OPEN ACCESS

2nd World Sustainability Forum
1-30 November 2012

www. wsforum.org

Article, Review, Communication

An Improved Method for Predicting Energy in Variable Occupancy Academic Buildings

Kevin P. Hallinan [1,*], Yosef Tesfay [1], Jesse Monn [1], Emily Krehnovi [1], Phil Brodrick[1], and Alvin Rivera Lopez [1]

[1] University of Dayton, Renewable and Clean Energy Program, Dayton, Oh, USA 45469-0238

E-Mails: kevin.hallinan@udayton.edu, tesfayy1@udayton.edu, monnj1@udayton.edu, emily.krehnovi@gmail.com, riveralopeza1@udayton.edu

* Author to whom correspondence should be addressed; Tel.: +1-011-937-229-2875; Fax: +1-011-937-229-4766

Received: / Accepted: / Published:

Abstract: Statistical energy savings calculations are fundamentally rooted in how well energy data can be normalized against influencing factors. Attempts to predict monthly energy use in academic buildings based strictly on weather as a driver for energy fail because of variable monthly occupancy. A genetic based energy model is used to characterize monthly energy consumption in academic buildings or any other buildings with variable occupancy. Such a model is essential for both estimating savings when changes are made and for continuously commissioning the building. Monthly average outdoor air temperature is considered to reflect the weather driver on energy use. Monthly occupancy is based upon the the historical academic year calendar; occupancy is considered a linear function of number of normal academic days per month.. The multi-functional model developed is tested on both simulated and actual academic building energy data. The results demonstrate universally improved correlations.

Keywords: statistical; prediction; energy use, regression, occupancy, weather.

---

1. Introduction

Since 2007, 22 U.S. states have adopted energy efficiency resource standards, bringing the total number of states up to 26 as of 2011. Typical annual energy reduction requirements range from 0.5 to 2 percent per year [1]. Additionally, there has been substantial growth of Energy Service Companies (ESCOs), which are companies contracted to realize guaranteed savings from implemented energy reduction measures. The U.S. ESCO industry grew at about 7% per year between 2006 and 2008, and 26% between 2009 and 2011.This industry now has annual revenues in excess of $7.1B. Key to this growth has been an increased spending rate in taxpayer-funded energy efficiency programs [2]. ESCOs are faced with determining two different types of savings:

1. Predicted Savings
    a. - based on outside temperature and occupancy with the use of a genetic algorithm or other statistical methods to predict energy use.
    b. – determined by a variety of methods before any ECMs have been put into effect, and consequently based only on data obtained prior to the ECM.
2. Actual Savings - determined from a comparison of measurements before and after at least one Energy Conservation Measure (ECM) has been put into effect.

An international standard has been developed by the Efficiency Valuation Organization (EVO) for the measurement a verification of Actual Savings called the International Performance Measurement and Verification Protocol (IPMVP) [3]. Savings are defined by the IPMVP as:

$$Savings(Benchmark\ Period\ Energy - Reporting\ Period\ Energy) \pm Adjustments \quad (1)$$

Subject to the definitions:

Adjustments – A factor that corrects for differences in conditions between the Benchmark and Reporting periods due to independent variables.

Benchmark[1] Period – The period of time ranging from the start of the energy data to the time ECMs were implemented.

Reporting Period – The period of time ranging from the time the ECMs were implemented to the end of the energy data.

Key to measuring actual savings is the proper characterization of Adjustments; in fact "simple comparisons of utility costs without such Adjustments report only cost changes and fail to report the true performance" [3]. Independent variables responsible for Adjustments come in many forms, including but not limited to meteorological variation, production rates, and facility size. The IPMVP groups Adjustments into two categories:

Non-Routine Adjustments – Adjustments that handle independent variables that are expected to remain constant throughout the Reporting and Benchmark Periods. Common examples of these independent variables include design and operation of equipment, facility size, and facility use type.

Routine Adjustments – Adjustments that handle independent variables expected to change throughout the Reporting and Benchmark Periods. Common examples of these independent variables include meteorological data, and production rates.

Independent variables responsible for Non-Routine Adjustments, or "Static Variables", must be monitored for change throughout the Reporting Period, and Non-Routine Adjustments must be made in the event that any of these Static Variables change. While Non-Routine Adjustments may be difficult to calculate in the event that the Static Variables are poorly monitored or too numerous, it will be assumed for the purposes of this paper that all Non-Routine Adjustments can be effectively calculated.

Independent variables responsible for Routine Adjustments, or "Dynamic Variables", can be difficult to account for due to both a lack of clear definition of Dynamic Variables as well as a high degree of variation in these variables. In the case of buildings where energy devoted to heating or cooling is a dominant factor, outdoor air temperature and occupancy are the most dominant Dynamic Variables and their effects have traditionally been taken into account with a regression analysis (Rabl et al., 1988 [4], Rabl et al., 1992 [5] , Fels 1986 [6], Ruch 1993 [7]). The IPMVP claims that when energy measurements used for savings calculations are performed on an entire facility, Savings should exceed 10% of the Benchmark Period Energy if the Savings are to be distinguishable when the Reporting Period is less than two years [3]. This claim stems from the assumption that a substantial portion of the Benchmark Energy is unexplained by the Adjustments – thus, improving the accuracy of the Adjustments would allow for the accurate measurement of Savings even when Savings are under 10% of the Benchmark Energy.

In total, the techniques for predicting / forecasting energy data, e.g., for making Adjustments, can be divided up into 5 approaches [8]:

- Engineering Methods
- Statistical Approaches
- Artificial Neural Networks
- Support Vector Machines
- Grey Models

Energy Methods are based upon physical energy models of buildings of varying detail. From these models, the energy model is first calibrated using historic energy use, and once calibrated can be used to both predict future energy use and estimate savings from specific measures. As noted by Zhao and Magoules, however, these techniques often rely upon detailed information about a building, which may not be available. Without precise inputs, the estimates from these types of models can have high uncertainty [8].

Statistical regression methods aim to utilize only historical energy data and possibly other energy drivers such as occupancy and incident solar energy, in order to model past energy use and thus be poised to predict future energy use and predict savings. Correlating energy use with weather variation permits estimation of heating and cooling energy. The nature of the weather information is variable. White and Reichmuth attempted to use average monthly temperatures to predict monthly building energy consumption; an approach that is more accurate because it accounts for how an individual user heats and cools their building rather than standard procedures which normally use heating and cooling degree days or temperature bins, which employ assumed 65 deg. F balance point temperatures for heating an cooling [9]. Westphal and Lamberts correlated monthly energy use with monthly average of maximum and minimum temperatures, atmospheric pressure, cloud cover and relative humidity [10]. Kissock et al. correlated monthly energy use with only monthly dry bulb temperature to develop building heating and cooling slopes (energy use / time period / degree temperature change) along with balance point temperatures for heating and cooling [11]. These balance point temperatures could be used to estimate heating and cooling degree hours in a typical year using local typical weather year data available from the NOAA(National Climatic Data Center). Finally, in buildings dominated by internal loads where energy use is far less or negligibly dependent on weather variation, it is essential that statistical methods utilize occupancy variation as a driver for energy use. But, occupancy as an input can be problematic. This data is generally less available than building energy characteristics. A research vein has emerged to measure occupancy in a building real time. RFID technology has been employed to measure real time the entrance of an occupant into a building [12]. Newsham and Birt used simple occupancy sensors and logical inference of sensor input to estimate occupancy [13] An interesting study by Bellala et al. of academic buildings based occupancy models upon computer network port-level logs. This study showed a correlation between occupancy and network activity [14].

Zhao and Margoules describe artificial neural networks as being quite effective in both modeling past energy use and in predicting future energy use [8]. However, unlike the regression approaches, these approaches are not capable of illuminating the drivers for energy use. When multiple drivers exist (such as occupancy and weather), it is impossible to determine the fraction of energy owing specifically to the drivers. Grey models are particularly useful when there is only incomplete or uncertain data. As noted by Zhao and Margoules, there has been little application of this approach to buildings.

2. Goal

The goal of this paper is to develop a method for predicting future energy use in academic buildings, as well as to identify buildings for upgrade. This goal is only achievable if the variation in functioning of the building due to the the academic calendar can be accounted for.

3. Analysis

This study begins with a statistical analysis of energy use in the 14 primary academic buildings at the University of Dayton which have calendar year variation in use. Included in this mix of buildings are those dedicated to teaching/research, service, administration, and residential living. The initial statistical analysis seeks to correlate electrical energy use in these buildings with monthly average outdoor temperature (using the approach of White and Reichmuth [9]), of the form:

$$\text{Monthly Electric Use} = \text{Baseline} + \text{Cooling Slope} * H(T - T_{balC}) * (T - T_{balc})$$
$$+ HS*H(T_{balc} - T_{balc1})^{*}(T_{balc} - T_{balc1})$$

However, as shown in Table 1, which summarizes the correlation r-squared value for each building, this correlation yields generally poor results, with r-squared values raning from 0.08 to 0.69.

Table 1. r-squared values from standard statistical analysis

Bldgs	1	2	3	4	5	6	7	8	9	10	11	12	13	14
r-squared	0.14	0.41	0.11	0.23	0.69	0.17	0.1	0.09	0.08	0.29	0.19	0.29	0.175	0.093

These buildings are cooling load dominated; this explanation partially explains the poor correlations since the cooling often compensates for human loads. However, there certainly should be increased cooling energy use in the summer. Figure 1, which shows the electrical energy use in a representative academic building – where there is significant summer variation in use, seems to sometimes contradict this expectation. This figure highlights in yellow the summer months. While often over the time period from 2002-2009 the summer energy data exceeds the mean energy data given by the trendline, this isn't always the case. Moreover, some of the peak energy months aren't in the summer.

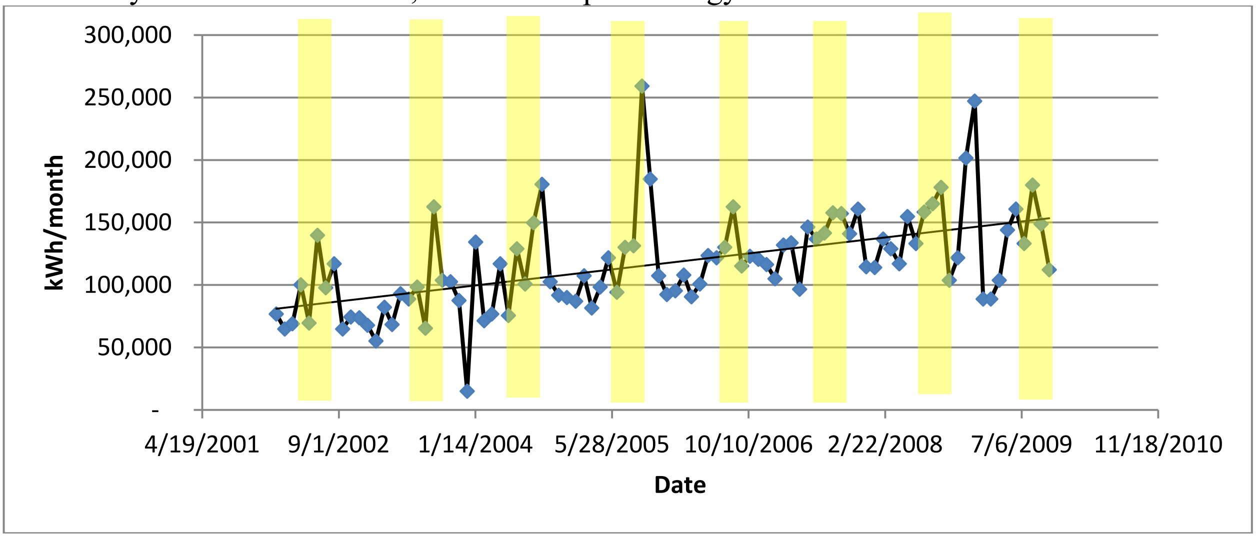

Figure 1.Monthly electrical energy (kWh) vs Date

Figure 2 shows the monthly electrical energy use as a function of mean monthly temperature during the meter period. The summer data points are indicated by square red blocks, while the energy data for the rest of the year are shown as blue diamonds. There are some obvious points to illuminate. First, there is in general only a weak correlation of energy use with temperature. Secondly, the summer months show only a nominally improved dependence upon temperature. Thus, there must be other influences on energy use which aren't considered in the statistical analysis given by equation (1).

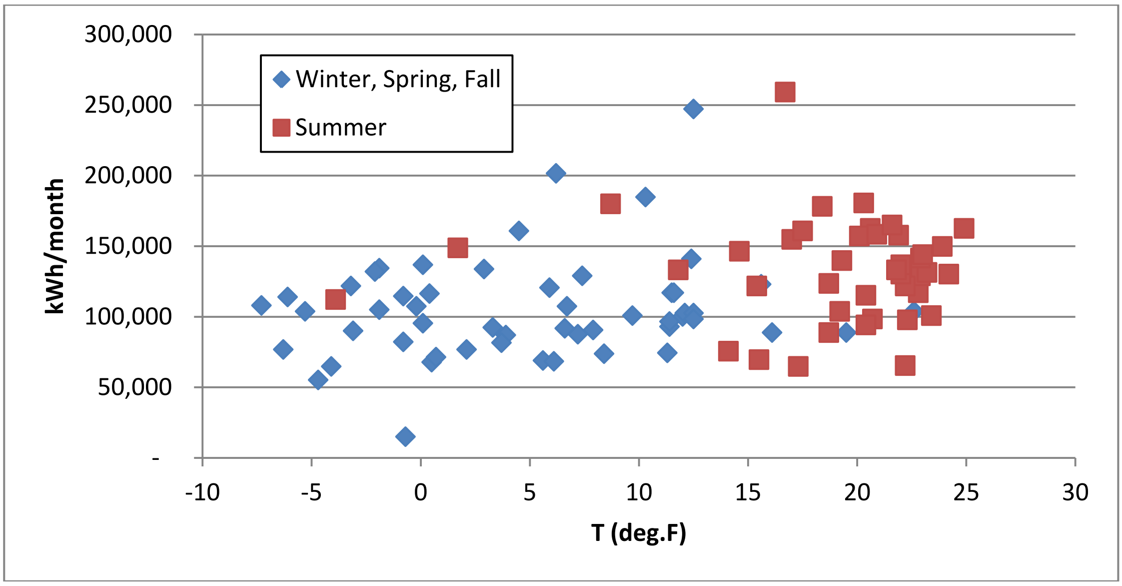

Figure 2. Monthly electric energy use as a function of mean outdoor temperature during the billing period.

With a goal to be able to predict monthly energy use in order to evaluate the possibility of problems in the building, it is clear from Figure 1 and 2 that the varied use of the academic building during summer months must somehow be accounted for.

The building characterized in Figures 1 and 2 has nearly all classrooms fully occupied during the academic year (Aug. 22 – May 5), and is only fractionally occupied during the summer and during a winter break. At the same time, faculty/staff offices see a slight summer decrease in occupancy, either because of summer study abroad participation, off-site research, or simply because of vacations taken. The key to improving predictive capability for the energy use is to somehow account for the variation in functioning during the summer and holidays. Ideally, this variation could be correlated to occupancy. Occupancy sensors ideally could be used for measurement. However, in this application such data is unavailable. Melfi et al. estimated occupancy via building wide network utilization [15]. This approach, while clever and novel, could not be employed in this study.

Another approach is used to account for functional variation. Given the fact that the summer 'bump' in energy use due to an expectedly increase in cooling load is not realized almost certainly because the internal human loads (and even computer loads) decrease during the summer, weekends, and holidays, a modified statistical model representing energy use dependence on weather and occupancy is posed. This approach assumes a linear occupancy functionality to the number of normal class days per month. For example, during the nomral academic year in April 2009, there were 16 class days because of the Easter holiday and there were 20 days in which the residences were occupied. Occupancy for academic buildings was defined as the ratio of class days in a month to total days in the month. For residential buildings, occupancy is defined as the ratio of days in which the residence hall was occupied in a month to the total number of days in the month. In the summer months, the occupancy value for all academic and residential buildings was set to 0.

In order to determine prior occupancy in each building, the historical academic calendar for the University of Dayton was used. Table 2 presents the 'occupancy' measure for the various types of academic buildings over the period from Jan. 15, 2002 to Sept. 15, 2009.

Table 2. Academic building type monthly high occupancy data obtained from correlation with the historical academic year calendar

Date	Days in Month Total	Residence Hall: Days Occupied in Month	Classrooms: Days Occupied in Month	Residence Hall: Percentage of Month Occupied	Classrooms: Percentage of Month Occupied
1/15/2002	31	27	19	0.87	0.61
2/15/2002	28	23	17	0.82	0.61
3/15/2002	31	23	17	0.74	0.55
4/15/2002	30	28	20	0.93	0.67
5/15/2002	31	5	2	0.16	0.06
6/15/2002	30	0	0	0.00	0.00
7/15/2002	31	0	0	0.00	0.00
8/15/2002	31	8	5	0.26	0.16
9/15/2002	30	29	20	0.97	0.67
10/15/2002	31	31	23	1.00	0.74
11/15/2002	30	25	17	0.83	0.57
12/15/2002	31	12	9	0.39	0.29
,,,	,,,	,,,	,,,	,,,	,,,
,,,	,,,	,,,	,,,	,,,	,,,
,,,	,,,	,,,	,,,	,,,	,,,
10/15/2008	31	26	21	0.84	0.68
11/15/2008	30	25	16	0.83	0.53
12/15/2008	31	19	14	0.61	0.45
1/15/2009	31	26	19	0.84	0.61
2/15/2009	28	26	19	0.93	0.68
3/15/2009	31	29	22	0.94	0.71
4/15/2009	30	20	16	0.67	0.53
5/15/2009	31	3	1	0.10	0.03
6/15/2009	30	0	0	0.00	0.00
7/15/2009	31	0	0	0.00	0.00
8/15/2009	31	9	4	0.29	0.13
9/15/2009	30	29	21	0.97	0.70

In this context, an objective function is hypothesized to estimate past energy use. This objective function, shown below, considers as influences:

- Base loads (lighting / appliances / computers);
- Weather (dry-bulb temperature); and
- Occupancy

Predicted Monthly Electric Use$_i$ = Baseline + Cooling Slope * Heaviside(T- T_{balc}) * (T- T_{balc})^{CSE}

$$+ \text{Occupancy Slope} * \text{Occ} \quad (2)$$

In this equation, Baseline refers to energy use which is not dependent upon weather or occupancy; thus it represents the monthly lighting/appliance/computing energy load during the lower occupied periods. The Cooling Slope (kWh/month/deg.F) and Occupancy Slope are the sensitivities of energy use to respectively mean monthly average temperature, T, and Occ (occupancy). T_{balc} is the cooling balance point temperature; e.g., the outdoor dry-bulb temperature above which cooling occurs. The monthly outdoor temperatures for Dayton, OH, USA over the study period were obtained from the NOAA [16]. The exponent CSE enables a step change in predicted monthly energy use for temperatures above T_{balc}, while permitting an increase of cooling energy for temperatures exceeding this temperature. This situation is common in some of the academic buildings, where the chiller(s) may be turned off during winter As noted previously, the Occ variable is normalized to total days in a month, such that it represents the % of days in a month that class is in session (e.g., the last two columns in Table 2).

A genetic algorithm optimization approach was used to find the optimal values of the presumed independent factors (Baseline, Cooling Slope, T_{balc}, CSE, and Occupancy Slope). The optimization process attempts to maximize the r-squared, where r-squared is the goodness measure of the fit between the Predicted Monthly Electric Use and the billed electric energy use, Actual Monthly Electric Use (kWh). This correlation was based only on the last two years of billed data. The objective function is shown below in equation (3).

$$f = r\text{-}squared \quad (3)$$

where

$$r\text{-}squared = \frac{\sum_{n=1}^{N}(predicted\ monthly\ energy_n - actual\ monthly\ energy_n)^2}{\sum_{n=1}^{N}(actual\ monthly\ energy_n - mean(actual\ emonthly\ energy_n))^2} \quad (4)$$

This optimization required the following non-linear constraint, which basically says that the sum of the monthly predicted energy over the study period (N months) must be equal to the sum of the actual energy use over the same period. Thus, the annual predicted energy use must be equal to the annual actual energy use.

$$\text{Total predicted energy} = \sum_{n=1}^{N} \text{Predicted monthly energy}_n = \sum_{n=1}^{N} \text{Actual monthly energy}_n \quad (5)$$

4. Results

The optimization given by equations (3) -(5) were applied to all of the academic buildings for which the functionality varied throughout the year. The summary results are shown in Table 3. Included in this table are the building names, the r-squared value of the standard fit (e.g., the fit obtained without consideration of occupancy influences), and the r-squared value of the improved fit which included occupancy data. It is apparent that in all cases but 1 there is significant improvement. Interestingly, the one building which didn’t realize improvement in the r-squared value, was the engineering building (Kettering). This building includes many labs, all of which operate year round. The energy change in the summer is certainly less; thus the occupancy dependence was small.

Table 3. Summary of r-squared goodness measures of standard and improved fits for all buildings in the study

Building #	r-squared, original	r-squared,improved
Anderson	0.138	0.56
Humanities	0.41	0.48
Kettering	0.110	0.114
Kennedy Union	0.230	0.37
Liberty	0.69	--
Marianist	0.169	0.357
Marycrest	0.100	0.221
Miriam	0.09	0.183
Recplex	0.083	0.307
Roesch	0.293	0.331
Sherman	0,19	0.44
Stuart	0.295	0.410
VWK	0.175	0.389
Wohleben	0.093	0.397

Figures 3 and 4 offer representative comparisons of actual energy use and predicted energy use as a function of month by the standard and improved fits for two of the buildings. Both figures highlight the summer months (yellow box), where the occupancy level is at a 0 level. A clear reason for the improvement in the fit is that energy use in lower occupancy times could be identified. Both figures show the predicted energy use during these periods beneath that predicted from the standard fit.

It is also clear that the regression fits still have room for improvement. One reason is that the energy data maintained by the university didn't include the actual billing date. Thus, the calculated monthly average temperatures can be off by as much as 2 degrees F. Secondly, what cannot be included in the statistical analysis are changes in the building made by facilities personnel in order to keep the building HVAC operational. There are significant building fixes which routinely cause the energy use in all buildings to fluctuate significantly.

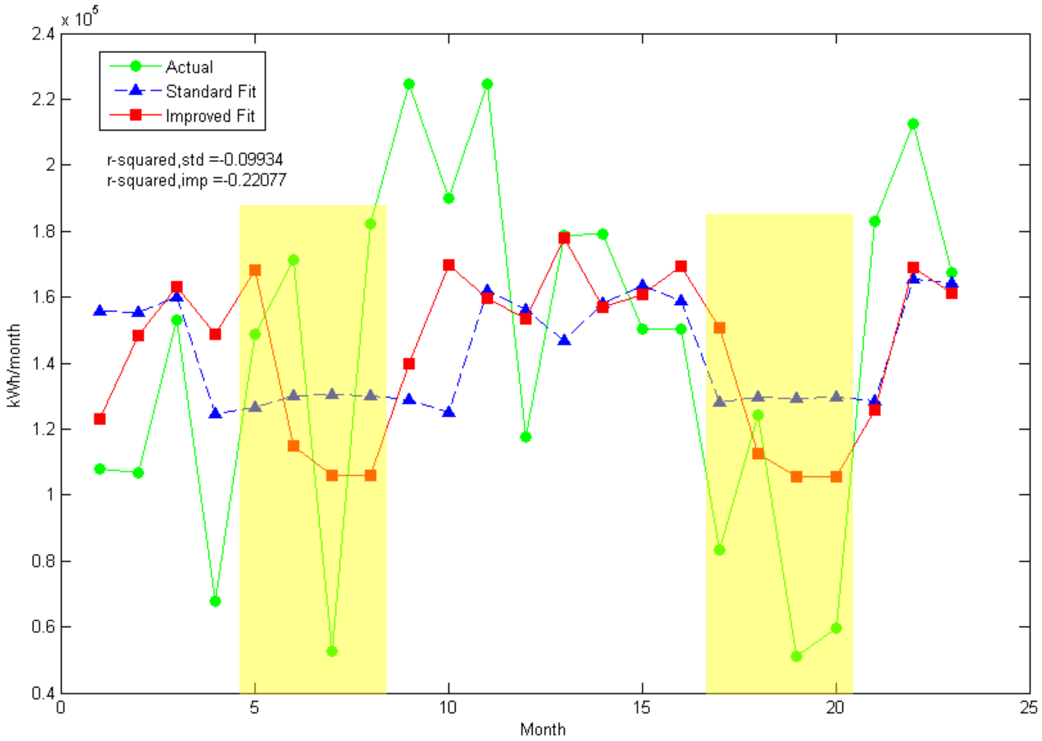

Figure 3. Actual monthly versus predicted monthly energy use for month for standard and occupancy influenced fit (Marycrest).

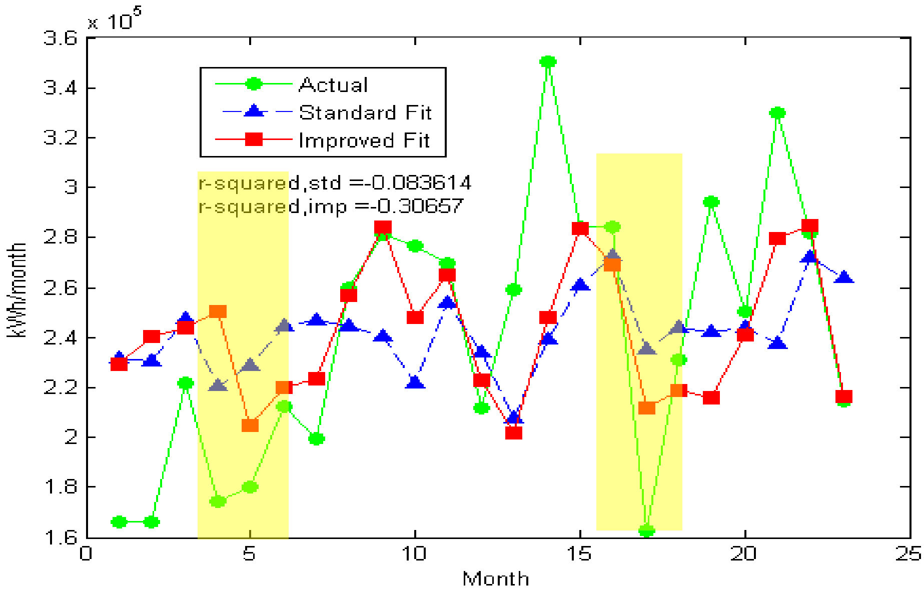

Figure 4. Actual monthly versus predicted monthly energy use for month for standard and occupancy influenced fit (Recplex).

4. Conclusions

A new approach for predicting monthly electrical energy use in cooling load dominated academic buildings has been posited. This approach models occupancy as a binary function, considering high occupancy during the normal academic year and low occupancy characterizing building use during the summer, weekends, and holidays. The occupancy level is attainable from historical academic year calendars. Results have shown that inclusion of the occupancy level in the predictive equation for energy use improves significantly the regression goodness, as evidenced by r-squared values which often are two times higher.

While the regressions developed still do not explain all variation with the actual data, they offer at least an improved means to estimate monthly energy use. Such estimates can be compared to actual use as part of a continuous commissioning process. When actual energy use in any month is well higher than expected energy use, facilities personnel can be directed to investigate the cause.

Conflict of Interest

State any potential conflicts of interest here or "The authors declare no conflict of interest".
Main text paragraph (M_Text).

References and Notes

[1] J. White and R. and Reichmuth, "Simplified method for predicting building energy consumption using average monthly temperatures," in *31st Intersocieity of Energy Conversion Engineering Conference*, 1996.

[2] R. Melfi, B. Rosenblum, B. Nordman and K. and Chrstensen, "Measuring building occupancy using existing network infrastructure," in *International Green Computing Conference*, 2011.

[3] NOAAA, "On-Line Climate Directory," 2012. [Online]. Available: http://lwf.ncdc.noaa.gov/oa/climate/climatedata.html. [Accessed 12 10 2012].

[4] H. Allcott, "Social Norms and Energy Conservation," Massachuesetts Institute of Technology, Cambridge, MA, 2009.

[5] A. Satchwell, C. Goldman, P. Larsen, D. Gilligan and T. Singer, "A Survey of the U.S. ESCO Industry: Market Growth and Development from 2008-2011," Lawrence Berkeley National Laboratory, 2010.

[6] Efficiency Valuation Organization (EVO), "International Performance Measurement and Verification Protocol (IPMVP)," 2010.

[7] A. Rabl, "Parameter Estimation in Buildings: Methods for Dynamic Analysis of Measured Energy Use," *ASME Journal of Solar Energy Engineering,* vol. 110, pp. 52 - 62, 1988.

[8] A. Rabl, L. Norford and J. and Spadaro, "Steady State Models for Analysis of Commercial Building Energy Data," in *ACEEE Summer Study on Energy Efficiency in Buildings,* Pacific Grove, CA, 1992.

[9] M. Fels, "Special Issue Devoted to Measuring Energy Savings: The Scorekeeping Approach," *Energy and Buildings,* vol. 9, no. 1 & 2, pp. 5-18, 1986a.

[10] D. Ruch, J. Chen, J. Haberl and D. and Claridge, "A changepoint principal component analysis (CP/PCA) method for predicting energy usage in commercial buildings: The PCA model," *ASME Journal of Solar Energy Engineering,* vol. 115, no. 2, pp. 77-84, 1993.

[11] H. Zhao and F. Magoulies, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews,* vol. 16, pp. 3586-3592, 2012.

[12] F. Westphal and R. and Lamberts, "The use of simplified weather data to estimate thermal loads of non-residential buildings," *Energy and Buildings,* vol. 36, no. 8, pp. 847-854, 2004.

[13] J. Kissock and S. and Mulqueen, "Targeting Energy Efficiency in Commercial Buildings Using Advanced Billing Analysis," in *ACEEE Summer Conference*, 2008.

[14] S. Li, N. Li, B. Becerik-Gerber and C. Gulben, "RFID-Based Occupancy Detection Solution for Optimizing Hvac Energy Consumption," in *28th International Association for Automation and Robotics in Construction*, Seoul, Korea, 2011.

[15] G. Newsham and B. and Birt, "Building-level Occupancy Data to Improve ARIMA-based Electricity Use Forecasts," in *2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, BuildSyst '10*, 2010.

[16] G. Bellala, M. Marwah, M. Arlitt, G. Lyon and C. and Bash, "Towards and Understanding of Campus-Scale Power Consumption," in *BuildSys' 11*, Seattle, WA, USA, 2011.