Periodicals:

[1] H. Hippert, C. Pedreira and R. C. Souza "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation", *IEEE Transactions on Power Systems*, February 2001, Vol. 16, No. 1, pp. 44-55

[2] P. Santos, A. Martins and A. Pires, "Designing the input vector to ANN-based models for short-term load forecast in electricity distribution systems", *International Journal of Electrical Power and Energy System*, May 2007, Vol. 29 - Issue 4, pp. 338-34

[3] M. T. Hagan and M. B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm", *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, November 1994, pp. 989-993

[4] J. N. Fidalgo and J. A. Peças Lopes, "Load Forecasting Performance Enhancement When Facing Anomalous Events", *IEEE Transactions On Power Systems*, February 2005, Vol. 20, No. 1, pp. 408-415

[5] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta and V. Cencellli, "A neural network based technique for short-term forecasting of anomalous load periods", *IEEE Transactions on Power Systems*, Nov.1996, Volume: 11, Issue: 4, pp. 1749 – 1756

[6] D. Gerbec; S. Gasperic, I. Smon and F. Gubina, , "Allocation of the Load Profiles to Consumers using Probabilistic Neural Networks", *IEEE Transactions on Power Systems*, May 2005, Vol. 20, No. 2, pp. 548- 555

[7] J. Jardini, C. Tahan, S. Ahn and F. M. Figueiredo, "Daily Load Profiles for Residential, Commercial and Industrial Low Voltages Consumers", *IEEE Transactions On Power Delivery*, January 2000, Vol. 15 - No. 1, pp. 375-380

[8] M. Espinoza, C. Joye, R. Belmans and B. DeMoor, "Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series", *IEEE Transactions on Power Systems*, August 2005, Volume: 20, Issue: 3, pp. 1622- 1630

Books:

[9] H. L. Willis, "Power Distribution Planning Reference Book", Marcel Dekker, Inc. 2004, Second Edition, Revised and Expanded

[10] M. T. Hagan, H. B. Demuth and Mark Beale. (1996) "Neural network design", PWS Publishing Company, Boston MA.

Papers from Conference Proceedings (Published):

[11] H. Chen, C.A. Canizares and A. Singh, "ANN-based short-term load forecasting in electricity markets", *IEEE Power Engineering Society Winter Meeting,* 2001, vol.: 2, pp. 411-415

[12] P. Campbell and K. Adamson, "Methodologies for Load Forecasting", *3rd International IEEE Conference on Intelligent Systems*, 2006, pp. 800-806

[13] X. Guo, Z. Chen, H. Ge and Y. Liang, "Short-term load forecasting using neural network with principal component analysis", Proceedings of the *International Conference on Machine Learning and Cybernetics*, 26-29 Aug. 2004, Vol. 6, pp. 3365-3369

[14] S. R. Abbas, and M. Arif, "Electric Load Forecasting Using Support Vector Machines Optimized by Genetic Algorithm", *INMIC'06 IEEE Multitopic Conference 2006*, 23-24 Dec. 2006,pp. 395-399

[15] J. Fidalgo and M. A. Matos, "Forecasting Portugal Global Load with Artificial Neural Networks", *ICANN2007 - International Congress on Artificial Neural Networks 2007*, pp. 728-737

[16] G. Chicco, R. Napoli and F. Piglione, "Load pattern clustering for short-term load forecasting of anomalous days", *Proceedings of the IEEE PowerTech 2001*, Porto, 10-13 Sept 2001, Vol. 2, pp. 6pp

[17] M. Matos, J. Fidalgo, and L. Ribeiro, "Deriving LV Load Diagrams for Market Purposes using Commercial Information", *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems*, 6-10 November 2005, pp. 105-110

Standards:

[18] H. Demuth, M. Beale and M. Hagan, "Neural Network Toolbox 5 User's Guide", Revised for Version 5.1, September 2007, Copyright 2005-2007 by the Mathworks Inc.

[19] ERSE – Energy Services Regulatory Authority (2007) , "Guide of Data Measuring, Reading and Availability", document nº 4591-A/2007

VIII. Biographies

J. C. Sousa was born in Coimbra, Portugal, on December 18, 1977. He obtained the MSc. and Eng. Degree from the Department of Electrical Engineering and Computers, of the University of Coimbra in 2006 and 2001 respectively.

He's currently assistant professor in School of Technology and Management in Polytechnic Institute of Leiria since 2003.

In 2007 he joined INESC Coimbra as a researcher and is actually working on his Ph.D degree related with electric energy consumers' characterization and load forecasting.

His main research interest is the use of artificial intelligence in energy planning.

L. P. Neves received his Electrical Engineering degree in 1992 and his PhD degree in 2005, both from the University of Coimbra. He is auxiliary Professor at the School of Technology and Management of the Polytechnic Institute of Leiria. His research areas include Demand-Side Management, Power Systems Planning and Analysis and Operational Research.

H. M. Jorge received his Electrical Engineering degree in 1985 and his PhD degree in 1999, both from the University of Coimbra. He is auxiliary Professor at the Department of Electrical Engineering of University of Coimbra. His research areas include load research, load forecast, load profile, power quality, power distribution and energy efficiency. He is an IEEE member since 1992 (nº. 3181112).