The Phantom Control Group:

Predicting Electricity Usage in University Residential Colleges

Marthinus J. Bekker and Louis S. Leland, Jr.

Department of Psychology

University of Otago

Dunedin, New Zealand

During 2008 the people of New Zealand were once again reminded that electricity supply is not always guaranteed, with headlines reporting a 'looming power crisis' when lake levels got down to just 53% of the average expected for the time of year, which was the lowest since the 1992 power crisis (Coleman, 2008). These situations, combined with a global effort to reduce the world's reliance on non-renewable resources such as fossil fuels, make the need for electricity conservation obvious. One area that has been targeted to conserve electricity is collective, master metered, accommodation such as University Residential Colleges (Bekker, et al., 2007; Petersen, Shunturov, Janda, Platt, & Weinberger, 2007). Residents of collective accommodation often use electricity with no regard for conservation at all, as they are not responsible for the bill and do not receive feedback as to how much they are using or saving. Bekker et al.'s (2007) study set out to introduce a combination of feedback, incentive and education to the residents of a University College of residence. The results indicated a 12.60% decrease during the day and a 4.17% decrease during the night as compared to a baseline and controlling for environmental factors by using a control college. The use of the control college was an essential part of this experiment because, in order to calculate savings for feedback and reward levels, factors that contribute to electricity usage such as temperature, humidity, and luminance need to be controlled for. This is not, however, always practical, as a comparable control residence might not always be available and daily electricity recording at both colleges may not be practical. Also, if all residential colleges in an area such as Dunedin were to implement such an effective intervention, at least one would need to remain un-intervened to act as a control. One theoretical solution would be a regression equation predicting usage from weather and other relevant variables that could be used to compare actual usage to predicted usage. This would enable colleges to calculate savings attributable to the intervention and give feedback to the residents regarding their savings (Bekker et al., 2007). The current project proposed to gather past daily electricity usage data for several residential colleges in Dunedin, available through the University's Energy Manager, and to then integrate this with other relevant variables for the same periods in order to produce a regression equation that could, with some accuracy, predict expected usage from predictive variables. Predictive variables included previous usage patterns, temperature, humidity, luminance, and time variables (e.g. time of day, day of the week, and day of the year). Applications of such a predictive equation would potentially range from Residential Colleges to many other types of residence including Hotels and private residences.

Method

Settings

The settings of this study consisted of some of the 14 tertiary Colleges of Residence within the Dunedin area. Selection of the colleges to be included in the current study depended primarily on the availability of detailed historic electricity data. Five colleges had sufficient data to be included in the analysis; they consisted of Aquinas College, Selwyn College, University College, City College, and Cumberland College.

Aquinas College is located approximately 20 minutes walk from campus and accommodates 152 residents, most in single rooms. It is University owned and uses electricity for heating and hot water. The college is purpose built and was established in 1954.

Selwyn College is located on campus and accommodates 170 residents in single study bedrooms. It is privately owned and uses electricity for heating and hot water. Selwyn was Dunedin's first purpose built Residential College and was established in 1893, with much of the accommodation still in the historic building.

University College (Unicol) is located on campus and was purpose built by the university. It was first opened in 1969 to accommodate just over 400 residents; an additional 107 rooms were opened in 2004 which has increased the college's capacity to 518 residents, mostly in single study bedrooms. The college has its own coal burner from which the heating and hot water is run.

City College (Citycol) is the newest of the residential colleges and was officially opened in 2000. It is privately owned and located five minutes walk from campus. It accommodates 211 residents in 37, 5 and 6 bedroom apartments. All heating and hot water uses electricity and is run through an electronic building management system.

Cumberland College is located five minutes walk from campus in a building first opened in 1916 as the Dunedin hospital's nurses home. It was first used as a Residential college in 1989 and now accommodates 326 residents, in single study bedrooms. All heating and hot water is supplied from a nearby steam plant which also provides the hospital and other buildings with steam.

Apparatus

The main apparatus used in this project has been version 15 of the Statistical Package for Social Sciences (SPSS) and Microsoft Excel 2003. These programs were used to collate and analyse the data. All the weather data was obtained from is the University of Otago, Energy Studies Program weather station. Located at latitude 45º 52'south and longitude 170º 31' east longitude, on the top of the University of Otago's Physics building approximately 45m above mean sea level (University of Otago, Energy Studies Program, n.d., a). Another essential source of data in this study was the energy management system employed by the University of Otago's property management unit, who supplied all of the electricity data used in the study. The electricity meters at all of these Colleges of Residence are operated by a company called *Stream* based in Auckland. Meters are accessed through cell phone connections and readings are then passed on to retailers and customers *(Hans Pietsch, Energy Manager, University of Otago Property Services, Personal Communication, Jan. 26, 2009).*

Data Gathering and Transformation

Electricity data was obtained through the University's Energy manager for half hour intervals from June of 2004 till October of 2008 for nine Colleges of Residence. Since the majority of the University runs on a time of use contract, which has different pricing schedules across four hourly intervals starting from midnight, it was decided to transform the data into four hourly time intervals *(Hans Pietsch, Energy Manager, University of Otago Property Services, Personal Communication, Nov. 6, 2008).* The times of interest for this study were the periods during which the majority of student residents would be living at the college. Therefore all holiday periods, including summer break, mid year break, and both mid semester breaks were removed from the data starting from the last weekday before the break till the first weekday after the break according to University of Otago's (2004-2008) calendars. Exam periods starting from the last day of lectures were also removed due to the varying departure dates of students during these times. Every electricity data point was then matched to the equivalent day in the year before, for data starting from June of 2005 till October of 2008 for each of the respective Colleges of Residence.

Weather data were obtained through the University of Otago's, Energy Studies Program (n.d., b) weather station archive. Five minute interval data was obtained for temperature, relative humidity, wind speed, wind direction, solar radiation, UVA, UVB, visible micro molecules per meter squared, rain, atmospheric pressure, and the maximum wind gust, for the periods corresponding to the transformed electricity data. In order to convert the weather variables to four hourly figures, variables were either averaged, summed or the maximum across the four hour period was taken. Wind direction was first converted to radians from degrees in order to calculate an average, since the difference between 1° and 360° is only 1° rather than 359°.

The third set of variables were time related including the date and time in seconds, day of the year, hour of the day, day of the week, and the year. Since many of these variables do not have linear relationships with electricity use they were converted into suitable forms. Two variables, 'hour of the day' and 'day of the week' were converted into dummy variables where each time period had its own variable which reads "0" every time except when the specific variable is true in which case it reads "1". For instance 'hour of the day 8' would read "0" for every line of data except for the lines that correspond to the 8am time slot. Day of the year and year were converted to a variable called day multiplied by year squared in order to make what is essentially a parabolic relationship (an increase in electricity use up to the middle of winter and then a decrease) into a more linear one.

These variables were then combined into a single data base, using Microsoft Excel 2003.

Data Analysis

The final Excel file containing all of the variables was then imported into SPSS for analysis. The data for each college was then entered into separate linear regression analyses. The order of variables depended on the contribution of each, as determined by an initial analysis. Previous usage was first entered then the hour of day and day of the week dummy variables, followed by the year, day of the year, and day times year squared in one block. Following this, the rest of the variables were entered in a stepwise manner with an entry criteria of $F<=0.05$ and a remove criteria of $F>=0.10$. The variables for each of the colleges then produced a range of equations for each college. From these equations, the one with the highest predictive value calling for the least variables was used. Equations were then transcribed into a form suitable for Excel and applied to all of the colleges of residence to produce predicted values from each of the equations for all the colleges. From the predicted values, variance (r^2), and standard errors of estimate were calculated for all possible combinations of equations and colleges. In order to then compare standard errors of estimate across colleges with different absolute electricity usage, the standard errors of estimate were converted to a percentage of the average for each respective college.

Results

The results for this project are displayed as predictive values (r^2= amount of variance explained) and as standard errors of estimate, expressed as percentages of the respective college's average usage, in order to make comparisons between colleges. Predictive values are expressed as decimals between zero and one, with zero describing a situation where none of the variance is explained and therefore there is no predictive value, and one is describing a situation where all of the variance is explained and there are very precise predictions. Not surprisingly, the best predictive value for each college was obtained for the equations which had been derived from its own historic data. These predictive values ranged between 0.816 and 0.930 which means that a very large portion of the variance could be accounted for in each of the colleges. No one equation produced excellent predictions for all the colleges, however, the best out of the various equations appeared to be the equation derived from Selwyn's data which produced predictive values between the first and third best for all of the colleges (Figure 1.). Also it is worth noting that both steam heated colleges, Unicol and Cumberland, not only had excellent predictive values for their own equations but also for each other's equations.

Electricity usage for a four hour period =(-21871.254)+(0.251*Previous usage)+(-118.046*Hour of the day Dummy 4)+(-121.155*Hour of the day Dummy 8) + (95.972*Hour of the day Dummy 12)+(100.549*Hour of the day Dummy 16)+(55.956*Hour of the day Dummy 20)+(11.014*Year)+(-0.006*Day multiplied by year squared)+(2.247*Day of the year)+(-12.813*Temperature)+(3.630*UVA)+(42.880*Day of the Week 2)+(41.297*Day of the Week 3)+(42.531*Day of the Week 4) +(40.619*Day of the Week 5)+(34.822*Day of the Week 6)+(9.773*Day of the Week 7)+(-0.765*Relative Humidity)+(-0.117*Visible micromols per m^2)

Figure 1. Multiple regression equation derived from Selwyn College of Residence

Table 1. *Predictive values for the equations derived from five Colleges applied to each other*

	Predictive (r^2) values for equations derived from:				
	Aquinas	Selwyn	Unicol	Citycol	Cumberland
Aquinas	**0.836**	0.816	0.611	0.776	0.606
Selwyn	0.796	**0.816**	0.629	0.706	0.623
Unicol	0.706	0.779	**0.930**	0.498	0.923
Citycol	0.808	0.772	0.584	**0.859**	0.583
Cumberland	0.614	0.682	0.903	0.416	**0.924**

The second measure used in the study was the standard error of estimate which is derived from the square root of the average squared error (Figure 2.).

$$\sigma_{est} = \sqrt{\frac{\sum (Y-Y')^2}{N}}$$

Figure 2. Standard error of estimate equation.

This is a measure of the accuracy with which the equation can predict a value. However since this is in the units of measurement in which the prediction is made (in this case kWh) it can not be compared across different colleges who use different absolute amounts of electricity. For this reason the standard error of estimate was transformed into a percentage of the average for each respective hall in order to make comparisons. The best error of estimate for each college ranged between 7.88% of the average usage for Unicol to 17.04% of the average usage for Selwyn. This means that, on average, a single prediction is likely to be out by 7.88% to 17.04% depending on the college. Once again the lowest standard error of estimate for each college was derived from the equation computed from its own historic data.

Table 2. *Standard errors of estimate for equations derived from five Colleges applied to each other expressed as percentages of the respective College's averages*

	Standard errors of estimate as percentages of the average usage for equations derived from:				
	Aquinas	Selwyn	Unicol	Citycol	Cumberland
Aquinas	**13.89 %**	14.78 %	22.92 %	32.68 %	21.53 %
Selwyn	17.94 %	**17.04 %**	24.85 %	36.90 %	24.69 %
Unicol	16.00 %	13.95 %	**7.88 %**	29.50 %	10.84 %
Citycol	19.69 %	19.91 %	23.36 %	**13.34 %**	24.87 %
Cumberland	24.07 %	21.97 %	16.08 %	56.35 %	**9.28 %**

Discussion

The results of this study indicate that with the equations derived, errors are likely to average between 7.88% and 17.04% of the respective college's usage, either side of the true value, with some outliers being well in excess of these standard error terms. So even though excellent predictive values, as high as 0.93, were achieved the idea does not quite seem plausible due to the magnitude of the likely savings. In the intervention study by Bekker et al. (2007) the results indicated a 12.60% decrease during the day and a 4.17% decrease during the night as compared to a baseline and controlling for environmental factors by using a control college. This essentially means that, by using these equations, a day on which an attributable saving was made could show no saving or a vastly exaggerated saving, with the same being true for losses. This could be acceptable if it was done across a longer period of time, since an extended period of time should have over and under exaggerations which, in theory, would even out to produce a valid estimate of the average savings across a period of time. However, this study was hoping to achieve day to day accuracy so that feedback could be given to residents of the colleges as a part of the intervention. Therefore, despite great predictive values the current equations do not seem plausible for the proposed function.

Another interesting finding of the current study was that the two steam heated Colleges both produced highly predictive equations for themselves and for each other. This could be due to not having to account for any variation in the way heating or hot water is used. This is further supported by the fact that their equations also predict each other's electricity usage much better than the electrically heated Colleges, suggesting that there are similar factors that drive these colleges' electricity usage. The opposite also holds true for most of the equations derived from electrically heated Colleges, these equations predicted their own and other electrically heated College's electricity usage better than the usage of the steam heated Colleges. This finding would suggest that any future research into predictive equations for large scale accommodation should develop separate equations for electrically heated and steam heated residences.

References

Bekker, M. J., Cumming, T. D., Leland, L.S., Jr., Osborne, N. K. P., Bruining, A. M. & McClean, J. I. (2008, August). *Encouraging electricity savings in a University hall of residence through a combination of feedback, visual prompts, and incentives.* In N. Blampied (Chair), *Environmental Psychology*. Symposium conducted at The Annual Conference of The New Zealand Psychological Society, Christchurch, New Zealand.

Coleman, J. (2008, June 5). NZ's looming power crisis edges closer: an interview with Patrick Strange (Transpower CEO). On *TV3 News*: TVWorks Limited. Retrieved on June 12, 2008, from, http://www.tv3.co.nz/VideoBrowseAll/NationalVideo/tabid/309/articleID/58470/Default.aspx#video

Petersen, J. E., Shunturov, V., Janda, K., Platt, G., & Weinberger, K. (2007). Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives. *International Journal of Sustainability in Higher Education*, *8*(1), 16-33.

University of Otago. (2004-2008). *Calendar.* Dunedin: Author

University of Otago, Dunedin, Energy Studies Programme. (n.d., a). *About the weather station.* Retrieved January 26, 2009, from http://www.physics.otago.ac.nz/eman/weather_station/about_station.html

University of Otago, Dunedin, Energy Studies Programme. (n.d., b). *About the weather station.* Retrieved November 3, 2008, from http://www.physics.otago.ac.nz/eman/weather_station/archive.html

.