Predicting Energy Usage Using Historical Data and Linear Models

Majeed Safa[a], Jeremy Allen[b], and Mahdi Safa[c]

[a] Department of Agricultural Management and Property Studies, Lincoln University, New Zealand
[b] Energy Solution Providers Ltd, Auckland, New Zealand
[c] Department of Civil and Environmental Engineering, University of Waterloo, Canada
E-mail: Majeed.safa@lincoln.ac.nz, jeremy.allen@espnz.co.nz, msafa@uwaterloo.ca

Abstract:

This paper presents a method to predict energy usage, based on weather conditions and occupancy, using a multiple linear regression model (MLR) in research office buildings. In this study, linear regression models of four research office sites in different regions of New Zealand were selected to show the capability of simple models to reduce margins of error in energy auditing projects. The final linear regression models developed were based on monthly outside temperatures and numbers of full time employees (FTEs). Comparing actual and predicted energy usage showed that the models can predict energy usage within acceptable errors. The results also showed that each building should be investigated as an individual unit.

Keywords: Energy prediction, energy auditing, linear regression model, office buildings, energy saving, optimization

1. Background:

Recent years have witnessed a remarkable increase in energy costs and increasing environmental concerns [1]. The optimal operation of office buildings is essential for reducing energy. Optimizing the energy consumption of office buildings requires robust energy monitoring and energy auditing systems [2]. In many commercial projects energy savings are estimated based on a comparison of current usage and the previous years data [3, 4]. Due to environmental conditions, changing occupation and other causes, energy usage in different years could change significantly. Therefore, modelling energy usage based on historical data has the potential to provide more accurate energy saving estimations. The objective of this study was the development of a new forecasting model to predict office building energy usage to measure energy savings during energy auditing projects.

Offices and the retail sector are the most intensive energy consumers in the non-domestic building sector and it is estimated that over 50% of energy usage was for non-domestic buildings [5]. In New Zealand, standard office buildings mainly consume electricity for their operation. A survey of energy sources by the Building Research Association of New Zealand (BRANZ) showed that 11% of non-residential buildings consumed natural gas while only 3.5% of non-residential buildings used diesel and/or fuel oil, mainly for heating [6]. The major electrical consumption in the buildings surveyed was for lighting, air-conditioning, plug loads and hot water [5, 6].

There have been several studies that analysed electricity consumption in different buildings. However, due to the lack of current advanced metering technologies, in most early energy modelling studies monthly electricity bills were used to analyse energy consumption in buildings [7].

As the first step to analyse energy usage in buildings, it was important to understand which factors were more important for energy consumption in different buildings. Kavousian, et al. [7] identified four main categories for energy consumption in residential buildings: weather and location, appliances and electronic stock, physical characteristics of the buildings, and occupancy. Nonetheless/However??, Issacs, et al. [6] classified the businesses in their survey based on business sector and activities, staff numbers, client numbers and operating periods.

Office buildings are very different in terms of their design, construction, occupancy and activity, which makes it too difficult to identify small numbers of them to represent the majority of office buildings [8]. Comparing different studies showed that environmental parameters were the main factors in most studies for estimating energy usage in buildings. Temperature, humidity and lux level were the main environmental parameters which can directly affect energy consumption in buildings

[6, 9].

In energy auditing projects and sustainability plans, predicting energy usage is a very important challenge. Several optimization methods have been applied for the energy usage estimation and a variety of modelling methods have been developed over the last decade [10-18].

2. Models:

In this study, four research buildings, including offices and laboratories, of the same research institute in different New Zealand regions, were investigated (Figure 1). Each site was investigated individually based on available historical monthly data.

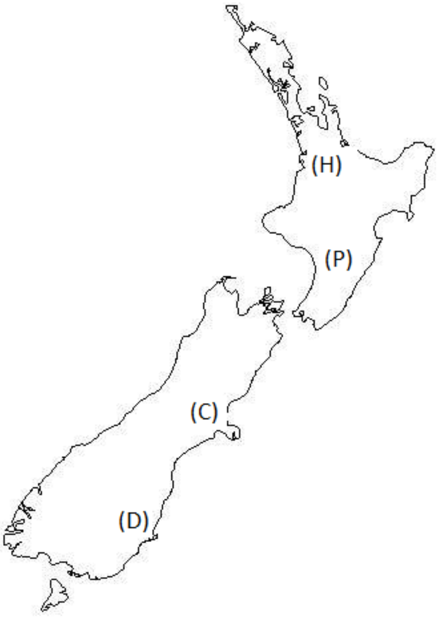

Figure 1. Site locations [Hamilton (H), Palmerston North (P), Christchurch (C), and Dunedin (D)] in New Zealand map,

The multiple linear regression model was selected and developed for predicting energy consumption. This model has been extensively used in energy modelling projects [9, 19-25]. Compared with nonlinear models, linear regression models are easier and more practical for solving the different problems [21].

For use in the model, it was necessary to select a limited number of variables without any selective bias [26]. A simple model with the highest r^2 was designed through a combination of forward, backward and stepwise regression adjustments. Terms were maintained in the final model if they were significant at $p=0.05$ [27]. In the first step, the relationship between energy consumption and each input variable was tested with a simple linear regression using the r^2 as the decision criterion. Then, a multiple linear regression model was developed for predicting the energy consumption as:

$$Y=a_0+a_1V_1+a_2V_2+\cdots+a_nV_n \qquad (1)$$

where a_0-a_n are the regression coefficients and V_0-V_n are the independent variables.

The model was in a linear form to represent the linear relationships of the dependent variable with the independent variable. After running the model, predictions on the validation data were estimated. In this study, the model was developed with a minimum number of variables to capture energy consumption in as simple a form as possible. After investigating several variables, occupancy (full time employees) and outside temperature were selected as the independent variables with which to develop the models. The monthly temperature was collected from a national weather database (NIWA) and property managers provided the number of employees.

Several methods of error estimation were proposed. The mean square error (MSE) over all training patterns (Eq. 2) was the most commonly used error indicator. MSE was very useful to compare different models; it showed a network's ability to predict the correct output. The MSE can be written as:

$$MSE=\frac{1}{2N}\sum_{i}^{N}(t_i-z_i)^2 \qquad (2)$$

where t_i and z_i are the actual and predicted outputs for the i^{th} training pattern, and N is the total number of training patterns [21, 28]. The root mean square error (RMSE) is another error estimation method, which shows the error in the units of actual and predicted data.

In this study, several models were developed and compared to find the best fit between the predicted and actual data. As shown in Figure 2 the final model of the Palmerston North, Christchurch, and Dunedin sites were developed based on 30 months of available historical data and the model of the Hamilton site was developed based on 18 months of data. The Christchurch model was affected by the 4 September 2010 Canterbury earthquake, which significantly reduced energy usage during September and October; therefore, the data of September and October 2010 were removed from the model. It was notable that the models developed, based only on outside temperature for all sites, were accurate with a very high correlation coefficient between actual and predicted energy usage. However, comparisons between models with one independent variable (temperature) and two independent variables (temperature and FTEs) showed that the FTEs can improve the correlation coefficients between actual and predicted energy usage.

You need to label these figures from A to D

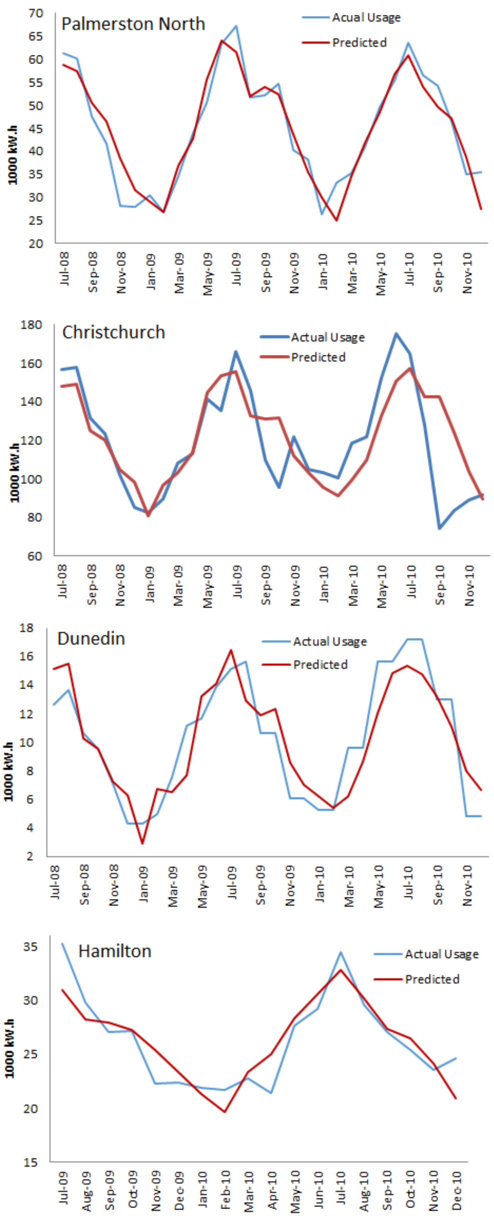

Figure 2. Predicted and actual energy usage in the four research offices.

3. Results

A comparison of the models showed that the monthly energy usage in research office buildings can be predicted by temperature and occupancy data. The energy usage mostly depended on weather conditions. The low usage in summer would be influenced mostly by outside temperature and Christmas and New Year holidays (summer in the Southern Hemisphere). It should be noted that energy usage in buildings would also be affected by a number of other factors.

Figure 2 shows that the predicted and actual data are matched in most months; however, for different numbers of days per month, for different numbers of holidays per month and some other factures can influence the model and the predicted data.

Multiple linear regression models could be fitted to the energy consumption data and accounted for around 89%, 81%, 79%, and 76% of the variance in the sites (a to D) studied, respectively (Figure 3). The final RMSEs were calculated as 2751.6, 9508.2, 1904.6, and 1411.8 kW.h for Plamerston North, Christchurch, Dunedin, and Hamilton research office buildings, respectively (Table 1).

Table 1. Model equations and MSE and RMSE (kW.h)

	Equation	MSE	RMSE
Palmerston North	87561.6+ (31.68* Tem) - (3405.34*FTE)	7571201.0	2751.6
Christchurch	262840.9- (5830.67*Tem) - (337.17*FTE)	90405451.5	9508.2
Dunedin	31043.8- (1116.2*Tem) - (758.86*FTE)	3627612.3	1904.6
Hamilton	55496.3 – (953.168*Tem) - (496.2*FTE)	1993100.6	1411.8

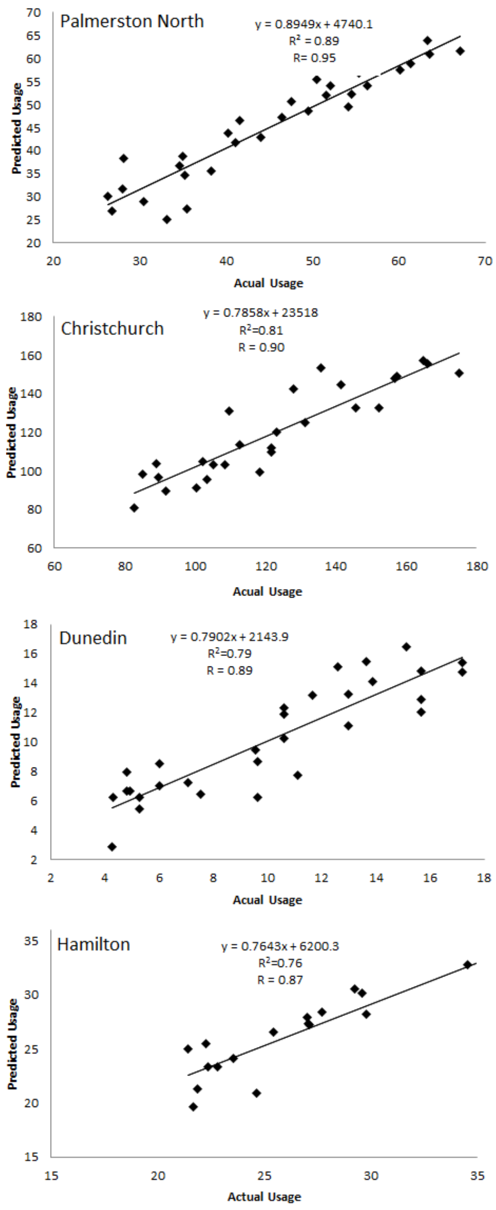

Figure 3. Correlation between the actual and predicted energy usage (1000 kW. h) in studies research offices

The differences between the models showed it would be too difficult to develop a model estimate even in similar buildings. However, the development of the models was simple and accurately compared actual and predicted energy usage and estimate energy auditing in buildings. The final linear models were very simple (two variable regression model) compared with complex nonlinear modelling methods and can even be used in MS Excel.

This modelling method can be used by building owners, property managers and consultants to monitor and manage energy usage in existing buildings. The hypothesis of this study can be used in other similar projects. Based on the introduced models, several models have been developed in different commercial buildings. For example, similar modelling methods have been used to forecast energy usage in a number of swimming pools and libraries using outside temperature and numbers of visitors.

4. Conclusion

This study presents a modelling method to predict energy usage to estimate energy savings in energy auditing projects in existing office buildings, when investigated separately. The results showed that the linear model with simple independent variables can predict energy usage within acceptable errors. The main challenge of this method was finding accurate data in an acceptable period of time.

References

[1] Azhar S., Brown J., and Sattineni A. A Case study of Building performance analyses using building information modelling. In. *27th International Symposium on Automation and Robotics in Construction (ISARC 2010)*, 213-222, Bratislava, Slovakia, 2010.

[2] Juan Y.-K., Gao P., and Wang J. A hybrid decision support system for sustainable office building renovation and energy performance improvement. *Energy and Buildings*, 42(3): 290-297, 2010.

[3] Brandon G. and Lewis A. Reducing Household Energy Consumption: A Qualitative and Quantative Field Study. *Journal of Environmental Psychology*, 19(1): 75-85, 1999.

[4] Abrahamse W., et al. A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3): 273-291, 2005.

[5] Pérez-Lombard L., Ortiz J., and Pout C. A review on buildings energy consumption information. *Energy and Buildings*, 40(3): 394-398, 2008.

[6] Isaacs N., et al., *Building Energy End-Use Study (BEES) Year 3*. 2010, BRANZ.

[7] Kavousian A., Rajagopal R., and Fischer M. Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy*, 55(0): 184-194, 2013.

[8] Korolija I., et al. UK office buildings archetypal model as methodological

approach in development of regression models for predicting building energy consumption from heating and cooling demands. *Energy and Buildings*, 60(0): 152-162, 2013.

[9] Gugliermetti F., Passerini G., and Bisegna F. Climate models for the assessment of office buildings energy performance. *Building and Environment*, 39(1): 39-50, 2004.

[10] Mathews E.H., et al. HVAC control strategies to enhance comfort and minimise energy usage. *Energy and Buildings*, 33(8): 853-863, 2001.

[11] Rubinstein F., Neils D., and Colak N. Daylighting, Dimming, and the Electricity Crisis in California. In *Invited paper for the 2001 IESNA National Conference*, Ottawa, Canada, 2001.

[12] Wang W., Zmeureanu R., and Rivard H. Applying multi-objective genetic algorithms in green building design optimization. *Building and Environment*, 40(11): 1512-1525, 2005.

[13] Roche P. and Milne M. Effects of combining smart shading and ventilation on thermal comfort. In *Solar Energy National Conference*, Orland , Florida, 2005.

[14] Mukherjee S., et al. Closed Loop Integrated Lighting and Daylighting Control for Low Energy Buildings. In *ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, 2010.

[15] Magnier L. and Haghighat F. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. *Building and Environment*, 45(3): 739-746, 2010.

[16] Vakiloroaya V., Su S.W., and Ha Q.P. HVAC integrated control for energy saving and comfort enhancement. In *Int. Symp. Automation and Robotics in Construction,*, 245-250, Seoul, Korea, 2011.

[17] Pandharipande A. and Caicedo D. Daylight integrated illumination control of LED systems based on enhanced presence sensing. *Energy and Buildings*, 43(4): 944-950, 2011.

[18] Üçtuğ F.G. and Yükseltan E. A linear programming approach to household energy conservation: Efficient allocation of budget. *Energy and Buildings*, 49(0): 200-208, 2012.

[19] Korolija I., et al. Regression models for predicting UK office building energy consumption from heating and cooling demands. *Energy and Buildings*, 59(0): 214-227, 2013.

[20] Athienitis A., Kalogirou S.A., and Candanedo L., *3.11 - Modeling and Simulation of Passive and Active Solar Thermal Systems*, in *Comprehensive Renewable Energy*, Sayigh A., Editor. 2012, Elsevier: Oxford. p. 357-417.

[21] Catalina T., Iordache V., and Caracaleanu B. Multiple regression model for fast prediction of the heating energy demand. *Energy and Buildings*, 57(0): 302-312, 2013.

[22] Catalina T., Virgone J., and Blanco E. Development and validation of regression models to predict monthly heating demand for residential buildings. *Energy and Buildings*, 40(10): 1825-1832, 2008.

[23] Lam J.C., et al. Multiple regression models for energy use in air-conditioned office buildings in different climates. *Energy Conversion and Management*, 51(12): 2692-2697, 2010.

[24] Neto A.H. and Fiorelli F.A.S. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and Buildings*, 40(12): 2169-2176, 2008.

[25] Suganthi L. and Samuel A.A. Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, 16(2): 1223-1240, 2012.

[26] Safa M. and Samarasinghe S. Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand". *Energy*, 36(8): 5140-5147, 2011.

[27] Alvarez R. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *European Journal of Agronomy*, 30(2): 70-77, 2009.

[28] Samarasinghe S., *Neural networks for applied sciences and engineering : from fundamentals to complex pattern recognition* Boca Raton, FL: Auerbach, 2007