Load forecasting, the importance of the probability "tails" in the definition of the input vector.

P.J.Santos
Dep. of Electrical Engineering
Polytechnical Institute of Setubal
IPS/ESTSetúbal/INESCC
Setúbal, Portugal
paulo.santos@estsetubal.ips.pt

S. Rafael
Dep. of Electrical Engineering-
Polytechnical Institute of Setubal
IPS/ESTSetúbal
Setúbal, Portugal
silviano.rafael @estsetubal.ips.pt

A.J. Pires
Dep. of Electrical Engineering-
Polytechnical Institute of Setubal
IPS/ESTSetúbal/UNINOVA
Setúbal, Portugal
armando.pires@estsetubal.ips.pt

***Abstract*— The load forecast is part of the global management of the electrical networks, namely at the transport and distribution levels. This type of methodologies allows to the system operator, to establish and take some important decisions concerning to the mix production and network management, with the minimum of discretionarity. The load forecast in particularly the peak load forecast, represents an important economic improvement in the global electrical systems. Also in certain circumstances, allow reducing the contribution of the non-renewable units, in the daily mixing production. The regressive methodologies specially the artificial neural networks, are normally used in this type of approaches, with satisfactory results. In this paper is proposed a careful analysis in order to define the best-input vector in order to feed the regressive methodology. It was establish careful analyses of the load consumption series. It makes use of a procedural sequence for the pre-processing phase that allows capturing certain predominant relations among certain different sets of available data, providing a more solid basis to decisions regarding the composition of the input vector to ANN. The methodological approach is discussed and a real life case study is used for illustrating the defined steps, the ANN and the quality level of the results.**

Keywords- Transport and distribution electrical networsk; load forecasting; Smart-Grids; input vector; regressive methods. Load bheaviour

I. INTRODUCTION

Load forecasting, is an important resource in the modern electrical energy systems. Particularly, the daily peak load forecast, is an important auxiliary for the management of the electrical system [1-3]. Between other functions, allow establish the global spinning reserve for the production system, and represent an important economic aspect. In figure 1 is shown an example of the actual type of load forecast made by the Portuguese dispatch operator [4] and represents the load diagram collected from the Portuguese transmission network grid (Fig.2), collected in 2012- 05-17 [4]. The blue line represents the evolution of the consumption during the day, and the gray line represents the forecast value. These type of information is part of the actually data base information of the national transmission grid operator - REN .

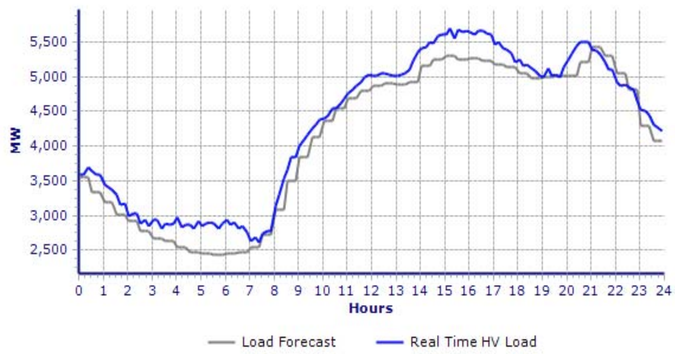

Figure 1 - Real time information collected directly in to the REN (2012-09-06).

The peak load diagram register during this day was 5627 MW, the difference between the gray line (active power forecast value) and the blue line (real time active power value) is around - 400 MW The minus signal give the information that it was necessary to request some spinning reserve [5] in order to fill the gap between the forecast and the real value of active power. The price of the traded energy changes during the day [6]. If we consider, for example the value 60 €/MWh for the traded energy [6], the error between the forecast value and the real value represents during one hour (16:00 pm) the value 30k€. This example shows the importance of these auxiliaries methodologies in the dispatch operator in order to help to establish and take some important decisions, with the minimum of discretionarity [7]. The load forecasting in generally, and the peak load forecast in particularly, represents an important economy improvement in the electrical systems [8].

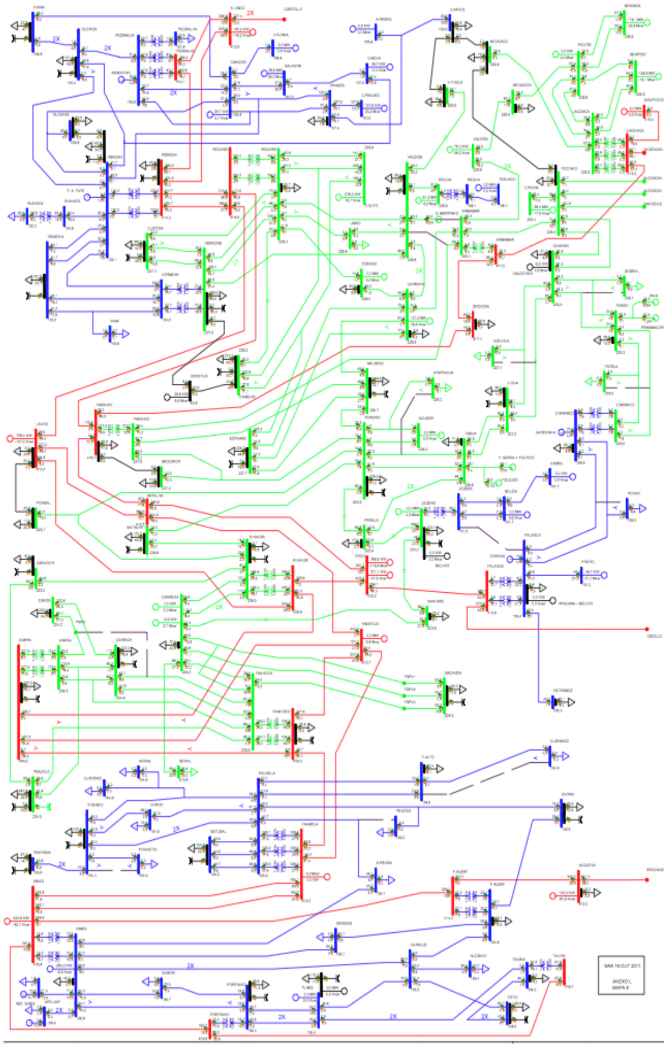

Figure 2 - Portugal Transmission Network ; - 400kV, - 220kV and - 150 kV (REN)

II. METHODOLOGY

The first step is to classify the information of the endogenous signal (active power). Despite the evolution of the signal represented in figure 1, this type of signal presents some different behavior during a day (Fig. 3). The evolution of the daily load diagram of active power, presents two distinct zones. The first one, during a night period presents some low values of active power, and represents the natural decreasing of consumption. A second zone more or less, during the period from 10h: 00m until 22h: 00m represents the increasing of industrial or services activities during the day and the consequent increasing of consumption and the natural increasing of active power.

Despite this daily behavior the load diagram present others periodicities, has the weekly and the annual behaviors, strictly connected to the cycling of consumption for example during the winter the daily peak increase caused by the growing of the heating loads (Fig. 3).

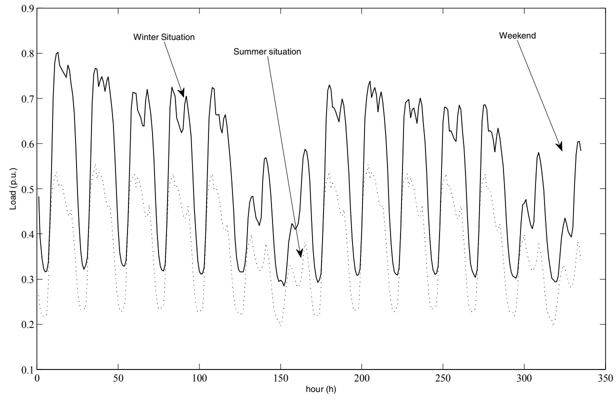

Figure 3 - Two typical week load diagrams for different climatic situations.

As consequence of the cycling behavior certain zones of the active power signal presents a random behavior and other zones have another behavior related with the tendency. In figure 4 is presented a probability plot of the time series, and are marked two zones. A central zone with a random behavior and left and right zones, usually these zones are known as "pig tail".

These two different zones must treat with different composition of the regressor. The first zone implies one feed with more contiguous information. The second zone implies other type of information, based on a tendency of the signal [10].

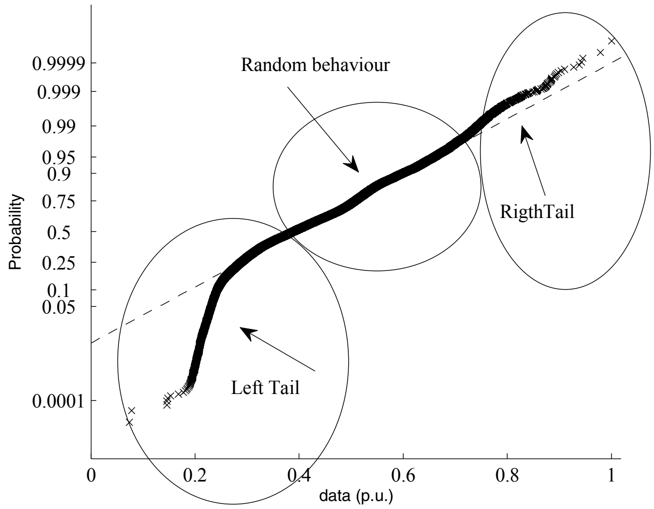

Figure 4 -Probability plot of the active time series.

The final model has developed based in this concept; the forecast value is based in the application of different input vector for each period. It was defined two distinctive zones of actuation for the different vectors. (Fig.5).

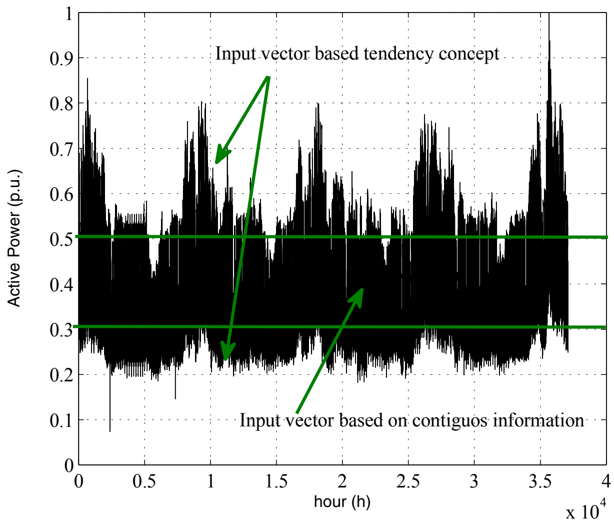

Figure 5 - Zones with of actuation of the different input vectors several years .

A standard feedforward backpropagation artificial neural network has been used for designing the forecast methodology. The structure used having a fully connected architecture with a single hidden layer where hyperbolic tangent is used as the common activation function, the output being activated with linear functions. This is a well-proven arrangement, adequate when, as in the present case, the relations between the variables at stake have a strong non- linear behavior. Mostly trial and error processes, based on engineering judgment and a certain degree of discretion, are the basis for designing input vectors to ANN [1]

Two models with different composition of the input vector have been developed. The first one includes contiguous information (1). And a second vector, with information of the tendency (2).

$$p(t\text{-}1)\ p(t\text{-}2)\ p(t\text{-}3)\ p(t\text{-}4)\ p(t\text{-}5)\ p(t\text{-}6) \qquad (1)$$

$$p(t\text{-}1)\ p(t\text{-}167)\ p(t\text{-}168)\ p(t\text{-}169)\ p(t\text{-}335)\ p(t\text{-}336)\ p(t337) \qquad (2)$$

In the next section is presented the simulation results and the values of error obtained with two different input vectors.

III. SIMULATION RESULTS

Simulations have been carried out with data not used in training, testing or validating the ANN. In Figure 6 is represented the result of the simulation of the input vector without the tendency, only including contiguous information. In this figure is represented only the peak zone. The simulation results denotes a lack of coverage of the peak areas. The behavior is explained buy the composition of the input vector.

In figure 7 is represented for the same period the simulation using the second input vector (2). Is clearly that the results presents a different level of accuracy. And denotes a better result in the prediction of the peak load. The using of the input vector with tendency concept is preferable. In opposition with the input vector (1).

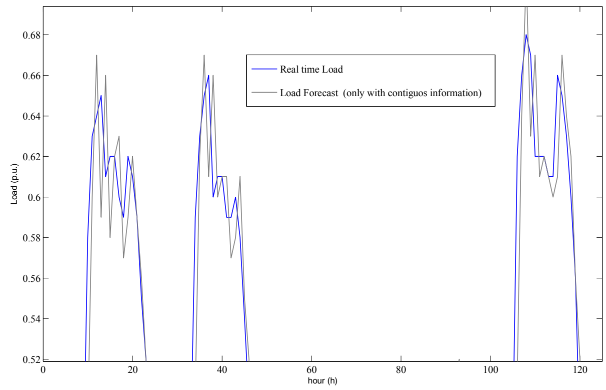

Figure 6 - Peak simulation using contiguous information.

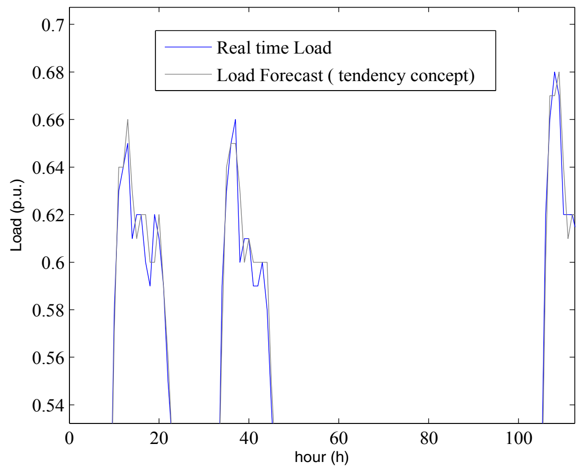

Figure 7 - Peak simulation results using tendency information.

In figures 8 and 9 is represented the similar situation, but for the night period. Is clearly that for this zone of the load diagram, the inferior zone of the load.

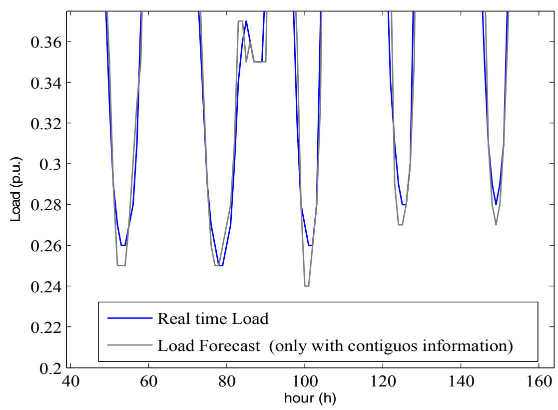

Figure 8 - Night period, simulation results using only contiguous information.

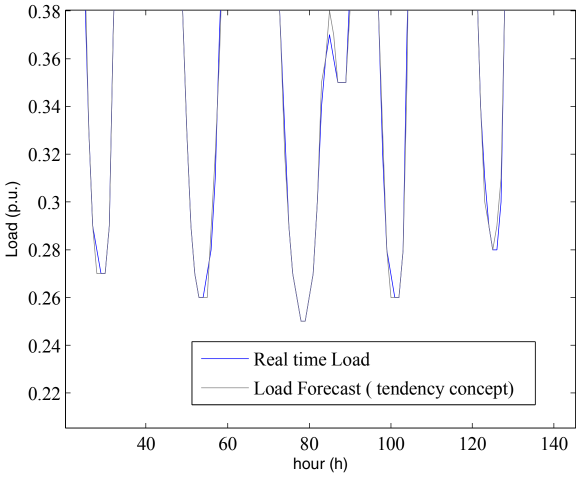

Figure 9 - Night period, simulation results using tendency information.

The mean absolute percentage error (MAPE) value is the most common error indicator, as is generally accepted for comparing different forecast approaches. Other statistical indicators are also necessary to provide a more comprehensive view of forecast results such as the percentage error (PE) and the mean percentage error (MPE) Table 1 presents the forecasting MAPE and MPE results for the three regressors used in the experiments for de service sector. According to Table 1 the best mean absolute percentage error (MAPE) obtained during the test period (two winter weeks, for a service load type) is 1.39%, and was achieved by regressor.

Table 1- Different regressors

Regressor	MAPE(%)	MPE(%)
(#1)	4,95	0,15
(#2)	1,37	0,01

IV. CONCLUSIONS

The load peak forecast assumes a great importance, in dispatch activities, and represents an accuracy improvement in the daily load peak forecast. STLF, represents an important economic aspect in the definition of the spinning reserve in the system and contributes to the security of the production system. Usually ANN short-term load forecasting models, based the composition of the input vector in endogenous information, normally in the active power. The choose of this information represents an important aspect.

The proposed model reduces the discretionary in the selection of endogenous information. For the time series is marked two different zones with different behaviors. The central zone with a random behavior and the left and right zones with a behavior based on the tendency. It was selected and approach based on the two distinctive input vectors. The results obtained with these two approaches allow to get a MAPE value in order of 1,37%. Other error indicators have also been calculated, with the goal of providing a more comprehensive quality analysis of the results obtained.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of EDP Distribuição (Portugal) for all the provided data.

REFERENCES

[1] Hippert, H.S.; Pedreira, C.E.; Souza, R.C.; , "Neural networks for short-term load forecasting: a review and evaluation," *Power Systems, IEEE Transactions on* , vol.16, no.1, pp.44-55, Feb 2001.

[2] (2012, Oct.) The REN website. [Online]. Available: http:// www.ren.pt.

[3] Tang, G.Q.; , "Smart grid management & visualization: Smart Power Management System," *Emerging Technologies for a Smarter World (CEWIT), 2011 8th International Conference & Expo on* , vol., no., pp.1-6, 2-3 Nov. 2011.

[4] Tasre, M.B.; Bedekar, P.P.; Ghate, V.N.; , "Daily peak load forecasting using ANN," *Engineering (NUiCONE), 2011 Nirma University International Conference on* , vol., no., pp.1-6, 8-10 Dec. 2011

[5] (2012, Oct.) The REN website. [Online]. Available: http:// www.ren.pt.

[6] Chowdhury, N.; Billinton, R.; , "Interruptible load considerations in spinning reserve assessment of isolated and interconnected generating systems," *Generation,Transmission and Distribution, IEE Proceedings C* , vol.137, no.2, pp.159-167, Mar 1990.

[7] . (2012, Oct.) The INESCPorto website. [Online]. Available: http:// www2.inesc.pt.

[8] Hao-Tian Zhang; Fang-Yuan Xu; Long Zhou; , "Artificial neural network for load forecasting in smart grid," *Machine Learning and Cybernetics (ICMLC), 2010 International. Conference on* , vol.6, no., pp.3200-3205, 11-14 July 2010.

[9] Makarov, Y. V.; Etingov, P. V.; Huang, Z.; Ma, J.; Chakrabarti, B. B.; Subbarao, K.; Loutan, C.; Guttromson, R. T.; , "Integration of wind generation and load forecast uncertainties into power grid operations," *Transmission and Distribution Conference and Exposition, 2010 IEEE PES* , vol., no., pp.1-8, 19-22 April 2010.

[10] O'Toole, E.; Clarke, S.; , "Dynamic forecasting and adaptation for demand optimization in the Smart Grid," *Software Engineering for the Smart Grid (SE4SG), 2012 International Workshop on* , vol., no., pp.30-33, 3-3 June 2012.