NEURAL NETWORK CONFIGURATIONS EVALUATION FOR ENERGY CONSUMPTION FORECAST

Flávio Augusto Sanzovo Fiorelli, fiorelli@usp.br
Esteban Patricio Manuel Fernandez Arancibia, ep.fdez@gmail.com
Alberto Hernandez Neto, ahneto@usp.br
Escola Politécnica da Universidade de São Paulo
Departamento de Engenharia Mecânica
Av. Prof. Mello Moraes, 2231 - Cidade Universitária
05508-900 - São Paulo (SP) - Brasil

Abstract. *This work presents a comparative study on the use of different artificial neural network configurations in a simulation program to predict the daily energy consumption of a commercial building. Taking as reference a simpler feed-forward model, more complex network arrangements were implemented, such as the recurrent and self-organizing maps models. Daily consumption data of the University of São Paulo Administration Building at Cidade Universitária "Armando de Salles Oliveira" in São Paulo was used as study case. The main hypotheses assumed is that the variations in energy consumption is primarily due to air-conditioning equipment, and therefore the climatic conditions at the building location obtained from the USP meteorological station were used as input parameters for the networks. After implementing and training each configuration with the available data for 2004, it was performed a validation process was performed using the energy consumption of the first three months of 2005, resulting in improvements of up to 10%, in the best scenario, in the forecasts made with more complex network configurations in comparison with the simpler feed-forward network.*

***Keywords**: Energy consumption forecast, Artificial neural network, Feed-forward configuration, Recurrent networks, Self-organizing maps, Modular networks.*

1. INTRODUCTION

Adequately managing the building energy demands has always been a struggle for facility managers. The proper use of energy in a building provides lower operational costs in two aspects. The first one is achieved by evaluating the energy end-uses (lighting, electrical equipments and HVAC) and implementing actions to reduce the amount of energy for one or more of these end-uses. The second one is related to the penalties imposed by electricity companies in Brazil and in other countries due to the increase in the peak energy demand beyond a limit previously agreed in the energy supply contract. If the facility manager could anticipate the energy demand profile and also the energy consumption of the building, he could implement actions to reduce one or both of them and, therefore, reduce the operational cost of the building.

The use of analytical models in this forecast process is very difficult because there is a great amount of variables to consider, and the need of knowing in detail properties and characteristics of the building facility and its components, which very often are not available (Hernandez & Fiorelli, 2008). A possible solution, when there is available previous data about building energy consumption, is the use of models that extrapolate the consumption for new situations based on such previous data.

One of these models is the artificial neural network (ANN), a generic denomination for several simple mathematical models that try to simulate the way a biological neural network (for instance the human brain) works. The main characteristic of such models, which is important for this study, is the capability of learning the "rule" that controls a physical phenomenon under consideration from previously known situations and to extrapolate results for new situations.

There are several works on ANN usage for forecasting energy consumption in buildings (for instance Kalogirou, 2000; Kalogirou & Bojic, 2000; Amjady, 2001; Ben-Nakhi & Mahmoud, 2004; Pao, 2006), most of them using the feed-forward configuration for the neural network, since this is the most known and simple network arrangement.

Campoleone et al (2006) implemented a feed-forward ANN for forecasting the daily energy consumption of the Adminstration Building of University of São Paulo. It was assumed that consumption variations was basically due to air conditioning equipments, and therefore weather conditions (temperature, humidity and solar radiation were used as ANN inputs. The implemented network was trained using weather and consumption data from 2003 to 2004, and validation against data for the three first months of 2005 indicated that ANN predicts energy consumption within ± 10% error range, and that humidity and solar radiation has a second-order influence for the considered building.

Hernandez Neto & Fiorelli (2008) compared the simpler ANN model developed by Campoleone et al (2006) and a detailed building HVAC design and simulation software (EnergyPlus) as forecasting tool for the energy demand. Results show that both models are suitable for energy consumption forecast for the Adminstration Building of University of São Paulo. The authors also carried out a parametric analysis for the considered building on EnergyPlus in order to evaluate the influence of several parameters such as the building profile occupation and weather data on such forecasting. Besides the second-order effect of humidity and solar radiation pointed out by Campoleone et al. (2006), such analysis showed that internal heat gains and equipment performance are more significant for the present case.

As a continuation of these previous works, this paper analyses the influence of different network arrangements on the building energy consumption forecast .

2. ARTIFICIAL NEURAL NETWORKS

The development of artificial neural networks is based on the observation of the biological neural network behaviour (cf. Sabbatini, 2003). A biological neuron, shown in Fig. 1, receives electrical inputs from the nearby neurons through connections called synapses. Combination of such inputs and a minimal level of sensibility result in an electrical output response trigged by this neuron to the other interconnected neurons. The electrical inputs can be excitators or inhibitors of neuron activity, and consequently of its output response.

The artificial neuron, also shown in Fig. 1, is a mathematical element with multiple inputs (x_1, x_2, ..., x_n) and one output (Y). Each input is weighted by a coefficient (w_1, w_2, ..., w_n). The artificial neuron combines its weighted inputs and compares the result to a reference (t). After that, the result is sent to the activation function (F) to determine the neuron output value. This is mathematically expressed for a given neuron j by Eq. (1) (cf. Haykin, 1994). Table 1 presents some examples of activation functions that can be used.

$$U_j = \sum x_i w_{ij} \tag{1}$$

$$Y_j = F_j\left(U_j, t_j\right) \tag{2}$$

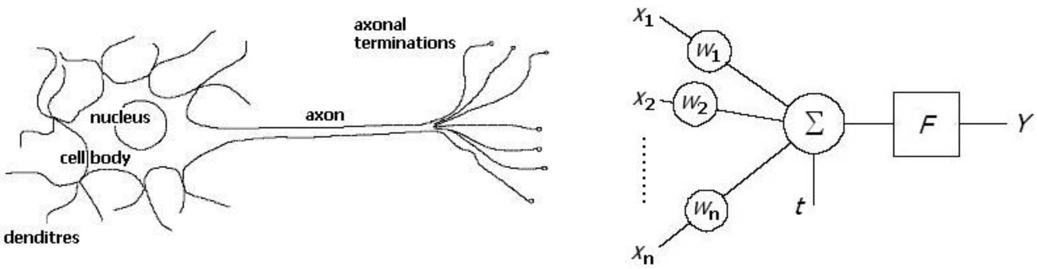

Figure 1. Biological and artificial neurons

Table 1. Some activation functions

Activation function	Expression
Linear	$y = \alpha x$
Threshold[(1)]	$\begin{cases} y = 0 \text{ for } x < 0 \\ y = 1 \text{ for } x \geq 0 \end{cases}$
Sigmoid[(1)]	$y = \dfrac{1}{1+e^{-\alpha x}}$
Gauss	$y = e^{-\alpha x^2}$

[(1)] threshold and sigmoid functions can be used with lower value equals to -1 instead of 0, and in this case they are called "simetrical".

During the network training, the neural network compares its response for a given input with the expected one and evaluates the network error, which will be used to adjust the weights w_{ij} and reference values t_j of each neuron, starting from the last layer to the first one. Such correction method is called back propagation, and is given by:

$$e_j = Y_j\left(1 - Y_j\right)\left(d_j - Y_j\right) \tag{3}$$

$$t_j = t'_j + \lambda e_j \tag{4}$$

$$w_{ij} = w'_{ij} + \lambda e_j x_i \tag{5}$$

where e_j is the error of neuron j, d_j is the difference between the expected output and the obtained one, and λ is the learning rate, which defines a more (or less) significant contribution of calculated error on weigth adjustment.

After evaluating the errors and adjusting weights w_{ij} and reference values t_j of each neuron of the ANN last layer, the process is repeted for the previous layers until the first one is reached. Since the expected intermediary output values are unknown it is employed a modified form of Eq. (3), in which d_j is replaced by the sum of the product of the weights w'_{jk} and the errors e_k of the neurons of the following layer:

$$e_j = Y_j\left(1-Y_j\right)\sum\left(e_k w'_{jk}\right) \tag{6}$$

2.1. ANN Configurations

Since it is not well known how a biological neurons is arranged, several possible arrangements for artificial networks were suggested, generating distinct network models. The following arrangements were used in this work:

- ***Feed-forward***: the most known, simple and used network arrangement is the *feed-forward* model (Fig. 2). In this model, the neurons are placed in several layers. The first one is the *input* layer, which receives inputs from outside. The last layer, called *output* layer, supplies the result evaluated by the network. Between these two layers, a network can have none, one or more intermediate layers called *hidden* layers. The input layer is usually considered a distributor for incoming signals, the hidden layers are signal classifiers, and the output layer is the organizer of obtained responses. It must be pointed out that in the feed-forward model the neurons of a given layer are only connected with the previous layer and the next one, and data is propagated linearly from input to output.
- ***Recurrent***: a more complex network arrangement is the recurrent network, in which there is at least one recurrent loop (feedback connection). The feedback connection has a significant impact on the network learning capacity and performance, and provides a non-linear dynamic behaviour to the network. Depending on the arrangement, the networks can be classified as full or partial recurrent. In this work it was used the Hopfield full recurrent network (Fig. 3), as well as the Elman (Fig. 4) and Jordan (Fig. 5) partial recurrent networks.
- ***Kohonen Self-organizing maps (SOM)***: Inspired on cortex maps, the Kohonen network (Fig. 6) is based on competitive learning. The neurons compete to answer to a given input. During learning, topologically organized neuron groups are formed, and each group is responsible for answering to a given class of input. The self-organizing characteristic that names this class of neural networks is connected to the fact that the only information presented to the network is the input pattern, and the synaptic connections are defined to award the winning neuron without comparison to desired patterns.
- ***Modular networks***: All the previous arrangements described present a good performance when the input data group is small. Nevertheless, network complexity increases and performance decreases quickly as the input data group increases. A way to overcome such problem is the use of ANN modules that works with a subgroup of input data and gives a partial output that is used as input for an ANN decision module that gives the final output of the network.

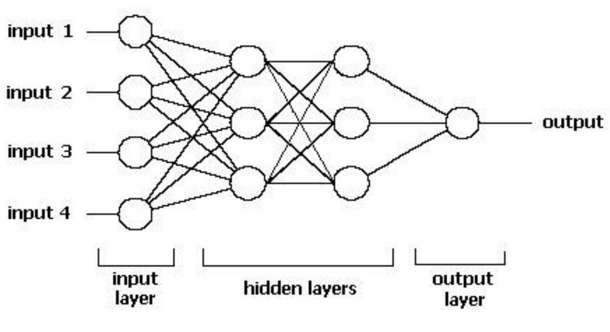

Figure 2. Feed-forward network configuration

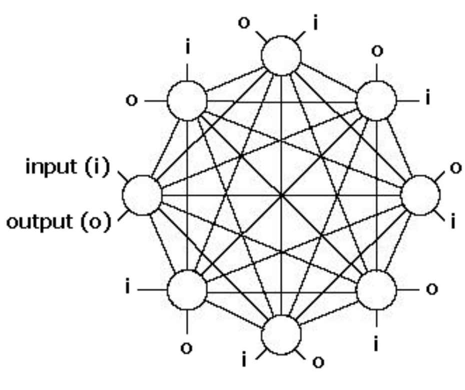

Figure 3. Hopfield network configuration

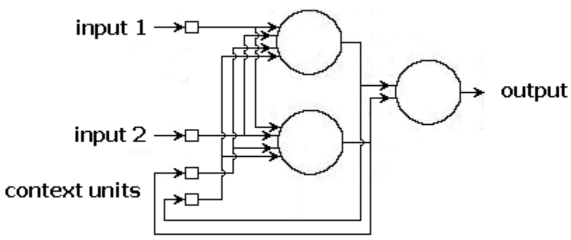

Figure 4. Elman network configuration

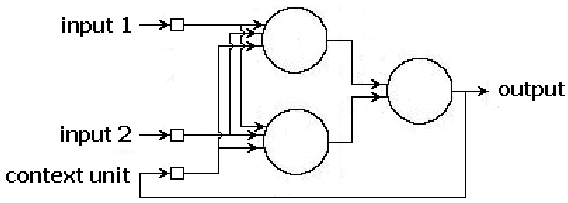

Figure 5. Jordan network configuration

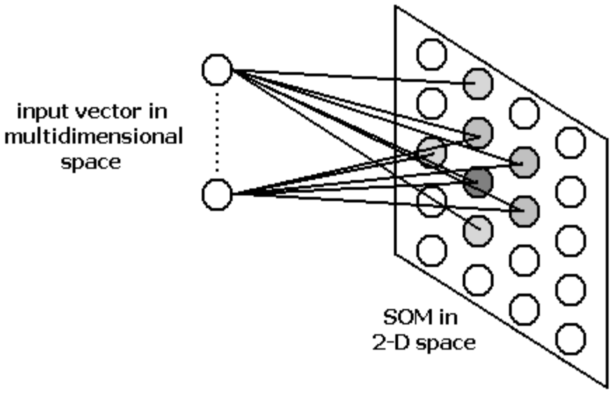

Figure 6. SOM network configuration

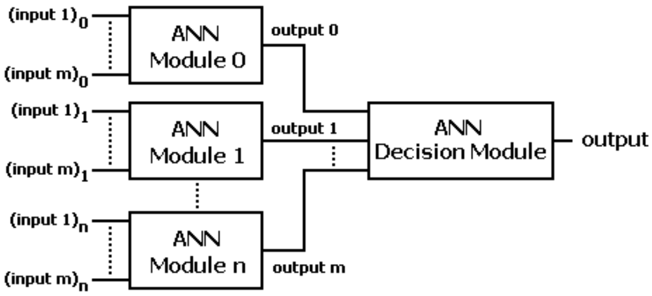

Figure 7. Modular network configuration

3. CASE STUDY

The University of São Paulo (USP) started a program called PURE-USP (Permanent Program of Efficient Use of Energy) (Saidel et al., 2003) in 1997 to design and implement actions in order to reduce energy consumption. Among the several actions implemented by the program, an on-line measurement system for energy consumption that allows the development of building energy consumption profile database can be pointed out, which has become a very important tool for planning the retrofitting actions.

Since the beginning of the program, several retrofits have been implemented in air conditioning systems in use at the University. One of those retrofits was implemented at the University Administration building (Fig. 8), built in the 1970s. That building has two six-floor blocks , with gross floor area of 3.000 m^2. Both blocks are oriented 43° NW and most of the building occupancy occurs between 8:00 and 18:00 from Monday to Friday. Building population is of almost 1000 employees.

There are three major end-uses responsible for energy consumption in the building: air conditioning system, lighting and office equipment (computers, printers, copy machines, etc.). Table 2 shows the end-use breakdown distribution, and Fig. 9 presents a typical profile for measured building energy demand. These profiles were evaluated by the previously mentioned measurement system developed by PURE-USP. It is important to point out that the building air conditioning system is composed of unitary window-type and split air conditioners spread along each floor and individually controlled by the users (cf. Aquino, 2005).

Figure 8. USP Administration building

Table 2. USP Administration building end-use breakdown

End-use	Contribution (%)
Air conditioning	45
Lighting	30
Office equipment	21
Others	4

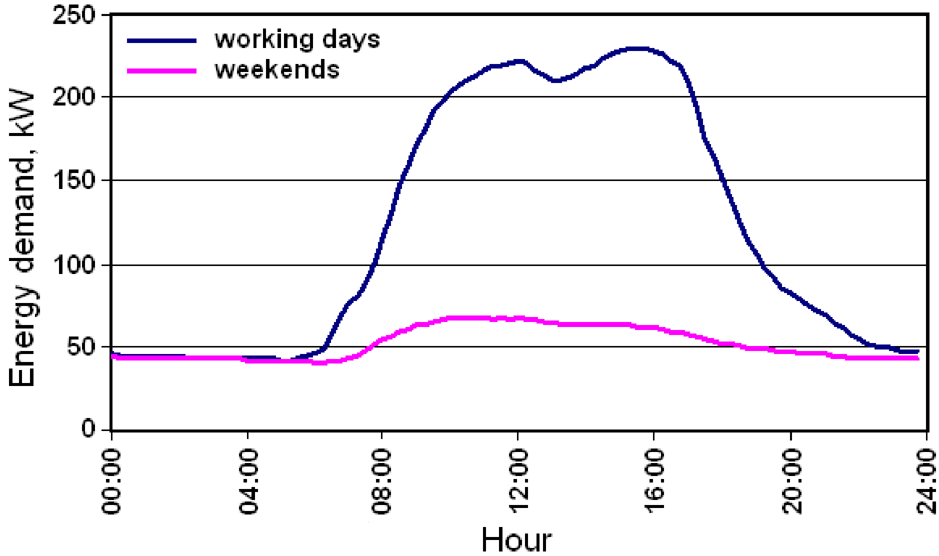

Figure 9. Typical measured building energy demand profile

4. ANN IMPLEMENTATION

As part of the PURE-USP program developments, Campoleone et al. (2006) implemented a simple feed-forward ANN model to forecast the daily energy consumption in the University Administration building. Three different networks were implemented: one for all days (both working days and weekends), and two others for working days only and weekends only.

Based on the assumption that consumption variation during the day is mainly due to air conditioning equipment, the climatic conditions (temperature, humidity and solar radiation) have been considered as parameters for the analysis. The University campus has a meteorological station where the most important parameters have been registered (dry bulb temperature, relative humidity, solar radiation, etc.) for the last ten years, providing a reliable weather database.

Using climatic and energy consumption data acquired in 2003 and 2004 to train the network, it was possible to forecast the load from January to March 2005 within a ±13.5% error range for 85% of the cases with the working-days network. The all-days network, in spite of receiving the day type (working day or weekend) as an extra input, presented a higher error between predicted and actual data than the distinct networks for working days and weekends. Another result of this previous work was that relative humidity and solar radiation have a negligible effect on energy consumption for the considered building.

In order to try to reduce the forecast error range, the more complex ANN arrangements mentioned in section 2 was implemented using a spreadsheet and an ANN add-in (NeuroDimension, 2009). As in the previous work of Campoleone et al. (2006), three different networks were implemented (all-days, working-days, and weekends). Table 3 presents input, output, and training parameters for all networks, and Table 4 shows the network configurations implemented. The parameters adopted for each configuration, like number of layers, etc., are those that presented the best results for the considered ANN configuration. The daily minimum and maximum external temperatures ($T_{\min}$ and $T_{\max}$, °C) were adopted as network input, and the network output was the daily energy consumption (C, kWh).

The comparison parameter of the performance for the different network arrangements was the mean squared normalized error (*MSNE*), Eq. (7),

$$\text{MSNE} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_{norm,i} - f_{norm,i}\left(x,w\right)\right)^2 \tag{7}$$

where $Y_{norm,i}$ represents the actual energy consumption value normalized to [-1,1], and $f_{norm,i}(x,w)$ is the network output, also normalized to [-1,1]. Normalization is used to ensure that networks with multiple outputs will be trained so that accuracy of each output is treated as equally important. Without normalization outputs with larger values, and therefore larger errors, would be treated as more important (MathWorks, 2009).

Table 3. Input, output and training parameters for all networks

Parameter	**All-days**	**Working-days**	**Weekends**
Inputs		$T_{\min}$, $T_{\max}$	
Output		C	
Training database	231	163	78
Validation database	81	56	25

Table 4. Network configurations

Network arrangement		Number of layers	Input layer neurons	Output layer neurons	Hidden layer(s) neurons
Feed-forward	all-days	3	2	1	5
	working-days				14
	weekends				8
Hopfield	all-days	3	2	1	6
	working-days				4
	weekends				2
Elman	all-days	3	2	1	8
	working-days				7
	weekends				2
Jordan	all-days	3	2	1	8
	working-days				4
	weekends				4
SOM(1)	all-days	4	2	1	8 / 4
	working-days	3			4
	weekends	3			2
Modular(2)	all-days working-days weekends	4	2	1	2

Obs: (1) for SOM configuration, the neuron plane dimensions are 5x5 (all-days & working-days) and 4x4 (weekends);
(2) for modular configuration, the data refers to the number of modules in a given layer. Each module uses a feed-forward configuration with two neurons at inlet layer, two hidden layers with 5 (superior modules) or 4 (inferior modules) neurons, and an output layer with one neuron.

5. RESULTS

The *MSNE* for the ANN arrangements implemented in this work are presented in Table 5. It can be verified that the higher *MSNE* values are observed for all-days and weekends networks, and the best results are observed for working-days networks, excepting the Hopfield configuration. This best performance of working-days networks is a similar result to that observed by Campoleone et al. (2006), and confirms that it is better to work with two different networks for working days and weekends instead of a single network for all days for the building considered.

Table 5 also shows that, for the working-days network arrangements, the best results are given, in order, by Elman, SOM and feed-forward arrangements. Elman arrangement reduces *MSNE* of about 10% in relation to simpler feed-forward configuration. Figures 10 and 11 show the comparison of the actual data and the network outputs for Elman and feed-forward, which also indicates a better agreement of Elman arrangement forecasts to actual data.

A possible explanation to the best performance of Elman arrangement would be the feedback from the hidden layers to the network input neurons, which gives to this network higher generalization capacity and speed up learning process in comparison to feed-forward. A similar behaviour would be expected for the Jordan arrangement since it is also a partial recurrent network, but the feedback from the output layer to the input layer seems to make the learning process slower and to reduce the generalization capacity.

The worse performance was observed for the Hopfield configuration, but if one considers that t Hopfield networks are normally employed for image perception and recognition or for optimization problems, and in principle are not suitable for such application, it would explain its worse performance.

Table 5. *MSNE* values for the ANN arrangements implemented.

Network arrangement	*MSNE* all-days	working-days	weekends
Feed-forward	0,890	0,650	0,937
Hopfield	0,969	0,905	0,966
Elman	0,917	0,587	0,946
Jordan	0,943	0,809	0,949
SOM	0,933	0,631	0,951
Modular	0,919	0,743	0,945

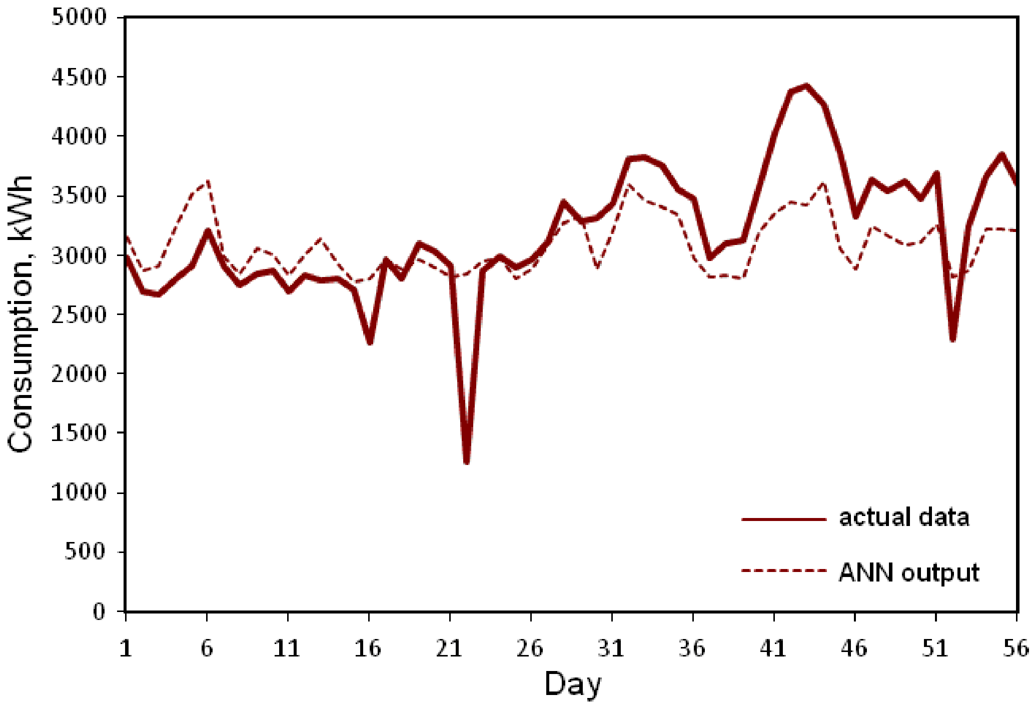

Figure 10. Comparison of actual data and ANN output for working-day feed-forward ANN configuration

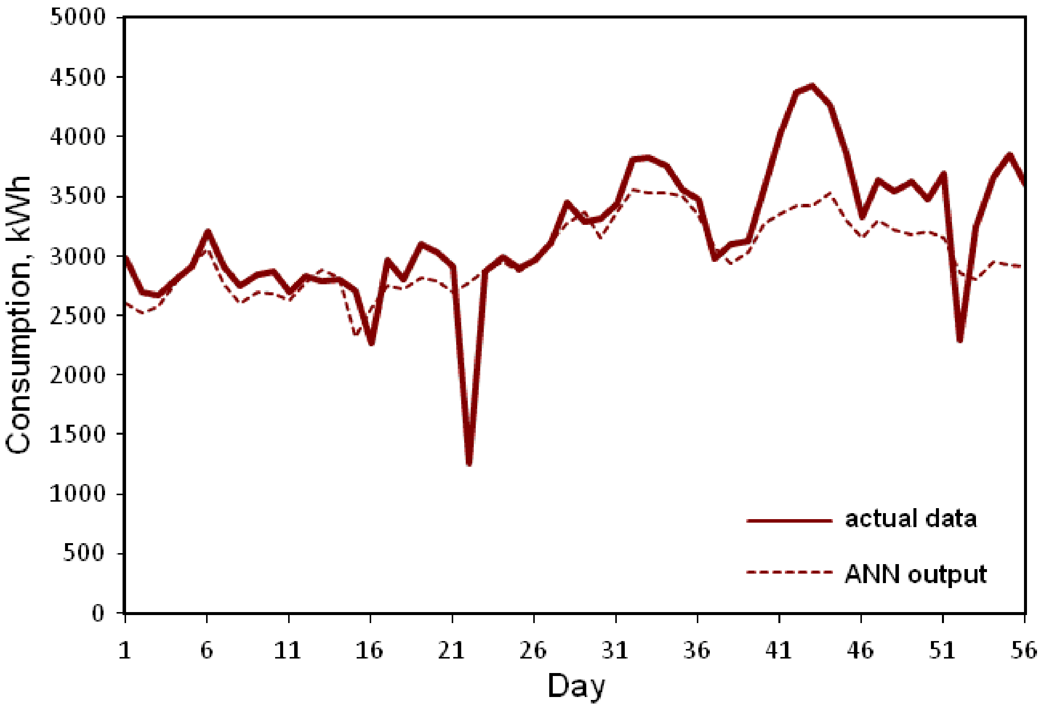

Figure 11. Comparison of actual data and ANN output for working-day Elman ANN configuration

6. CONCLUSIONS

Considering the implemented network arrangements, the Elman configuration presented a better performance than the feed-forward for working-days ANN, with a mean square error 10% lower than the simpler configuration. Although for all-days and weekends the feed-forward presented a slightly better performance, since the critical situation for energy consumption forecast occurs for working days, the result indicates that the Elman arrangement would be the most suitable arrangement for this case study. It is important to point out that the database is relatively small, and should be increased in order to enhance the confidence of the ANN forecast for all the arrangements considered.

Another aspect that should be addressed in future works is to analyse buildings with different air conditioning systems configurations, particularly with centralized systems, as well as different types of buildings, such as hotels, museums, schools, etc.

It should also be pointed out that the occupant's behavior in a building where the air conditioning equipment are mainly unitary systems (window-type air conditioners and split systems) can significantly affects the energy consumption profile, making its forecasting more difficult.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge the staff of the Permanent Program of Efficient Use of Energy of University of São Paulo (PURE-USP) for providing the energy consumption data, as well as the staff of the Institute of Astronomy, Geophysics and Atmospheric Sciences of the University of São Paulo (IAG-USP) for providing the weather database used in this paper.

8. REFERENCES

Amjady, N., 2001, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability", IEEE Transactions on Power Systems, Vol. 16, No. 3, pp. 498-555.

Ben-Nakhi, A.E., Mahmoud, M.A., 2004, "Cooling load prediction for buildings using general regression neural networks", Energy Conversion and Management, Vol. 45, Nos. 13-14, pp. 2127–2141.

Aquino, R., 2005, "Maintenance management of window-type and split air conditioners", Undergraduate Final Project, Escola Politécnica da USP, São Paulo, Brazil (in Portuguese).

Campoleone, E.T., Fiorelli, F.A.S., Hernandez Neto, A., 2006, "Neural Network for Energy Consumption Forecast", Proceedings of 4th National Congress of Mechanical Engineering, Recife, Brazil, CD-ROM (in Portuguese).

Haykin, S., 1994, "Neural networks - a comprehensive foundation", MacMillan College Publishing Co., New York, USA, 696 p.

Hernandez Neto, A., Fiorelli, F.A.S., 2008, "Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption", Energy & Buildings, Vol. 40, No. 12, pp. 2169–2176.

Kalogirou, S.A., 2000, "Applications of artificial neural-networks for energy systems", Applied Energy, Vol. 67, Nos. 1-2, pp. 17-35.

Kalogirou, S.A., Bojic, M., 2000, "Artificial neural networks for the prediction of the energy consumption of a passive solar building", Energy, Vol. 25, No. 5, pp. 479-491.

MathWorks, 2009, "MatLab Neural Network Toolbox™ Documentation", available at <http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/index.html?/access/helpdesk/help/toolbox/nnet/&http://www.mathworks.com/support/product/product.html?product=NN>. Last access 1 May 2009.

NeuroDimension, 2009, "NeuroSolutions for Excel". available at: <http://www.neurosolutions.com/products/nsexcel/>. Last access 1 May 2009

Pao, H., 2006, "Comparing linear and nonlinear forecasts for Taiwan's electricity consumption", Energy, Vol. 31, No. 12, pp. 2129–2141.

Sabbatini, R., 2003, "Neurons and Synapses – A History of its Discover ", Cérebro & Mente, May-August, São Paulo, Brazil (in Portuguese).

Saidel, M.A., Galvão, C.R., Fadigas, E.A.F.A., Affonso, O.F., Silva, F.A.T., Bussolini, M. Gimenes, A.L., Hernandez, A.N., Moreira, J.R.S., Coelho, S.T., Silva, O.C., Altafin, R.A. Ziles, R.A., Romero, M.A., 2003, "Energy Efficiency and Alternative Energy Sources at the University of São Paulo". Proceedings of 17th International Congress of Mechanical Engineering, São Paulo, Brazil, CD-ROM.

9. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.