Available online at www.sciencedirect.com

SciVerse ScienceDirect

Energy Procedia 14 (2012) 45 – 55

Energy Procedia

ICAEE 2011

Using Recurrent Artificial Neural Networks to Forecast Household Electricity Consumption

Antonino Marvuglia[a], Antonio Messineo[b*]

[a]CRP Henri Tudor/CRTE, 66 rue de Luxembourg, L-4002 Esch/Alzette, Luxembourg
[b]Engineering and Architecture Faculty, University of Enna Kore, Cittadella Universitaria, Enna, 94100, Italy

Abstract

The electricity consumption related to the civil sector (residential and tertiary) in the most developed countries has considerably increased during the last years, especially in the summer season. One of the reasons for this rise can be found in the drastic growth of the sales of mono and multi-split systems for air-conditioning. In this context is very important to assess the correlation between electricity demand and utilization of electric appliances (especially air-conditioners). This paper describes a model based on an Elman Artificial Neural Network (ANN) for the short-time forecasting (1 hour ahead) of the household electric consumption related to a suburban area in the neighbours of the town of Palermo (Italy). One of the aims of the study is the assessment of the influence of the use of air-conditioning equipments on the electricity demand.

Keywords: Short-term load forecasting, artificial neural networks, air-conditioning systems, carbon dioxide, sensitivity analysis.

1. Introduction

The primary energy consumption related to the civil sector in the European Community accounts for more than 40% of the total. This means that the building sector is responsible for about 40% of the total CO_2 emissions [1]. In 2007 the CO_2 emissions related only to the residential sector approached 10% of the total CO_2 emissions of EU-27 [2]. The building sector thus represents a crucial sector for the accomplishment of initiatives aiming to energy optimization and consumption reduction.

Concerning Sicily, the electricity consumption related to the household sector increased steadily until

* Corresponding author. Tel.: +39 0935 536448; fax: +39 0935 5369513.
E-mail address: antonio.messineo@unikore.it.

1876-6102 © 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of the organizing committee of 2nd International Conference on Advances in Energy Engineering (ICAEE).
doi:10.1016/j.egypro.2011.12.887

2006, then it was affected by a slight decrease in 2007, to climb up again, until 2010, where a new very small decrease (0.45% with respect to 2009) was observed. In Fig. 1 it is showed the electricity consumption trend from 2001 to 2010. The biggest growth is related to the tertiary sector (40.7%), followed by the household sector (8.3%); the agriculture and industrial sectors are instead characterized by a decrease of the electricity consumption (respectively 9.9% and 8.7%) [3]. In 2010 the household sector was responsible for the second largest share (30.6%) of the total Sicilian electricity consumption. The largest percentage of the total consumption was related to the industrial sector (37.5%), with an important role played by industrial activities with elevated cooling requirements [4, 5]. The tertiary and agriculture sectors were respectively responsible for 29.7% and 2.2% of the overall consumption of the region.

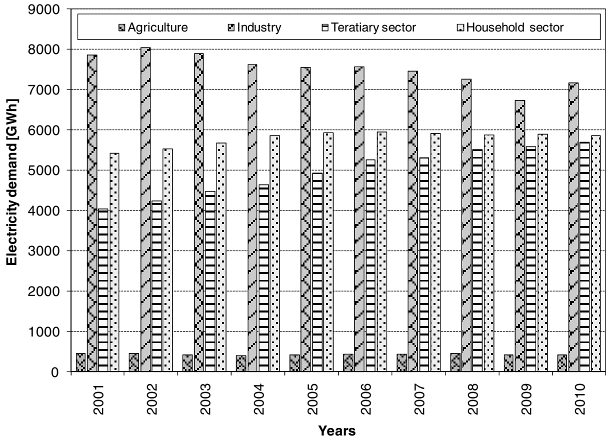

Fig. 1. Electricity demand in Sicily from 2001 to 2010 divided by sector (Source: Terna S.p.A.).

From a more general point of view, taking into account the electricity consumption of the civil sector in the most developed countries, it is possible to observe that what has considerably increased is above all the summer demand [6, 7]. This fact is mainly due to the increased level of the indoor comfort expectations by the population, accompanied by the spread of the air-conditioners (AC) [8] (thanks to the rapid reduction of their unit cost).

The market for heat pumps is steadily increasing throughout the world. However, it is worthy nothing that large efforts are made to decrease the energy use in buildings through numerous regulations and directives. Passive house concepts and net-zero energy buildings are also becoming increasingly popular and this could change the worldwide market conditions for heat pumps in the future.

The Italian market for air-conditioners (at least the single-split and multi-split systems) has been increasing up to 2007. There was then a decrease in the subsequent two years and a new growth in 2010. The data plotted in Fig. 2 were collected by an Italian association of aeraulics plants and equipments [9-11] after a survey carried out on 46 companies working on the Italian AC market. The massive market penetration of these systems, coupled with the occurrence of sudden heat waves that several times have been observed in Mediterranean climates have often caused unforeseen peak loads that have often brought the Italian national power grid to the point of collapse. It is thus possible to understand the importance that reliable short and very short-term forecasting models for the electricity demand at urban scale could have for electricity dispatchers. A very precise forecast allows the identification of any emergency situation in advance and the consequent realization of adequate countermeasures by the electricity operators.

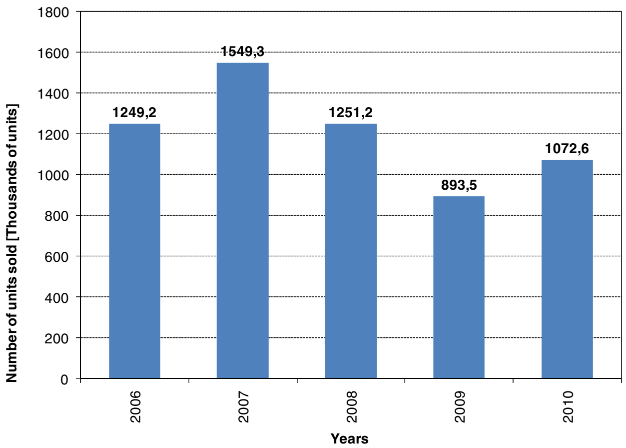

Fig. 2. Split and multi-split air-conditioning systems sold in Italy from 2006 to 2010 (Source: [7-9]).

In the paper we apply a forecasting model to the residential electricity consumption of a certain area of the town of Palermo (Sicily). The model exploits the predictive ability of a particular type of recurrent neural network (RNN) called Elman network [12]. In a neural network model, simple nodes (called *neurons*, *neurodes*, *PEs* (*processing elements*) or simply *units*) are connected together to form a network. Some specific algorithms are used to update the strength (weights) of the connections in the network to produce a desired signal flow.

The utility of artificial neural network (ANN) models lies in the fact that they can be used to infer a relation from observations. This is particularly useful in applications where the complexity of the data or task makes the elicitation of such a relation impractical.

The use of neural networks for load forecasting and also electricity price forecasting is very common [13-16]. The most applied model is the Multi-Layer Perceptron (MLP), consisting of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer.

However, when applied to problems different than pattern recognition or classification, the network's performances are significantly improved by the utilization of a short-term memory like the one implemented in RNNs [17].

For the application presented in this paper, several Elman networks were tested and the best model was sought in terms of prediction error.

2. A short description of the study

An Elman neural network has been implemented in order to forecast the electric current intensity at time t, knowing the values of the same variable at times (*t-3*), (*t-2*) and (*t-1*). The work started with the collection of different types of data:

1. weather data: air dry-bulb temperature (*T*), relative humidity (*U*), wind speed (*W*); global solar radiation (*R*), atmospheric pressure (*P*);
2. historical series of the hourly mean values of electric current intensity (*I*);
3. a variable (*HC index*) which takes into account the estimated number of residential air conditioning units likely to be used hour by hour in the dwellings located in the studied area;
4. a variable, called *Humidex* (*H*), linked to *T* and *U*, whose aim is to estimate the rate of discomfort

felt by the inhabitants.

The data related to electricity consumption were supplied by the Italian National Grid Operator (GSE) split according to the different end-uses (household, industrial and commercial sector) and, in consideration of the aims of our work, only data related to the household sector were used for the network's training.

Several trainings have been carried out by using different network architectures; in particular some simulations were also performed by excluding one input variable at a time and comparing the different results obtained.

3. Elman neural networks

Elman neural networks are also known as partial recurrent networks or simple recurrent networks. These are MLPs augmented with one or more additional context layers which store output values of one of the layers delayed by one step. These layers are used to activate this or some other layer in the next time step, as is sketched in Fig. 3.

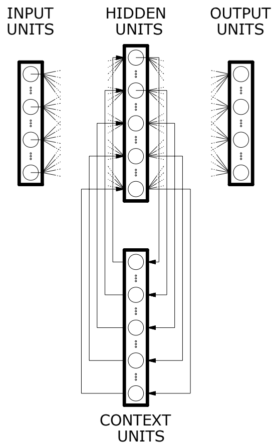

Fig. 3. Schematization of an Elman recurrent neural network.

The feedback from the hidden to the context layer allows Elman networks to learn, recognize and generate temporal patterns, as well as spatial patterns. Every hidden neuron is connected to only one neuron of the context layer through a constant weight equal to +1. Hence, the context layer constitutes a kind of copy of the state of the hidden layer, one instant before.

The number of context neurons is consequently the same as the number of hidden neurons. According to the method presented by Sarle [18], the whole data set was subdivided into a *training set* and a *validation set*. The whole training phase was stopped when the lowest error on the validation set was reached.

4. Input data

The historical series of the climatic data and the electric current intensity data are recorded on a hourly basis (hourly mean values) and cover a period of 79 weeks.

One of the aims of the work is to provide an estimation of the influence of the use of AC appliances on the overall domestic electricity demand. Because of the lack of any reliable survey regarding the presence of this kind of appliances in the small district investigated, this information was estimated by an “in situ” visual inspection.

In order to estimate the presence of AC devices installed during the whole period covered by the time series, national scale statistical data have been used and the data on a local scale were obtained by assuming the same yearly percentage variation as the national data. The trend of the devices installations has been assumed to be linearly growing for the periods ranging from 1st May to 31st August, with the same slopes as to the average national trends. It has been considered constant for the remaining periods (no new AC devices sales assumed after summer).

In order to estimate the number of AC appliances which are likely to be switched on in the dwellings located within the studied area, a thermal discomfort index was used. In fact it is normal to expect that the decision to turn a cooling system on or off is mainly based on a discomfort sensation felt by the population. The threshold conditions determining this discomfort sensation should be evaluated on the basis of the combined effect of temperature, thermal radiation, humidity, air speed, clothing insulation and metabolic rate [19]. An exhaustive list of the existing thermal comfort indexes can be found in the reference [20]. In our case study the most of the aforementioned variables were unknown and therefore only a simplified approach was possible. At this aim, the values of the Thom’s *Discomfort Index* (*DI*) [21] and of the *Humidex index* [22] were computed for each hour of the available data set. Thom [21] has proved that DI represents a very good approximation of the ASHRAE New Effective Temperature (ET) [23] for the common range of temperatures of the Mediterranean area. The values of DI were computed by the equation [24]:

$$DI = T - (0.55 - 0.0055 \cdot RU)(T - 14.5) \tag{1}$$

where T is the dry bulb air temperature (in °C) and RU is the air relative humidity (%).

This index provides reliable information about the discomfort conditions of a site [25]. According to the classification of the comfort conditions made by the *DI*, a subtle thermal discomfort sensation is felt by the population (10% of the people feel discomfort) when the values of *DI* are higher than 21. The most of the population feel a thermal discomfort sensation when *DI* is between 26 and 29.

The *Humidex index* is whereas defined as follows:

$$H = T + 5/9 \cdot (e - 10) \tag{2}$$

where:

T is the dry-bulb air temperature (in °C);

e is the air vapour pressure (in hPa) measured with a psychrometer.

If the value of the air vapour pressure is not available, it can be estimated through a function which combines relative humidity and temperature:

$$e = 6.112 * 10^{\frac{7.5 \cdot T}{237.7 + T}} \cdot R/100 \tag{3}$$

where R is the air relative humidity (%).

The *Humidex index* can be interpreted as the temperature actually perceived by the human body due to the combination of the dry-bulb air temperature and the relative humidity and it is measured in Celsius degrees.

Different values of *H* identify different categories of discomfort, corresponding to the following levels of alert:

Table 1. Comfort/discomfort categories related to the values of Humidex index.

Comfort/discomfort categories	Value
Full comfort	H′27
Subtle discomfort	27≤H≤30
Great discomfort	30′H≤40
Danger	40′H≤55
Imminent heat stroke	H∃55

By plotting the points of the data set which represent a discomfort condition according to the Thom's *Discomfort Index* (with a threshold value of 21) and according to the *Humidex index* (with a threshold value of 27) the graph showed in Fig. 4 was obtained.

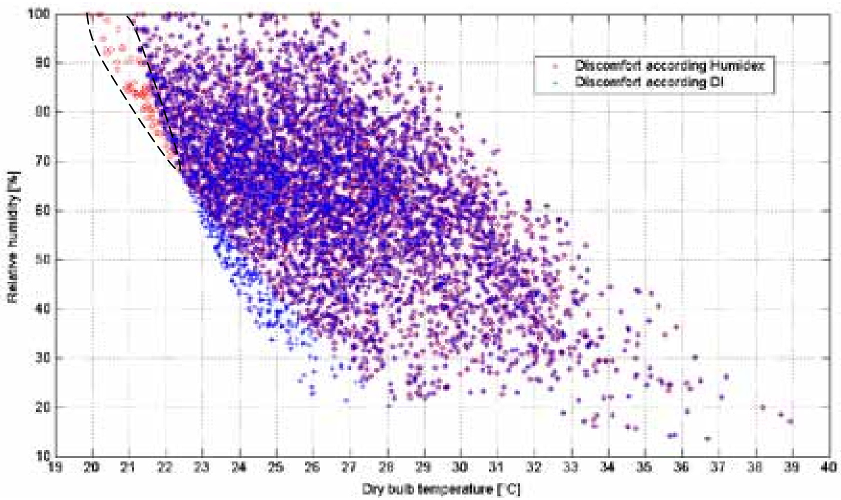

Fig. 4. Plot of the points of the data set which represent a thermal discomfort condition according to the *DI* (crosses) and the *Humidex index* (circles). The points contained into the area bounded by the dashed line in the upper left part of the diagram are representative of comfort conditions according to the *DI*.

It is possible to observe that the two indices are in a very good agreement as to the classification of the comfort/discomfort situations. In particular, it can be noted that the only points classified as representative of discomfort conditions by the *Humidex index* and not by the *DI* are contained into the area bounded by the dashed line in the upper left part of the diagram. These points have a relative humidity higher than 70%. If the *ID* would have been used, these points would have been considered as representative of comfort conditions and thus any air conditioning device should have been considered as turned on, even if the humidity is so high. For this reason we decided to evaluate the number of air conditioners used hour by hour through the utilization of the *Humidex index* instead of the *DI*. However it is worth noting that in our case both the indices have been computed by using the data referred to the outdoor environment, and not to the indoor one. However, in the lack of indoor data, they can give a first approximation of the indoor thermal conditions.

The neural network was therefore provided with an additional input calculated as follows:

$$HC = \begin{cases} ab - index \text{ if } H \le 33 \\ ab - index \times n - cond \text{ if } H > 33 \end{cases} \tag{4}$$

where *n-cond* is the estimated number of AC installed in the investigated area (estimated as described before) and *ab-index* is an index used to take into account the seasonal variations of the number of people living in the area and the consequent effects on the usage of domestic electric appliances. It was set to +1 for the periods ranging from June 15th to September 15th (presence of *resident population* + *fluctuating population*) and to +0.8 for the rest of the year (sole presence of resident population).

The threshold value 33 is associated with the beginning of a discomfort situation for indoor environments, inducing people to turn on the AC appliances. This value was selected because, after different trainings of the same network developed using different values of this threshold, it was the one that assured the best prediction performances.

5. Neural model selection and training

As previously stated, several combinations of *prediction order* (i.e. the dimension of the *lag space* of the input data), number of neurons in the hidden layer and number of input variables were tried in order to find the architecture able to model the data in the most effective way.

The available data set was split into a *training set* (about 80% of the original data set) and a *validation set* (about 15% of the original data set). The remaining data (about three weeks)were used as test data, to evaluate the generalization ability of the trained network.

The model was implemented by using the Stuttgart Neural Network Simulator (SNNS) v.4.2 [26] and the learning algorithm utilized is the *Resilient Backpropagation* (RPROP) modified for Jordan-Elman networks (JE_RPROP). RPROP is a local adaptive scheme, performing fast and robust supervised batch learning in neural networks [27]. In the training through this algorithm a method called *teacher forcing* [28] was used.

Before running the training phase of the network, all the data (except the *HC index*) were linearly normalized in the range [-1, 1]. The *HC index* was normalized by using the technique described by the following equations:

$$x' = \frac{x - \bar{x}}{Std_x} \tag{5}$$

$$x_{norm} = \frac{1}{1 + e^{-x'}} \tag{6}$$

where Std_x is the standard deviation of the generic vector component before normalization.

The forecasting ability of each trained model was tested by using the data related to the aforementioned week. The percentage prediction error (PPE) at time t_i was calculated as:

$$\varepsilon(t_i) = 100 \cdot \left| \frac{\hat{y}(t_i) - y(t_i)}{y(t_i)} \right| \tag{7}$$

where $\hat{y}(t_i)$ indicates the forecasted electric current intensity at hour t_i and $y(t_i)$ indicates the actual intensity at the same hour. Obviously, before comparing the network outputs with the actual electric current values of the period they are referred to, the first were de-normalized. Table 2 summarizes the architecture and the performances of the trained networks. The first column contains three numbers for each network, indicating the number of input, hidden and output neurons, respectively. The second column contains the list of the variables used as input of the model and the last two columns contain the average and the maximum PPE obtained on the test data.

Table 2. Summary of the tested Elman networks

Topology	Input*	Average PPE	Maximun PPE
22 – 40 – 1	*T, U, W, R, P, H, HC, I*	*1,5%*	*4,6%*
19 – 40 -1	*T, U, W, R, P, HC, I*	*3,9%*	*13,8%*
19 – 40 -1	*U, W, R, P, H, HC, I*	*5,1%*	*20,0%*
19 – 40 -1	*T, U, W, R, H, HC, I*	*4,2%*	*13,4%*
19 – 40 -1	*T, U, R, P, H, HC, I*	*4,6%*	*18,1%*
19 – 40 -1	*T, W, R, P, H, HC, I*	*4,7%*	*12,2%*
19 – 40 -1	*T, U, W, P, H, HC, I*	*3,9%*	*18,1%*
22 – 20 – 1	*T, U, W, R, P, H, HC, I*	*1,8%*	*5,7%*
22 – 60 – 1	*T, U, W, R, P, H, HC, I*	*5,6%*	*23,6%*

**all the inputs are referred to time steps (t-3), (t-2) and (t-1), except the variable HC, which is referred only to step (t-1).*

The best model was the network with 40 neurons in the hidden layer, eight input variables and a *prediction order* equal to 3 (first row of Table 2).

The *brain-state-in-a-box* (BSB) activation function [29] was used for the hidden layer and for the output layer. The error function minimized during the training is expressed by the following equation:

$$E = \sum \left(y(t_i) - \hat{y}(t_i) \right)^2 + 10^{-\alpha} \sum w_i^2 \tag{8}$$

The weight-decay parameter α determines the relationship of two goals, namely to reduce the output error (the standard goal) and to reduce the size of the weights (to improve generalization, that is the capacity to maintain good predicting performances also when the network is fed with completely new inputs). The parameters used in the learning algorithm are the following: $\eta^+ = 1.2$; $\eta^- = 0.5$; $\Delta_{\min} = 1e^{-6}$; $\Delta_{\max} = 30.0$. The starting value for Δ_{ij} was set to $\Delta_0 = 0$. These values are quite close to those suggested by M. Riedmiller and H. Braun [27]. The value of the weight decay parameter α was set to 4. Finally, the value of the special parameter linked to the teaching forcing was set to 1.

The connection weights were initialized to zero-mean random values with adequate upper and lower bounds of -1 and +1. The network was trained with 150 epochs because higher values did not lead to better learning. An example of the forecasting ability of the model is given in Fig. 5. It shows the actual and the forecasted hourly current intensities for one of the test weeks, a summer week (from 23rd to 29th of July).

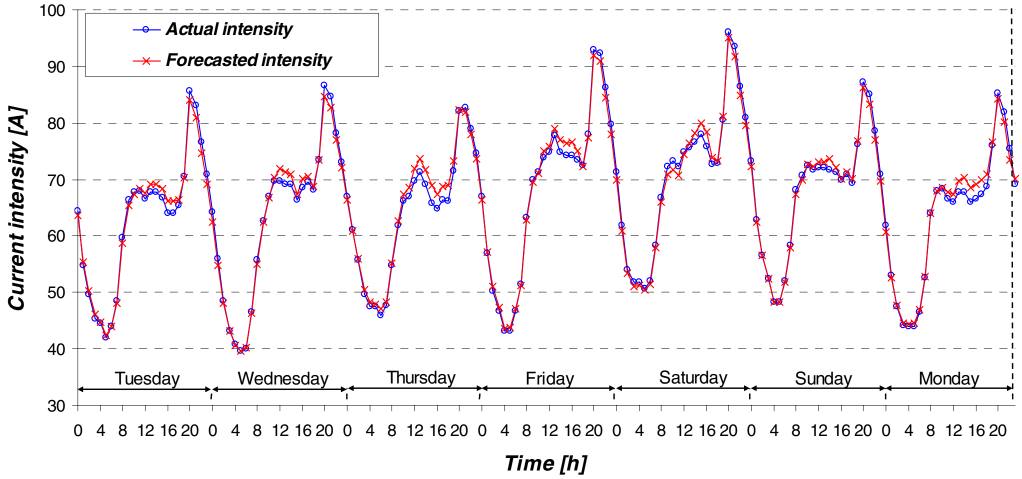

Fig. 5. Actual and forecasted current intensity curves for one of the test weeks.

6. Sensitivity analysis

In order to have a better insight into the relative importance of the various inputs of the model, a sensitivity analysis was accomplished by using a very simple technique: the *weight method*. This procedure was initially proposed by Garson [30] and then repeated by Goh [31].

The method essentially involves partitioning the hidden-output connection weights of each hidden neuron into components associated with each input neuron. Two steps have to be accomplished:

1. For each hidden neuron h, divide the absolute value of the input-hidden layer connection weight (W_{ih}) by the sum of the absolute value of the input-hidden layer connection weight of all the n_i input neurons:

$$Q_{ih} = \frac{|W_{ih}|}{\sum_{i=1}^{n_i} |W_{ih}|} \tag{9}$$

2. For each input neuron i, the relative importance of all output weights attributable to the given input variable is obtained as:

$$RI(\%)_i = \frac{\sum_{h=1}^{n_h} Q_{ih}}{\sum_{h=1}^{n_h} \sum_{i=1}^{n_i} Q_{ih}} \times 100 \tag{10}$$

where n_h is the number of the hidden neurons. Fig. 6 shows the results of the sensitivity analysis accomplished through the described method.

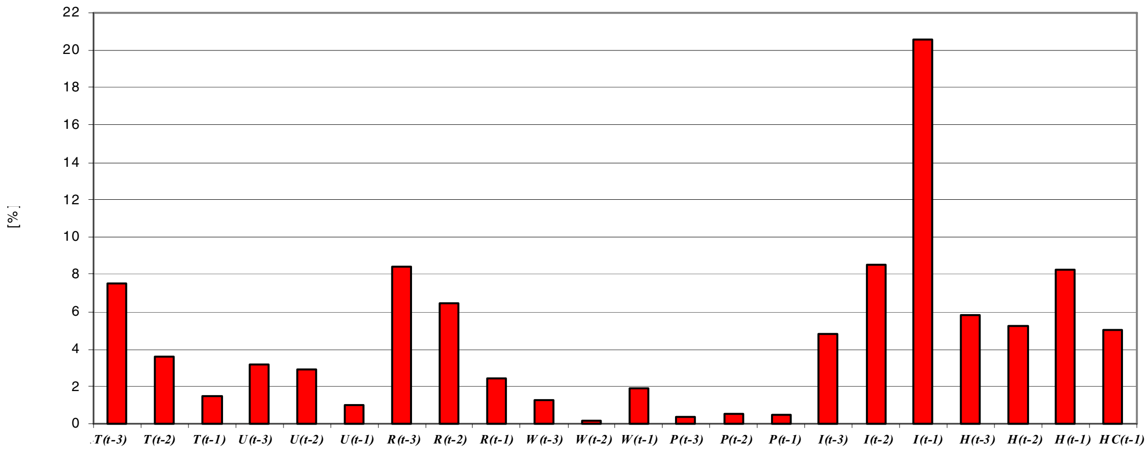

Fig. 6. Summary of the result of the sensitivity analysis accomplished on the selected network.

6. Conclusions

The paper presents a model based on an Elman recurrent neural network for the prediction, one hour ahead, of the intensity of the electric current supplied to the residential users located within a particular area of the town of Palermo (Italy). The model takes as inputs weather data as well as data related to the electricity consumptions. Furthermore, a special input (*HC index*) related to the presence and use of the AC appliances in the investigated area was added to the model.

The achieved results are quite interesting: the percentage prediction error computed for a test week are respectively 1.5% for the mean error and 4.6% for the maximum error.

It is noteworthy how the utilization of the *Humidex index* (*H*) improved the forecasting performances of the model (compare the first and the second row of Tab. 2) pointing out the importance of this thermal comfort index for the assessment of the influence of local thermal-hygrometric conditions on the comfort sensation felt by the occupants of the dwellings and, indirectly, on the utilization of the AC devices.

The results are certainly affected by the rough approximations which were necessary for the estimation of the number of AC installed and used in the investigated area. However, the described model represents an original approach able to supply some information about the household electricity consumption even in the lack of reliable data about the local spread of AC appliances.

References

[1] European Commission. Energy and transport in figures, Brussels; 2010.

[2] European Commission - Directorate-General for Energy and Transport (DG TREN). EU Energy in figures 2010.

[3] Terna – Rete Elettrica Nazionale, Dati statistici – Consumi energia elettrica per settore merceologico. Available on-line at: www.terna.it.

[4] Messineo A, Panno G. LNG cold energy use in agro-food industry: a case study in Sicily. *Journal of Natural Gas Science and Engineering* 2011; **3**: 356-363.

[5] Messineo A, Panno D. Potential applications using LNG cold energy in Sicily. *International Journal of Energy Research* 2008; 32: 1058–1064.

[6] Bessec M,Fouquau J. The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach. *Energy Economics* 2008; **30**:2705–2721.

[7] Izquierdo M, Moreno-Rodríguez A, González-Gil A, García-Hernando N. Air conditioning in the region of Madrid, Spain: An approach to electricity consumption, economics and CO_2 emissions. *Energy* 2011; **36**:1630-1639.

[8] Santamouris M, Argiriou A. Renewable energies and energy conservation technologies for buildings in southern Europe. *Int. J. Solar Energy* 1994; **15**:69–79.

[9] CoAer Gr.6. Indagine statistica sul mercato dei componenti per impianti di condizionamento dell'aria - Rilevazione annuale sul fatturato risultati anno 2008. Available online upon registration at: http://www.coaer.it/indagine.

[10] CoAer Gr.6. Indagine statistica sul mercato dei componenti per impianti di condizionamento dell'aria - Rilevazione annuale sul fatturato risultati anno 2009. Available online upon registration at: http://www.coaer.it/indagine.

[11] CoAer Gr.6. Indagine statistica sul mercato dei componenti per impianti di condizionamento dell'aria - Rilevazione annuale sul fatturato risultati anno 2010. Available online upon registration at: http://www.coaer.it/indagine.

[12] Elman JL. Finding structure in time. *Cognitive science* 1990; **14**:179-211.

[13] Xia C, Wang J, McMenemyc K. Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks. *Electrical Power and Energy Systems* 2010; **32**:743–750.

[14] Tso GKF, Yau KKW. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy* 2007; **32**:1761–1768.

[15] Amjady, N.; Keynia F. A New Neural Network Approach to Short Term Load Forecasting of Electrical Power Systems. *Energies* 2011; **4**:488-503.

[16] Lin WM, Gow HJ, Tsai MT. Electricity price forecasting using Enhanced Probability Neural Network. *Energy Conversion and Management* 2010; **51**:2707-2714.

[17] Hüsken M, Stagge P. Recurrent neural networks for time series classification. *Neurocomputing* 2003; **50**:223-235.

[18] Sarle WS. Stopped training and other remedies for overfitting. In: *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*. Convention Center and Vista Hotel, Pittsburgh, PA; 1995, vol. 27, p. 352-360.

[19] Fanger PO. *Thermal comfort*. McGraw-Hill, Danish Technical Press. Copenhagen, 1970

[20] Epstein Y, Moran DS. Thermal comfort and the heat stress indices. *Industrial Health* 2006; **44(3)**:388-98.

[21] Thom EC. The discomfort index. *Weatherwise*, 1959; **12**:57–60.

[22] Masterton JM, Richardson F.A. Humidex, a method of quantifying human discomfort due to excessive heat and humidity, CLI 1-79. *Environment Canada, Atmospheric Environment Service*, Downsview, Ontario, 1979

[23] ASHRAE. *ANSI/ASHRAE Standard 55-1992. Thermal environmental conditions for human occupancy*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. Atlanta, GA, 1992.

[24] Kyle, WJ. The Human Bioclimate of Hong Kong. In: Brazdil, R. & M. Kolar, editors. *Contemporary Climatology, Proc. of COC/IGU Meeting*, 15-20 August, 1994, Brno, Czech Republic, Masaryk University, p. 345-350.

[25] Tselepidaki I, Santamouris M, Moustris C, Poulopoulou G. Analysis of the summer discomfort index in Athens, Greece, for cooling purposes. *Energy and Buildings* 1992; **18**:51-56.

[26] SNNS - Stuttgart Neural Network Simulator, url: http://wwwra. informatik.uni tuebingen.de/SNNS.

[27] Riedmiller M., Braun H. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In: *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, CA , USA, 28th March - 1st April 1993 (ICNN 93), 1993.

[28] Williams RJ, Zipser D. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation* 1989, **1**:270-280.

[29] Anderson JA, Silverstein JW, Ritz SA. Jones R.S. Distinctive features, categorical perception, and probability learning: Some applications of a neuron model. In: Anderson JA & Rosenfeld E, editors. *Neurocomputing: Foundations of Research*, Cambridge, MA: MIT Press, 1988.

[30] Garson GD. Interpreting neural network connection weights. *Artificial Intelligence Expert* 1991; **6**:47-51.

[31] Goh ATC5. Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering* 1995; **9**:143-151.