

# Building lighting energy consumption modelling with hybrid neural-statistic approaches

F. Lauro<sup>1</sup>, C. Meloni<sup>2</sup> and S. Pizzuti<sup>2</sup>

<sup>1</sup>University of Calabria, fiorellalauro@gmail.com

<sup>2</sup>Energy New technologies and sustainable Economic development Agency (ENEA), Casaccia R.C., Via Anguillarese 301, 00123 Rome, Italy, {claudia.meloni, stefano.pizzuti}@enea.it

**Abstract.** In the proposed work we aim at modelling building lighting energy consumption. We compared several classical methods to the latest Artificial Intelligence modelling technique : Artificial Neural Networks Ensembling (ANNE). Therefore, in this study we show how we built the ANNE and a new hybrid model based on the statistical-ANNE combination. Experimentation has been carried out over a three months data set coming from a real office building located in the ENEA ‘Casaccia’ Research Centre. Experimental results show that the proposed hybrid statistical-ANNE approach can get a remarkable improvement with respect to the best classical method (the statistical one).

## 1 Introduction

Building energy consumption represents about 30%-40% of the global energy consumption [1] and it is the cause of about 40% of CO<sub>2</sub> emissions [2]. Therefore, the study of building energy demand has got in the recent years a remarkable relevance [3] in order to improve the management of existing buildings and the design of the new ones. In this context, having reliable energy estimations, and thus accurate models, is the key for energy efficiency with remarkable economic and environmental advantages.

In this scenario, at present there are three different approaches [4] for modelling energy consumption in buildings : Statistic Modelling (SM), Simulation Programs (SP) and Intelligent Computer Systems (ICS). The first one, known also as inverse modelling [5], is based on the building behaviour. With this approach a priori hypothesis about the model structure are made and the internal model parameters are tuned up through statistical analysis methods and the most popular techniques are linear regression and multivariate analysis [4]. Therefore the structure of the models is pretty straightforward but sophisticate statistical analysis methods are needed. The second approach, known also as direct modelling [5], starts from the physical description of the building which feeds a simulation program. Such methods need very accurate ambient information, a highly detailed building description and information about the occupants behaviour. All this information makes this approach computationally very expensive. The last approach, known also Artificial Intelligence (IA) approach [4], is based on Expert Systems (ES) and Artificial Neural Networks (ANN). ES are computer systems [6] that emulates the decision-making ability of a human expert. ES are designed to solve complex problems by reasoning about knowledge, like an expert, and it has a unique structure, different from traditional programs, which is divided into two parts, one fixed,

independent of the expert system: the inference engine, and one variable: the knowledge base. ANN [7,8] are a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. An ANN consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modelling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

The application of these methods depends on their own characteristics [4]. SM is mainly used in energy modelling of clusters of buildings [9,10] or in the design of areas where different end users are present. SM is mostly applied as energy estimation tools of single buildings [11] in design and retrofit interventions. ICS are somehow in between the two because can be applied to single as well as building clusters [12,13], moreover ICS are very effective in diagnosis, automation, control and management optimization.

Therefore, in this paper we describe a novel ICS for energy consumption modelling aimed at making diagnosis systems and controlling a complex building.

## 2 Methods

In this paragraph we shortly describe the modelling techniques we compared in the experimentation.

### 2.1 Naïve Model

In order to perform a meaningful comparison for the forecasting, a naive model should be introduced in order to quantify the improvement given by more intelligent and complex forecasting techniques. For seasonal data a naïve model might be defined as:

$$x_t = x_{t-s} \quad (1)$$

with S the appropriate seasonality period. This model gives a prediction at time t presenting the value observed exactly a period of S steps before. For this work, after the considerations of the previous section, we put the value of  $S = 24 \times 7 = 168$  which corresponds to a week given that the sampling considered is hourly.

### 2.2 Simple Linear Regression

In statistics, linear regression is an approach to modelling the relationship between a scalar variable Y and one or more explanatory variables denoted X. The case of one explanatory variable is called simple regression.

$$Y=aX+b \quad (2)$$

Therefore, the regression problem turns into the optimal setting of the parameters in order to explain the functional relation between X and Y.

The case with more than one explanatory variable is called multiple regression.

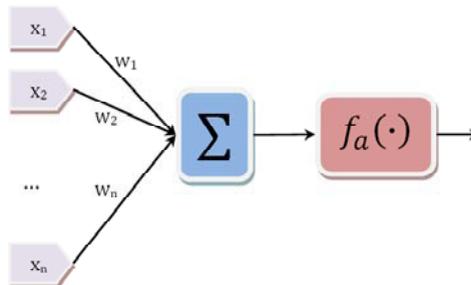
### 2.3 Statistical Model

One the simplest and most widely used models is to build an average weekly distribution of the consumption sampled hourly. Thus, from the data we compute for each day the average consumption hour by hour in such a way that we get an average distribution made of  $24 \times 7 = 168$  points.

## 2.4 Artificial Neural Networks

Artificial Neural Networks (ANN) [7,8] are computational models which try to simulate some properties of biological neural networks in order to solve complex modelling problems of non-linear systems. An ANN is an interconnected group of artificial neurons (called also nodes) that uses a mathematical or computational model for information processing based on a connectionistic approach to computation. In more practical terms ANN are non-linear data modelling or decision making tools which can be used to model complex relationships between inputs and outputs or to find patterns in data. ANN are referred also as black-box or data-driven models and they are mainly used when analytical or transparent models cannot be applied. Building such models needs several stages as input analysis and training through algorithms which minimize the error between the real values to be modelled and the ANN output. ANN demonstrated their effectiveness in modelling many real-world applications.

Once modelling an ANN model, we must take into account three basic components. First, the synapses of the biological neuron are modelled as weights. Let's remember that the synapse of the biological neuron is the one which interconnects the neural network and gives the strength of the connection. For an artificial neuron, the weight is a number, and represents the synapse. A negative weight reflects an inhibitory connection, while positive values designate excitatory connections. The following components of the model represent the actual activity of the neuron cell. All inputs are summed altogether and modified by the weights. This activity is referred as a linear combination. Finally, an activation function controls the amplitude of the output. Mathematically, this process is described in figure 1.



**Fig. 1.** Artificial neuron model

From this model the activity of the neuron can be shown to be:

$$y = f_a(\sum w_i x_i - \theta) \quad (3)$$

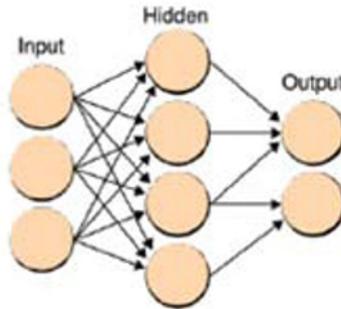
where  $\theta$  is a threshold called BIAS (Basic Input Activation System) which identifies the sensitivity of the neuron to respond to the external inputs. The most common function used to model  $f_a$  are the hyperbolic tangent, the sigmoid and the linear function.

Therefore each unit performs a relatively simple job: receive input from neighbours or external sources and use this to compute an output signal which is propagated to other units. Apart from this processing, a second task is the adjustment of the weights. The system is inherently parallel in the sense that many units can carry out their computations at the same time. Within neural systems it is useful to distinguish three types of units: input units which receive data from outside the neural network, output units which send data out of the neural network, and hidden units whose input and output signals remain within the network.

The way units are connected defines the network topology or architecture. In the past years many of them have been studied and the most widely used and is the feed-forward one. In this network

structure neurons are grouped into layers. There exists at least two layers, the input and the output one, which are those gathering the corresponding input and output variables. This basic structure is also known as perceptron [14]. Moreover, in order to let the model cope with non-linear problems, it is possible to add one or more intermediate layers, known as hidden layers. These models are also known as multi-layer perceptrons (MLP)[15].

The flow of data from input to output units is strictly in one direction (forward). The data processing can extend over multiple (layers of) units, but no feedback connections are present, that is, connections extending from outputs of units to inputs of units in the same layer or previous layers.

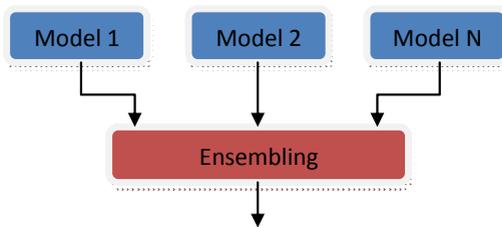


**Fig. 2.** Feed-forward neural network topology

A neural network has to be configured such that the application of a set of inputs produces (either 'direct' or via a relaxation process) the desired set of outputs. Various methods to set the strengths of the connections exist but the most used way is to 'train' the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule.

## 2.5 Ensembling Methods and Hybrid Models

The term 'ensemble' describes a group of learning machines that work together on the same task, in the case of ANN they are trained on some data, run together and their outputs are combined as a single one. The goal is obtain better predictive performance than could be obtained from any of the constituent models.



**Fig. 3.** Ensembling

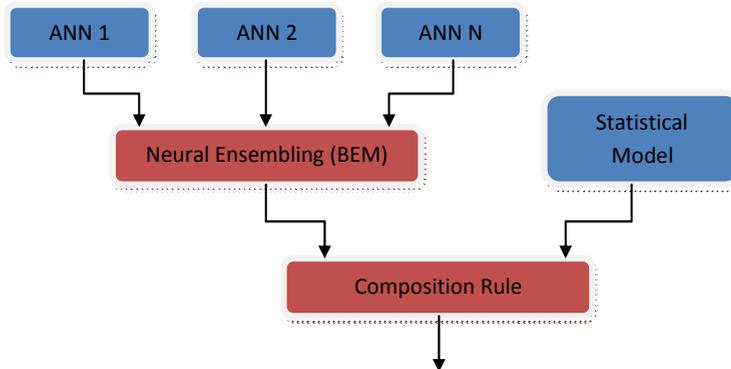
In the last years several ensembling methods have been carried out [16,17,18]. The first one, also known as Basic Ensemble Method (BEM), is the simplest way to combine  $M$  neural networks as an arithmetic mean of their outputs  $y_i$ . This method can improve the global performance [19,20] although it does not takes into account that some models can be more accurate than others. This method has the advantage to be very easy to apply.

A direct BEM extension is the Generalised Ensemble Method (GEM) [19,20] in which the outputs of the single models are combined in a weighted average where the weights have to be properly set, sometimes after an expensive tuning process.

Other methods are Bootstrap AGGREGatING (BAGGING) [21] and Adaboost [22,23].

Hybrid models are an extension of the ensembling approach in the sense that the final goal is to combine different models in such a way that the accuracy of the composition is higher than the best of the single models. The difference is that the combination is performed among highly heterogeneous models, that is models generated by different methods with different properties and thus the composition among them is a complex rule taking into account the peculiarities of the models and/or of the problem itself.

Therefore, in this work we propose an hybrid model which combines an ANN ensemble with the statistical model.



**Fig. 4.** Proposed Hybrid Model

The composition rule is the following :

$$\begin{aligned} &\text{If } |SM_t - X_{\text{real}_t}| < \theta \text{ Then } y_t = SM_t \\ &\text{Else } y_t = ANNE_t \end{aligned} \quad (4)$$

Where,  $SM_t$  is the consumption estimation of the statistical model at time  $t$ ,  $X_{\text{real}_t}$  is the real consumption measure at time  $t$ ,  $ANNE_t$  is the neural ensembling estimation at time  $t$ ,  $y_t$  is the final output.

This basically means that if we are in normal statistical conditions (thus where the statistical model makes a small error) then use as consumption estimation the statistical one (which is very accurate in this condition), else (when out of normal statistical situations) take the neural ensembling estimation.

### 3 Experimentation

In this paragraph we test and compare the methods presented in the previous section. The test case has concerned the energy consumption modelling of an office building (building ‘C59’) located in the ENEA ‘Casaccia’ Research Centre. The data set is made of 3 months of measurement (from September to November 2009) corresponding to 2184 hourly samples. each sample consists of 10 measurements : month, day of the month (1-31), day of the week (1-7), time, working day (true/false), occupancy, solar radiation, outdoor temperature, sunset time, light consumption. The first nine are used as input of the neural models and the last as the target (output). The hybrid model (4) has used a threshold value  $\theta = 1$  which turned out to be the optimal one (tab.1).

The ANN are MLP with 9 inputs, 10 hidden neurons and one output (the lighting consumption) with hyperbolic tangent as activation function for the hidden neurons and linear for the output.

Training has been performed with Matlab (ver. R2010a) through the Levenberg-Marquardt algorithm stopping after 1000 iterations. The reported result are averaged over 10 different runs (with standard deviation in brackets) and the ensemble is therefore made by the same 10 models.

**Table 1.** Hybrid model parameter tuning

$\theta$	MAE
0.1	0.91
0.5	0.84
<b>1</b>	<b>0.82</b>
1.5	0.84
2	0.89

The linear regression is made using as independent variable the occupancy and then by properly tuning the parameters.

The data set has been split in two parts : training (10 weeks) and validation (3 weeks, one for each month, corresponding to 360 samples) and the reported results refer to the last one.

The performance has been measured according to the Mean Absolute Error (MAE) and the Mean Nominal Error (MNE).

$$\text{MAE} = 1/N * \sum |y_i - y'_i| \quad (5)$$

$$\text{MNE} = \text{MAE} / ( \max(Y) - \min(Y) ) \quad (6)$$

Where  $y_i$  is the real output,  $y'_i$  is the estimated output,  $Y$  is the real values data set and  $N=|Y|$ .

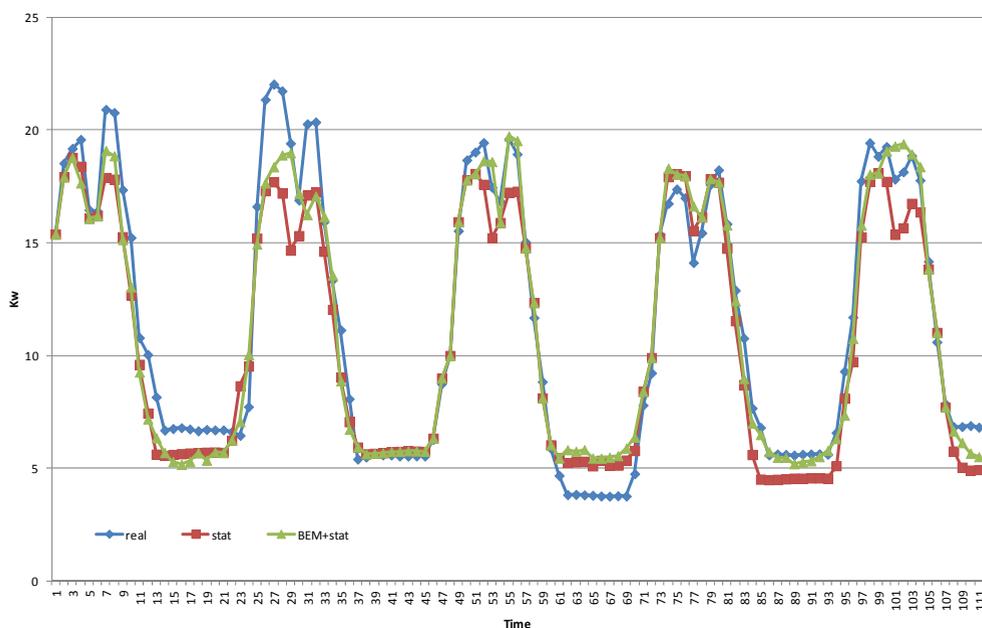
**Table 2.** Experimental results (testing)

	Naïve	Lin.Regression	Statistical	ANN	BEM	BEM+Stat
MAE (kW)	2.95	1.37	0.97	1.22 ( $\pm 0.10$ )	0.95	<b>0.82</b>
MNE	14.8%	6.8%	4.9%	6.1% ( $\pm 0.5\%$ )	4.8%	<b>4.1%</b>

From this experimentation we can see that the proposed method (BEM+Stat) clearly outperforms all the others. The reason for that is that the proposed approach is capable to provide reliable estimations when out of standard conditions (fig.5) because it takes into account several input features (as occupancy) which affect the energy consumption.

Moreover, it is interesting to point out that statistical modelling performs pretty well, even better than the single neural models (ANN). These get a remarkable accuracy, and an error slightly lower than the statistical model, only as an ensemble.

The accuracy achieved by the proposed hybrid model is such that it can be applied for intelligent monitoring, diagnostic systems and optimal control in order to reduce energy consumption.



**Fig. 5.** Modelling comparison (testing results).

## 4 Conclusions

In this work we proposed a new approach aimed at modelling building lighting energy consumption. The basic idea is to build a new hybrid model based on neural networks ensembling and statistic modelling in order to get the best of the two. Experimentation has been carried out over a three months data set coming from a real office building located in the ENEA ‘Casaccia’ Research Centre and experimental results show that the proposed method can get a remarkable improvement with respect to the best classical method .

The reason for that is that the neural ensembling model is capable to provide reliable estimations when out of standard conditions because it takes into account several input features (as occupancy) which affect the energy consumption.

The accuracy of the proposed model is such that it can be applied for intelligent monitoring, diagnostic systems and optimal control in order to reduce energy consumption.

As future work we are going to apply the same approach to model other building energy consumptions as conditioning and thermal flows. Moreover, we are also going to try applying more sophisticated ensembling methods in the neural modelling stage as well as to try different composition criteria between neural and statistical models.

## References

1. United Nations Environment Programme, Buildings Can Play Key Role In Combating Climate Change (2007).  
<http://www.unep.org/Documents.Multilingual/Default.asp?DocumentID=502&ArticleID=5545&l=en>.
2. Yudelson J., Greening Existing Buildings, Green Source/McGraw-Hill, New York (2010).
3. European Union, Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings. Official Journal of the European Communities (2003).

4. Pedersen L., Use of different methodologies for thermal load and energy estimations in buildings including meteorological and sociological input parameters. *Renewable and Sustainable Energy Reviews*, Elsevier, 11, 998-1007 (2007).
5. Rabl A., Rialhe A., Energy signature models for commercial buildings: test with measured data and interpretation. *Energy and Buildings*, Elsevier, 19, 143-154 (1992).
6. Jackson, Peter, *Introduction To Expert Systems* (3 ed.), Addison Wesley, p. 2, ISBN 978-0-201-87686-4 (1998).
7. Arbib M.A., *The Handbook of Brain Theory and Neural Networks*, The MIT Press, Cambridge (MA) (1995).
8. Haykin S., *Neural Networks, a comprehensive foundation* (2nd edition), Prentice Hall, New Jersey (1999).
9. Caldera M., Corgnati S. P., Filippi M., Energy demand for space heating through a statistical approach: application to residential buildings. *Energy and Buildings*, Elsevier, 40, 1972–1983 (2008).
10. Rauhala K., A simple computer model for estimating the energy consumption of residential buildings in different microclimatic conditions in cold regions. *Energy and Buildings*, Elsevier, 15–16, 561–569 (1991).
11. Farahbakhsh H., Ugursal V. I., Fung A. S., A residential end-use energy consumption model for Canada. *International Journal of Energy Research*, 22, 1133-1143 (1998).
12. Mihalakakou G., Santamouris M., Tsangrassoulis A., On the energy consumption in residential buildings. *Energy and Buildings*, Elsevier, 34, 727-736 (2002).
13. Tso G. K. F., Yau K. K. W., Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. *Energy*, Elsevier, 32, 1761-1768 (2007).
14. Rosenblatt, Frank, *The Perceptron--a perceiving and recognizing automaton*. Report 85-460-1, Cornell Aeronautical Laboratory (1957).
15. Rosenblatt, Frank. *x. Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington DC (1961)
16. Krogh A., Vedelsby J., Neural network ensembles, cross validation and active learning, in Tesauro G., Touretzky D. S., Leen T. K., editors, *Advances in Neural Information Processing Systems*, 7, 231-238, MIT Press (1995).
17. Liu Y., Yao X., Ensemble learning via negative correlation. *Neural Networks*, 12(10), 1399-1404 (1999).
18. Breiman L., Combining Predictors, in Sharkey A.J.C., (ed.): *Combining Artificial Neural Nets – Ensemble and Modular Multi-net Systems*, Springer, Berlin, 31-50 (1999).
19. Perrone M. P. e Cooper L. N.. When networks disagree: ensemble methods for hybrid neural networks. *Neural Networks for Speech and Image Processing* by R.J. Mammone, ed. Chapman-Hall (1993).
20. Bishop C. M.. *Neural Networks for Pattern Recognition*. Oxford University Press, 364 – 369 (1995).
21. Kohavi R. e Bauer E. An empirical comparison of voting classification algorithms: Bagging, boosting and variants. *Machine Learning*, 36 , 105 – 142 (1999).
22. Drucker H. Improving regressors using boosting techniques. In Douglas H. Fisher, ed., *ICML*, Morgan Kaufmann 107 – 115 (1997).
23. Avnimelech R. and Intrator N. Boosting regression estimators. *Neural Computation*, 11 491 – 513 (1999).