Zurich, June 15-17 2016
Sustainable Built Environment (SBE) Regional Conference

Expanding Boundaries:
Systems Thinking for the Built Environment

A METHODOLOGY FOR ENERGY AUDIT FOR COMMERCIAL BUILDINGS USING MACHINE LEARNING TOOLS

C. Deb[1]*, S. E. Lee[1], J. Yang[1], K. W. Shah[1]

[1] Department of Building, National University of Singapore, Singapore
*Corresponding author; e-mail: chirag.deb@u.nus.edu

Abstract
Energy auditing is a process for evaluating energy conservation opportunities in buildings. According to ASHRAE standards, energy audits can be performed at three levels based on the extent of the audit scope. This requires extensive assessment arrangements including instrumentation and data collection, data analysis and identifying energy saving potential of the building systems. However, these steps make the audit process time consuming and often expensive. This study presents an ongoing methodology that aims to enhance the audit process through a data driven approach. Building and energy related data from 24 commercial buildings is used to develop the methodology by analysing their energy audit reports. Several building factors that influence energy consumption have been analysed and ranked based on their correlation coefficients. An energy savings prediction model is developed using Artificial Neural Network (ANN) based on the energy savings data as gathered from the 24 audit reports. The results show that the ANN model predicts the energy savings accurately for a building by simply using the basic building data. The details of the ANN model development have also been highlighted. It is also intended to expand the dataset to about sixty buildings to make the results more accurate.

Keywords:
Energy auditing; Prediction model; Artificial Neural Networks; Building energy consumption

1 INTRODUCTION

The International Energy Agency has identified energy efficiency in buildings as one of the five measures to secure long-term decarbonisation of the energy sector (© OECD/IEA 2015 *World Energy Outlook Special Report*, IEA Publishing. License: [*http://www.iea.org/t&c/termsandconditions/*]) [1]. Along with environmental benefits, building energy efficiency also presents vast economic benefits. Buildings with efficient energy systems and management strategies have much lower operating costs. Many countries have now accelerated the implementation of energy codes and regulations for various building types. These regulations outline basic requirements to achieve an energy efficient design for new buildings with a view to reduce the final energy consumption and related CO_2 emissions. In addition, many computer softwares have also been developed and widely implemented for energy efficient design of new buildings. Some of the most popular ones are EnergyPlus, DOE-2, eQUEST, IES, ECOTECT etc. However, once the building is functional, the energy behavior of a building is governed by many factors such as weather conditions, occupancy schedule and behavior, thermal properties of building materials, complex interactions of the energy systems like HVAC and lighting etc. Due to these complex interactions, accurate, computation of energy consumption is very difficult. For these reasons, data driven techniques for building energy analysis of existing buildings are very crucial. These techniques are based on past recorded data and building factors like Gross Floor Area (GFA), occupancy etc.

1.1 Energy Auditing

Energy audits are the first step to improve the energy efficiency of buildings and industrial facilities. Generally, three levels of energy audits

G. Habert, A. Schlueter (eds.): Expanding Boundaries © 2016 vdf Hochschulverlag AG an der ETH Zürich
DOI 10.3218/3774-6_31, ISBN 978-3-7281-3774-6, http://vdf.ch/expanding-boundaries.html

can be distinguished as outlined by ASHRAE standards [2]:

- Walk-through audit - consisting typically of a short on-site visit of the facility to identify areas where simple and inexpensive actions can provide immediate energy use and operating cost savings. Usually it also involves utility cost analysis that includes a careful evaluation of metered energy uses and operating costs of the facility. Typically, the utility data over several years are evaluated to identify the patterns of energy use, peak demand, weather effects, and potential for energy savings.
- Standard energy audit - consisting of a comprehensive energy analysis for the energy systems of the facility. In particular, the standard energy audit includes the development of a baseline for the energy use of the facility and the evaluation of the energy savings and the cost effectiveness of appropriately selected energy conservation measures.
- Detailed energy audit – considered the most comprehensive and time-consuming energy audit type. It includes the use of instruments to measure energy use for the whole building and for some energy systems within the building (for instance by end uses: lighting systems, office equipment, fans, chillers, etc.). In addition, sophisticated computer simulation programs are typically considered for detailed energy audits to evaluate and recommend energy retrofits for the facility.

The purpose of the audit is to identify energy conservation measures (ECMs). Several tools have been developed to aid the audit process. Hong et al. [3] presented the Commercial Building Energy Saver (CBES), an energy retrofit analysis toolkit. This tool calculates the energy use of a building, identifies and evaluates retrofit measures in terms of energy savings, energy cost savings and payback. The CBES Toolkit includes a web app (APP) for end users and the CBES Application Programming Interface (API) for integrating CBES with other energy software tools. Hestnes and Kofoed [4] evaluated a set of retrofitting strategies designed for ten existing office buildings. This was done by examining different low energy retrofitting measures in terms of energy, indoor environment, and economy, and by using this as a basis for the development of general retrofitting strategies and design guidelines. Chuah et al. [5] introduced ROBESim (Retrofit-oriented building energy simulator). ROBESim is based on the popular EnergyPlus framework, and relies on EnergyPlus for most of the supported computations. By using the retrofit modules in ROBESim, the user can quickly and easily generate building models to perform retrofit comparison simulations. A recent review by Lee et al. [6] has divided such existing retrofitting tool into three divisions:

- Toolkits with empirical data-driven methods
- Toolkits using normative calculations
- Toolkits with physics-based energy modelling and simulation

The data-driven methods have been widely used to predict building energy usage, from simple benchmarking to more complex regression modelling. These models rely on past recorded data for energy consumption and operational parameters. Among these methods, the most widely implemented in modelling and forecasting is the machine learning method, which predominantly includes Artificial Neural Network (ANN) and Support Vector Machine (SVM) [7][8][9][10]. This study utilizes the ANN model that has gained momentum in building energy consumption studies for its application ranging from forecasting to prediction of saving potential for buildings. The objective of this study is to develop a prediction model for energy savings in commercial buildings using a data-driven approach. For this, 24 building energy audits have been analyzed and used to develop the prediction model using ANN. More details on the ANN methodology and the design of its parameters is presented in the next section.

2 METHODOLOGY

2.1 Data collection

This study is based on data collected from accredited Energy Service Companies (ESCOs) in Singapore. The data collection for model development involves reviewing through the energy audit reports as are provided by three such ESCOs. For cases where data is unavalibale or inconsistent across audit reports, linear regression or averaging is usually used to impute the missing values. However, in the cases studied, all data were available. An energy audit report contains detailed analysis of energy distribution and usage by the various energy consuming systems in a building. The three major energy consuming systems in a building are the following:

(i) Air Conditioning
(ii) Lighting
(iii) Plug loads

These three systems account for more than 80% of the total energy consumed in a building. The air conditioning system is the major energy consuming system (about 55% of total energy) out of the above three and is the central focus for this part of the study. The air conditioning system comprises of two parts. First is the chiller plant room and the second are the Air Handling Units (AHUs). Depending on the building configuration, there are different divisions for the amount of energy consumed by the chiller plant and the

G. Habert, A. Schlueter (eds.): Expanding Boundaries © 2016 vdf Hochschulverlag AG an der ETH Zürich
DOI 10.3218/3774-6_31, ISBN 978-3-7281-3774-6, http://vdf.ch/expanding-boundaries.html

AHUs. For a typical office building with central chiller plant, the amount of energy consumed by the chiller plant, (that includes the chiller power, chilled water pumps, condenser water pumps) corresponds to 60-70% of the air conditioning energy use. The rest is consumed by the AHUs. Along with the air conditioning energy use data, several building factors that influence building energy consumption and energy saving potential have been identified and studied for model development. The list of the eight factors studied is presented in the 'Results' section.

2.2 Artificial Neural Network (ANN) model

ANN is an intensely parallel network of processing units that can perform non-linear analysis. They learn the relationship between input and output variables by studying previously recorded data through a process called training. An ANN resembles the biological neural system, composed by layers of parallel elemental units, called neurons. The neurons are connected by a large number of weighted links, over which signals or information passes through. A neuron receives inputs over its incoming connections, combines the inputs, performs generally a non-linear operation, and then outputs the final results. In this study, the neural network adopted was a feed-forward multi-layer perceptron (MLP), which is among the most commonly used neural networks that learn from examples. A schematic diagram of the basic architecture is shown in Figure 1. It contains three layers – the input, hidden and output layers. Each layer is interconnected together by the connection strengths called weights. The training algorithm used in this study is the Levenberg-Marquardt (LM) algorithm. This algorithm performs well with function approximation problems like the one being dealt in this study. The learning process in an ANN involves determining the weight vectors. There are various algorithms that are used for this purpose. The aim of the training process is to minimize the squared error between the predicted and the measured outputs (1).

$$E = \sum \tfrac{1}{2}\left(O_p - O_m\right)^2 \qquad (1)$$

Here, E is the total error, O_p is the predicted output and O_m is the measured or desired output. E is minimized by the gradient descent method which involves computing the partial derivative of E with respect to each weight in the network. The most popular training algorithm is the backpropagation algorithm and its details can be found in the work of Rumelhart et al [11]. Once the weights are detemrnied using this backpropagation method, the ANN model is considered ready and can be tested for new data cases.

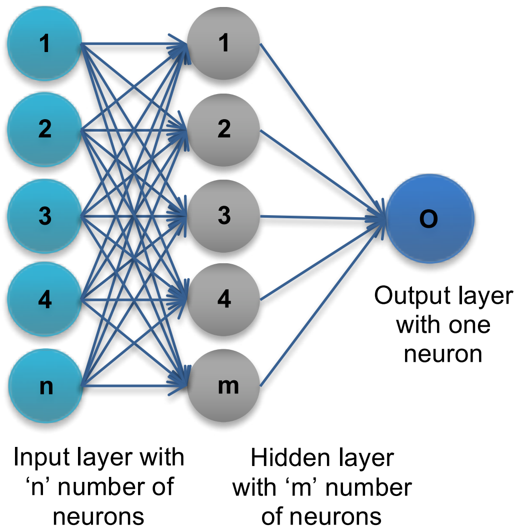

Fig. 1: A typical ANN structure with input, hidden and output layers.

3 RESULTS AND DISCUSSION

The mathematical modelling using ANN requires a set of dataset with certain inputs and their corresponding outputs. This input-output matrix is used to map the relationship between input and output variables by computing the weights of the connections. The connections are present between the inputs and the neurons in the hidden layer as well as the neurons between the hidden and the output layer. For model training and testing, 16 and 8 audit reports have been assigned respectively. The first step in ANN modelling is to determine the input-output matrix that is then used to train the model. The output in this case is fixed at the potential energy savings as proposed by the ESCOs in their audit reports of the buildings. To determine the inputs for the ANN model, an approach which deals with analyzing each variable independently is employed. The variables associated with the energy performance of each building that are collected from the ESCOs are eight different varaibles or factors. However, it is important to select only the most important variable and use them as inputs for the ANN model developement. A preliminary analysis was performed to determine these variables and the following results are obtained, as presented in Table 1.

Energy Indicator	**R^2**
Annual electricity consumption	**0.58**
Average cooling load per day (RT)	**0.86**
Energy Use Index (EUI)	**0.52**
Chilled water supply temperature	**0.37**

G. Habert, A. Schlueter (eds.): Expanding Boundaries © 2016 vdf Hochschulverlag AG an der ETH Zürich
DOI 10.3218/3774-6_31, ISBN 978-3-7281-3774-6, http://vdf.ch/expanding-boundaries.html

Chilled water return temperature	**0.33**
Gross Floor Area (GFA)	**0.88**
Chiller plant efficiency	**0.88**
Operation hours	**0.27**

Table 1. Corelation coefficients (R^2) of energy indicators with measured potential savings.

The correlation of the GFA with the potential savings calculation yields a positive and high coefficient of determination (R^2 = 0.88) as seen in Figure 2. This shows that there is a good correlation in the given dataset between the GFA of a building and the amount of energy that can be saved by retrofitting the facility. It is normally seen that a large facility has a much wider scope of energy conservation possibility because the facility requires high capacity building systems like air conditioning and lighting system. In such cases, even a simple retrofitting solution, like changing the fluorescent ballast lamps to LED (Light Emitting Diodes) could lead to a high saving potential. Similarly, for the air conditioning system, the larger GFA, in most cases ensures more potential savings unless there is an entire change in the air conditioning system.

Another variable that provides an estimate on the potential savings is average cooling load per day for a building. If the buildings that belong to a similar cluster of GFA have variation in their average daily cooling load, it shows that either the buildings differ heavily in occupancy numbers or that one of the buildings have high cooling load for other reasons. For the data on 16 buildings with water-cooled chiller water system, shows that there is a direct correlation between these two variables with a R^2 = 0.88. Therefore, the average cooling load is also considered as an input for the ANN model. A similar correlation is observed between the chiller plant efficiency and potential savings. The BCA Platinum green mark rating chiller plant efficiency target for non-residential buildings with a cooling load more than 500 RT is 0.65. Going with this target, there seem to exist a high potential for energy savings based on improving the chiller plant efficiency only.

The three inputs selected for developing the ANN model are the GFA, average cooling load and the chiller plant efficiency. These constitute the input layer with three neurons corresponding to the three inputs and the output neuron as the predicted potential saving (Figure 3). The LM training algorithm is used to train the model by using the data provided by the ESCOs for eight buildings. The data is divided into 16 buildings data for training and 8 buildings data for testing the model. Figure 4 shows the correlation between the predicted savings by the model and the ones as proposed by the ESCOs. The correlation coefficient of the predicted savings by the ESCOs and that predicted by the model is 0.94. Although this correlation is high, it is to be noted that this is for the training dataset and does not any testing with an independent set of testing data. ANNs are very good in accurately modelling the training data and such a high correlation should not be taken as evidence of an accurate model. The model developed is accurate for the data set involved in training alone. However, the dataset is currently being expanded and it is targeted to obtain the audit report data for another fifty buildings. Once the dataset is updated, it is targeted to partition the dataset into data for training and testing and validating the model.

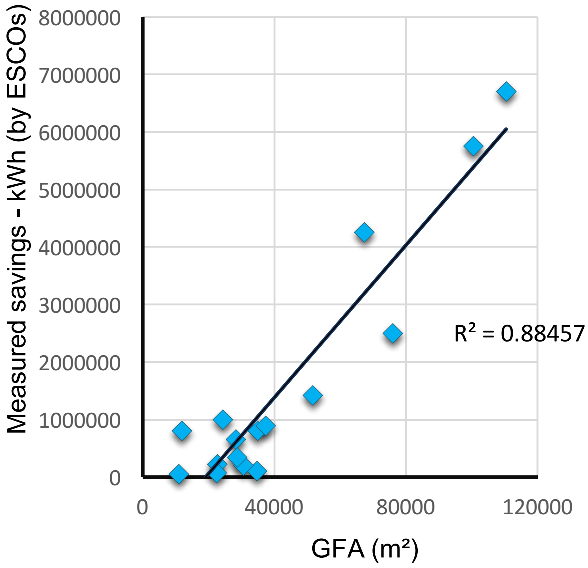

Fig. 2: Corelation of GFA with energy saving potential of a building as measured by the ESCOs.

G. Habert, A. Schlueter (eds.): Expanding Boundaries © 2016 vdf Hochschulverlag AG an der ETH Zürich
DOI 10.3218/3774-6_31, ISBN 978-3-7281-3774-6, http://vdf.ch/expanding-boundaries.html

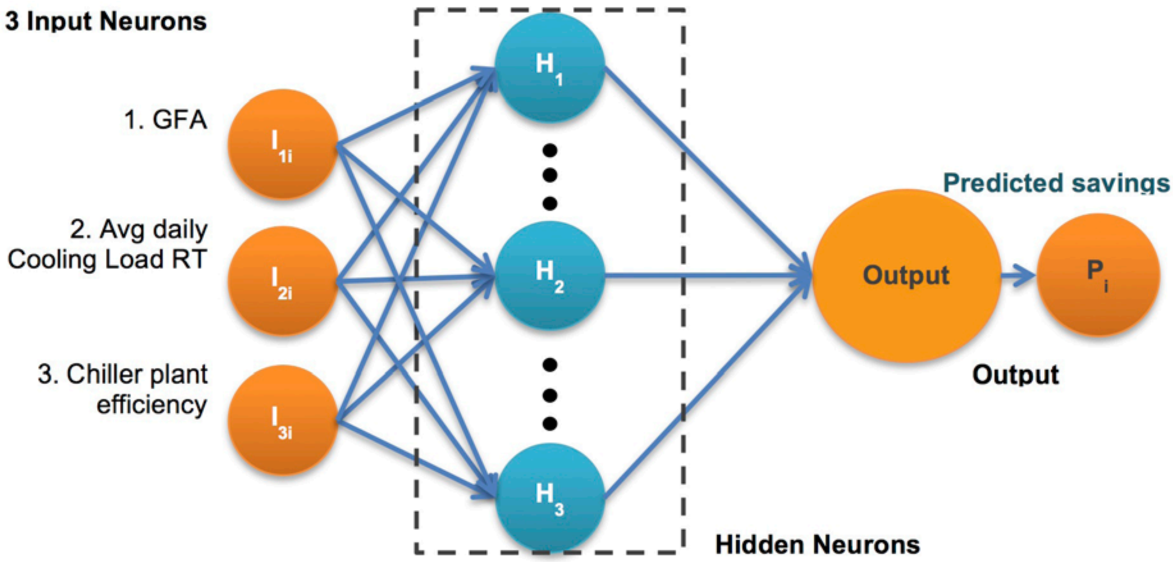

Fig. 3: The ANN structure for predicted savings potential.

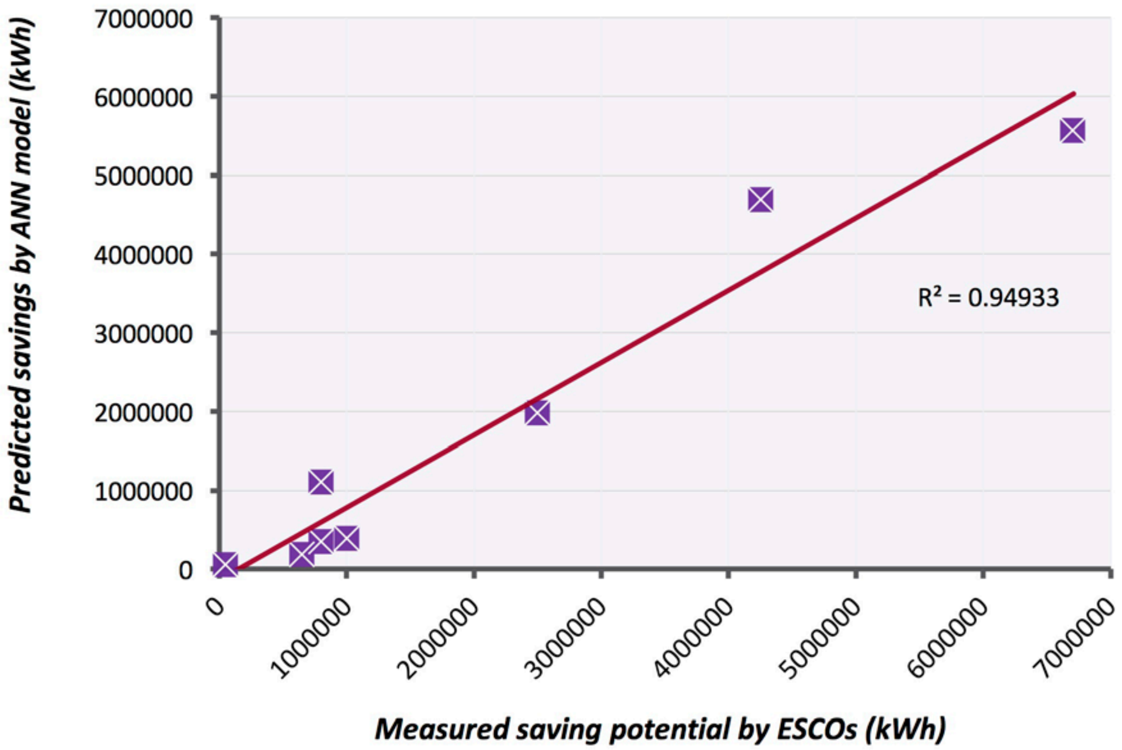

Fig. 4: Corelation between the measured savings by the ESCOs and the predicted savings by the ANN model.

G. Habert, A. Schlueter (eds.): Expanding Boundaries © 2016 vdf Hochschulverlag AG an der ETH Zürich
DOI 10.3218/3774-6_31, ISBN 978-3-7281-3774-6, http://vdf.ch/expanding-boundaries.html

4 SUMMARY

A detailed data-driven method using ANN is developed to predict the energy saving potential of a building. For this, 24 energy audit reports from 3 ESCOs are analysed and used to develop the prediction model. The inputs that have been selected are the GFA, average cooling load per day and chiller plant efficiency which exhibit R^2 values of 0.88, 0.86 and 0.88 respectively with potential savings. The model is trained using data from eight buildings. The developed model shows an accuracy of 0.94 between the potential accuracy as proposed by the ESCOs and the potential saving as predicted by the model. Such a model is expected to enhance the audit process by providing a guideline for potential savings calculation using simple building energy indicators like GFA, chiller plant efficiency etc. The process of expanding the dataset and data collection is underway with audit report data from another 50 buildings to be added to the dataset. Once the dataset is expanded, the accuracy of the model is also expected to improve.

5 ACKNOWLEDGEMENTS

The authors would like to thank the 4 ESCOs (G-Energy Pte Ltd; E2Green Pte Ltd; LJ Energy Pte Ltd; and EMSI Pte Ltd) for their collaboration and support to this study.

6 REFERENCES

1. IEA. Energy and climate change. 2015. doi:10.1038/479267b.

2. ASHRAE. Procedures for Conmmercial Building Energy Audit. Second Edi. Atlanta, USA: 2004.

3. Hong T, Piette MA, Chen Y, Lee SH, Taylor-Lange SC, Zhang R, et al. Commercial Building Energy Saver: An energy retrofit analysis toolkit. Appl Energy 2015;159:298–309. doi:10.1016/j.apenergy.2015.09.002.

4. Hestnes AG, Kofoed NU. Effective retrofitting scenarios for energy efficiency and comfort: results of the design and evaluation activities within the OFFICE project. Build Environ 2002;37:569–74. doi:10.1016/S0360-1323(02)00003-3.

5. Chuah JW, Raghunathan A, Jha NK. ROBESim: A retrofit-oriented building energy simulator based on EnergyPlus. Energy Build 2013;66:88–103. doi:10.1016/j.enbuild.2013.07.020.

6. Lee SH, Hong T, Piette MA, Taylor-Lange SC. Energy retrofit analysis toolkits for commercial buildings: A review. Energy 2015;89:1087–100. doi:10.1016/j.energy.2015.06.112.

7. Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16:3586–92. doi:10.1016/j.rser.2012.02.049.

8. Deb C, Eang LS, Yang J, Santamouris M. Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. Energy Build 2015.

9. Deb C, Eang LS, Yang J, Santamouris M. Forecasting Energy Consumption of Institutional Buildings in Singapore. Procedia Eng 2015;121:1734–40. doi:10.1016/j.proeng.2015.09.144.

10. Dong B, Cao C, Lee SE. Applying support vector machines to predict building energy consumption in tropical region. Energy Build 2005;37:545–53. doi:10.1016/j.enbuild.2004.09.009.

11. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6. doi:10.1038/323533a0.