Chaos in the Residential Sector: Evaluation of Conservation & Load Management Programs

Christopher Palmer, KEMA Inc., Middletown, CT

Abstract

This paper will explore the evaluation process from site selection to metering options to data analysis, while insuring that the ISO New England (ISO-NE) evaluation standards are met. Another aspect of residential program evaluation discussed in this paper is site characteristic diversity. Program success is often dependent on factors such as the number of occupants in the household and the daily schedule of these occupants, all of which vary from site to site. The importance of each of these factors and their impact on potential energy demand reduction will be presented. The paper will also discuss the practice of pre-post metering, and its role in determining the program's success through the calculation of the Demand Reduction Value (DRV).

The paper will demonstrate the effect of voltage drift on program evaluations involving long term interval power metering. This effect was analyzed through the use of multiple metering options throughout the 60 residential sites involved in the study. Some sites involved the installation of Elite Loggers, as these meters record interval power measurements, which show any change in voltage. Other sites were metered using Time-of-Use Current Transformers (TOU CT) Loggers, which do not monitor voltage change. Spot power measurements were taken with a handheld power meter that met the ISO-NE standards requiring a True Root Mean Square (True-RMS) power accuracy of ±2 %. The spot power measurements were used in conjunction with the time-of-use data to yield interval power data. The multiple metering scenarios offered two methods of acquiring interval power data; comparison of the data demonstrated the impact of voltage drift and the importance of metering equipment compliance when evaluating residential conservation and load management (C & LM) programs.

Introduction

According to the Energy Information Administration, residential sites consume nearly 40% of all electrical energy generated and sold in the U.S. Therefore, it is crucial that C & LM programs are offered to not only the commercial and industrial (C & I) sectors, but to the residential customer as well.

One such residential application is the water heater lease program. Customers are offered the opportunity to lease a time-clock-controlled electric water heater. The time clock regulates the operation of the heating elements within the water heater, forcing the majority of the work to be performed during off-peak hours. The application of time control devices at the residential level is coupled with the opportunity for customers to take advantage of time of use rates. Residential customers pay a rate which becomes dependent on on-peak and off-peak hours; encouraging customers to become more conscious of energy use.

The application of residential C & LM programs must be followed by program evaluation, which can prove to be a daunting task. Evaluation at the residential level introduces new obstacles not encountered in the C & I market. These obstacles can range from having to convince the customer that you are indeed offering a free energy evaluation and are not a telemarketer all the way to ensuring that your site selection will produce a sample that adequately encompasses the vast site characteristic diversities; while also complying with all standards set forth by ISO and any other governing bodies involved in the regulation of program evaluation. Often times only limited information regarding each site is available, and the site selection process must be performed within the boundaries of such limitations.

Methodology

This off-peak water heater study involved a three month period of data collection over the course of June, July, and August of 2008. It occurred at 66[1] residential sites located within the United Illuminated Company's utility domain, which is comprised of the south central portion of Connecticut. Each of these sites consisted of participants in the water heater lease program, in which customers received a time-clock-controlled electric water heater through the utility company. The study examined the energy savings achieved at each site through the use of a time-clock-controlled water heater. The program's overall success was quantified as a Demand Reduction Value (DRV), which is the average demand reduction achieved during the on-peak performance hours of weekdays from 1:00 pm to 5:00 pm. This time frame is defined as the hours of interest by ISO-NE for Forward Capacity Market (FCM) program evaluation.

Site Selection

The design and selection of the on-site sample was the first phase of the evaluation. To estimate the sample size required to achieve ±10% precision at the 80% level of confidence (referred to as 80/10) the following formulas were used:

$$n_0 = \left(\frac{z * cv}{D}\right)^2 \qquad n_1 = \left(\frac{n_0}{1 + \frac{n_0}{N}}\right)$$

Where,

n_0 = the required sample size before adjusting for the size of the population
z = a constant based on the desired level of confidence = 1.282 for 80%
cv = coefficient of variation = standard deviation / mean = 0.64
D = desired relative precision = 0.10
n_1 = the required sample size after adjusting for the size of the population
N = the population size (2,031), i.e., the number of program participants

Working with a population of some 2,000 participants, a sample size of 65 was determined to be adequate. The sample design was selected to comply with ISO-NE Metering and Verification (M & V) requirements for a stand alone project, but since it was anticipated that UI would include this measure as part of their energy efficiency portfolio it was not necessary to achieve 80/10. The ISO-NE M&V Manual requires that a project in its entirety reach 80/10, and the water heater study is merely one component of the project. Therefore, the utility company is capable of achieving this requirement at the portfolio level without ensuring that each individual component of the project reaches this degree of certainty.

Of the 2,000 participants information was made available regarding tank size and geographic location of the residential site. Tank size was considered an important factor; as this variable could potentially dictate the number of people living within a household. It is safe to assume that the total demand is directly linked to number of people in the household because one would expect the more people in a household the higher the hot water usage. The distribution of the various tank sizes within the population

[1] Initially 66 sites were scheduled and visited, however due to issues such as customer cooperation, faulty water heater controls, and data lost to logger issues the data analysis focused on 60 of these sites.

was later compared to the distribution within the randomly drawn sample. This would show if the sample was statistically representative of the total population in regards to tank size, and if any bias adjustment would be necessary as a result of this finding. It will be shown later that the information regarding tank size was not closely tied to household occupancy; therefore, the tank size data was not as valuable as originally assumed. Geographic location, on the other hand, is not a significant factor when designing a sample of this nature. Average demand reduction achieved by the use of a time-clock-controlled water heater would not be impacted by the town in which a household is located; especially when dealing with a total population in such a localized area. It was determined that a random sampling would be representative of these variables. Had more information been available, such as number of occupants, a stratified sample could have been designed and implemented.

A random sample was drawn from the table and it was verified that the sample was representative of the program's geographic distribution. A secondary sample was randomly selected to serve as a back up for the purpose of replacing refusals and participants who could not be reached. A total of 66 sites were scheduled, visited, and logged as this allowed for an extra site in case any issues arose that would require the removal of sites from the sample. However, during the course of the study issues such as customer cooperation, faulty water heater controls, and logger malfunctions resulted in the elimination of 6 sites from the data set. It was determined that the 60 residential sites would adequately represent the 2,000+ participants, while still allowing for a relative precision of ±10% at the 80% confidence level assuming a coefficient of variation (CV) of 0.64. The CV used in the sample design was derived from a previous study on time-clock-controlled water heaters.

Pre-Post Metering

The best method for data collection was determined to be through performing pre-post metering. Pre-post metering is the art of gathering data both before and after the implementation of, in this case, an energy efficiency program. The practice of pre-post metering allowed for the collection of two data sets; one illustrating the energy demand of the water heater without the time clock and the other with the time clock in operation. The comparison of the two data sets revealed an energy savings or demand reduction.

However, the installation of the time-clock-controlled water heaters took place prior to the start of the program evaluation. Therefore, in order to develop a scenario of pre-post metering, multiple site visits were required. Upon installation of the loggers during the initial visit, a field representative disabled the electronic time clock on approximately half of the water heaters. The remaining time clocks were disabled during the second visit, and those previously disabled were reconnected. The period where the time clocks were disconnected established a baseline of water heater use, pre, which could then be compared to the period when the water heater's operation was time-clock-controlled, post.

This method of pre-post metering could potentially yield skewed data as a result of behavioral changes within the household during either the pre or the post periods. In order to incorporate any discrepancies between the metering periods into the data analysis the homeowners where interviewed at the end of the study. The homeowners were asked if they were aware of any changes in their daily patterns from one phase of the study to the next or from before the start of the study. Whether or not the household was unoccupied for a length of time due to vacation or changes in work schedule or if they had house guests for an extended period are examples of specific behavioral changes that would impact the results. Through these end of study interviews it was determined that behavioral changes would have no impact on the results and that the mock-pre period would be representative of the actual water heater operation prior to the installation of the time-clock-control.

Metering Options

The field representative was directed to the water heater where spot wattage measurements were performed and recorded. The spot wattage power measurements yielded an instantaneous reading of the volts, amps, and kilowatts of the heating elements on the water heaters. These power measurements were performed using a clip-on power meter, calibrated prior to the start of the program evaluation in order to comply with ISO-NE Forward Capacity Market (FCM) project measurement requirements. The meter had a True-RMS power accuracy of ±2%.

Metering was done with one of two different meter types each of which allowed for the collection of interval power data. The two meter types used in the program evaluation were: Dent Instruments Elite Pro Power Meters (Elite Loggers), and Dent Instrument TOU CT Loggers. Nearly 20 % of the sites were metered using an Elite Logger, set up to record interval power data ever 15 minutes, as required by ISO-NE. The remaining sites were metered using TOU CT Loggers, which yield transition and time-of-use data. The time-of-use data was later converted into interval power data using the spot power measurements.

Each site was monitored using one of three metering scenarios. The most prominent metering scenario involved 2 TOU CT Loggers, in which both heating elements were metered separately. This data showed the individual operating schedules of the upper and lower heating elements. Other sites were metered with only one TOU CT Logger; this data showed the power usage of the entire water heater. The remaining sites were metered using an Elite Logger, yielding interval power data showing volts, amps, kilowatts, and power factor every 15 minutes.

Results

Site Characteristics

Once on-site, a field representative performed a brief survey with each customer. The survey allowed for the collection of household demographics and other information regarding the amount of water consuming appliance such as showerheads and dishwashers in the residence and the frequency with which each of these appliances were used. Water heater specs, including model number, installation date, rated wattage, location, and general condition were also gathered. This information allowed for an understanding of each customer's baseline hot water consumption while also supplying variables by which the sample could be post-stratified. Much of the information gathered during the on-site interviews is redundant, but in order for the evaluator to wade through the discrepancies of customer reported data this is necessary. Many times it becomes the evaluator's job to hone in on the truth and draw a confident conclusion about site specific characteristics.

Figure 1 illustrates how some of the information gathered during the on-site visits can be used to develop an understanding of the household demographics. The figure shows the number of occupants within a household based on tank size. Nearly all sites had either an 80 gallon or 120 gallon capacity water heater, but to determine the correlation between the water heater storage capacity and the number of occupants in a household it proved helpful to chart out the data. For example, it is now known that 65 % of households with an 80 gallon tank have two or fewer occupants, while a lesser majority at 58 % of households with 120 gallon tanks have two or less occupants. The figure below highlights the fact that the correlation between tank size and household occupancy is weaker than originally expected. It was anticipated that a statistically significant relationship would have existed between the 80 gallon tanks and smaller household and similarly 120 gallon tanks and larger household. This was not the case. It is clear that when water heaters were purchased by households, occupancy was not the key component of decision making.

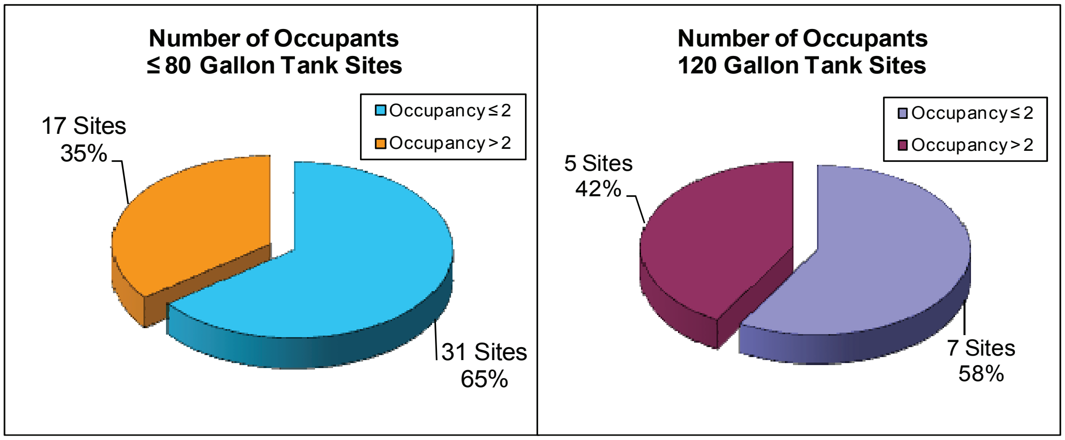

Figure 1: Example of household demographics gathered during on-site visit

Data Analysis / Demand Reduction Value

The data from each site consisted of power data for the water heater both with (post) and without (pre) the time clock enabled. For the purpose of quality control inspection all interval power data was brought into KEMA's Visualize-It® software. This allowed for a visual analysis of the power data in order to determine if there were any visible flaws in the metered data. Figure 2 is a Visualize-It® print out of the pre and post periods for one site. The blue line, representing the pre period, appears to be demand driven, meaning the water heater turns on and off based on the need for hot water. In contrast, the red line, representing the post period, shows that the water heater does not operate during midday on-peak hours, but instead functions during the middle of the night.

The data was then converted into hourly interval data from hours 1 through 24 for analysis. Each site consisted of two 24 hour data sets, pre and post. The difference between the two data sets at each hour was calculated illustrating the hourly demand reduction or increase. This change in energy demand based on the comparison of the pre and post metering is clearly portrayed in Figure 3. Hourly demand reduction is illustrated in the chart as a negative value, highlighted in blue, and the hourly demand increase is a positive value, highlighted in red. It is clear that savings, or demand reduction, is achieved between hours 6:00 AM to 10:00 PM, or 6 to 22 as shown in the figure. On-peak hours fall within this time frame where the demand reduction occurs. This verifies that the demand reduction is occurring during on-peak hours, as intended by the water heater lease program, forcing the majority of the work performed by the water heater to occur during off-peak hours.

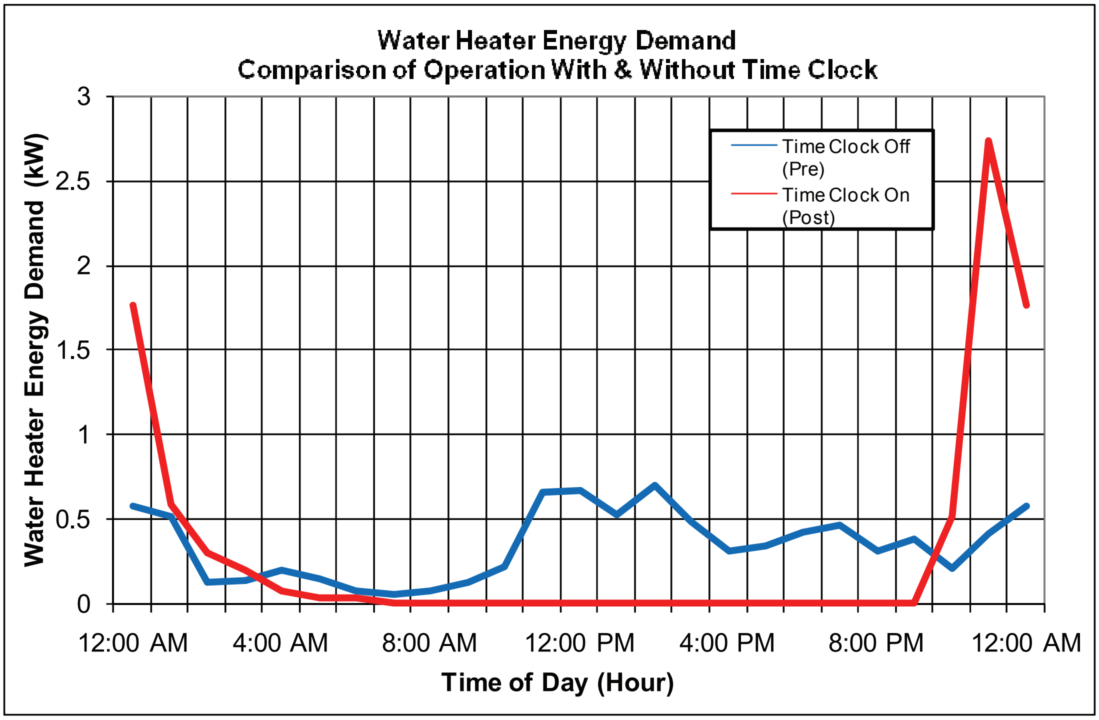

Figure 2: Visualize-It® printout of pre and post water heater data (KEMA Inc.)

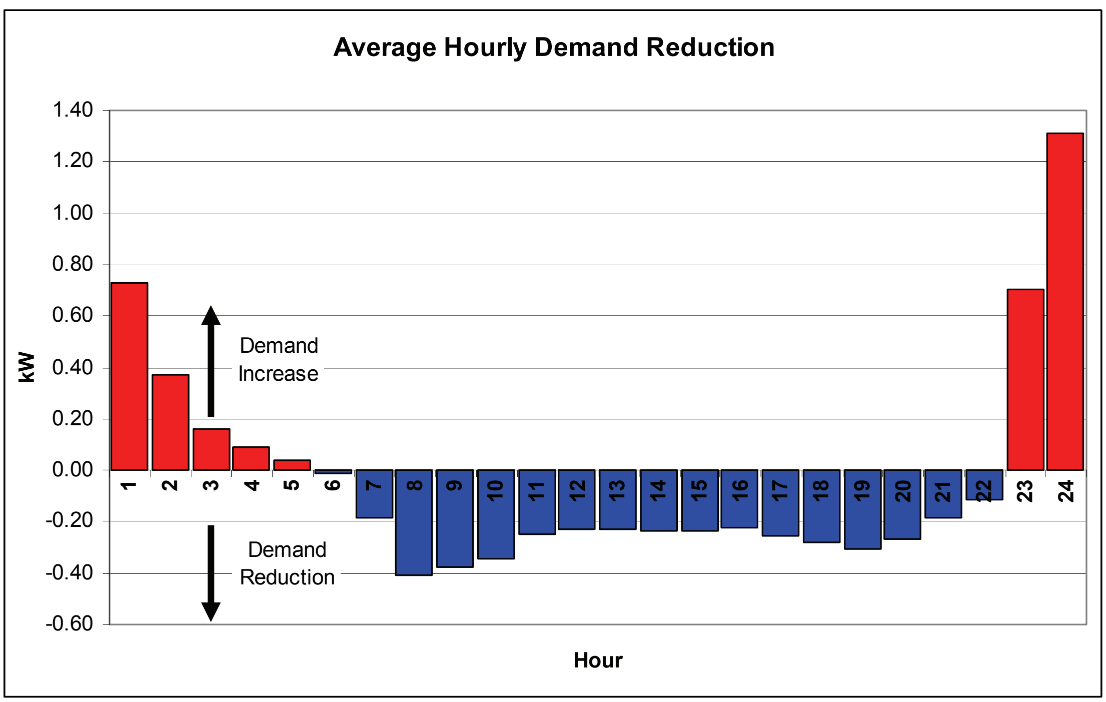

Figure 3: Average demand reduction/increase based on comparison of pre & post data

The data was analyzed to determine the Demand Reduction Value (DRV), a value used to quantify a program's overall success. The DRV represents the program's average demand savings achieved at all sites involved in the sample. This value is calculated during the ISO New England FCM summer on-peak performance hours, between hours 1:00 PM to 5:00 PM, or hours 13 to 17 in Figure 3. An average is calculated for each site during this time frame, and then the site specific values are averaged together yielding a program DRV. Table 1 shows the analysis results for the DRV calculation.

Table 1: Analysis results for all sites during ISO-NE Summer On-Peak hours

Unweighted DRV	0.24 kW
Standard Deviation	0.19
Coefficient of variation	0.79
Relative Precision	±13%

The initial calculation of the DRV utilized the unweighted data for each site during the pre and post metering periods. The overall DRV was calculated as the average of all the sites with no adjustments for sample selection bias. The practice of post-stratification was implemented to account for such bias.

Post Stratification

When enough information about a given population is made available prior to the drawing of a sample, stratification can be employed. Stratification of a population is the grouping of sites with similar characteristics into categories referred to as stratum. A stratified sample can then be drawn in order to proportionally represent the entire population. The use of stratification in sample design can reduce sampling error, and increase the accuracy of the estimation by decreasing the relative precision and the coefficient of variance. As is the case in this study, there was limited information made available prior to sample design. Without enough data to develop strata it was decided that a random sampling would be sufficient to develop a representative sample. Stratification would have been possible had more site characteristics been readily available prior to the program evaluation.

However, it is possible to stratify a sample during the data analysis phase of the study. This is referred to as post-stratification. Post-stratification is the stratifying of a sample after the sample has already been selected and data has been gathered and site characteristics have been defined. Certain factors specific to each site, such as number of occupants and tank capacity, can affect the expected demand reduction. Site characteristic diversity is a driving force in the achievable demand reduction, and can result in a large DRV spread within a sample. This was the case within this sample where the average program DRV of 0.24 kW fell within a range of -0.10 kW and 0.98 kW. To verify that trends specific to site characteristics existed within the sample and to determine their impact on the demand reduction, the sample was post-stratified using the information collected on site.

The primary reason for the post stratification of the data was to adjust for the sample selection bias and to investigate whether the overall relative precision of the DRV would improve when the data was stratified. Although there were a lot of variables evaluated for their potential impact on the data, including tank size, number of occupants, number of showerheads, number of showers per week, etc, only two variables, tank size and number of occupants, proved to have a clear correlation with the savings results. Since tank size was the only variable that was also available in the tracking data provided by the utility company, the distribution of number of occupants observed in the sample was assumed to be representative of the population.

Prior to stratification of the sample, the DRV was calculated assuming equal weights for each site. However, as shown in Figure 4, over-sampling of the 120 gallon water heaters did occur. The population consisted of about 15 % 120 gallon tank water heaters while the sample included 20 % of these larger tanks. This would clearly cause a sample bias in the results causing an increase in the calculated demand reduction.

Therefore, it was necessary to evaluate the demand reduction value using population weights that reflect the frequency of the tanks within the total population. This would adjust the sample back to the population ratio so that the bias would not be included in the final results.

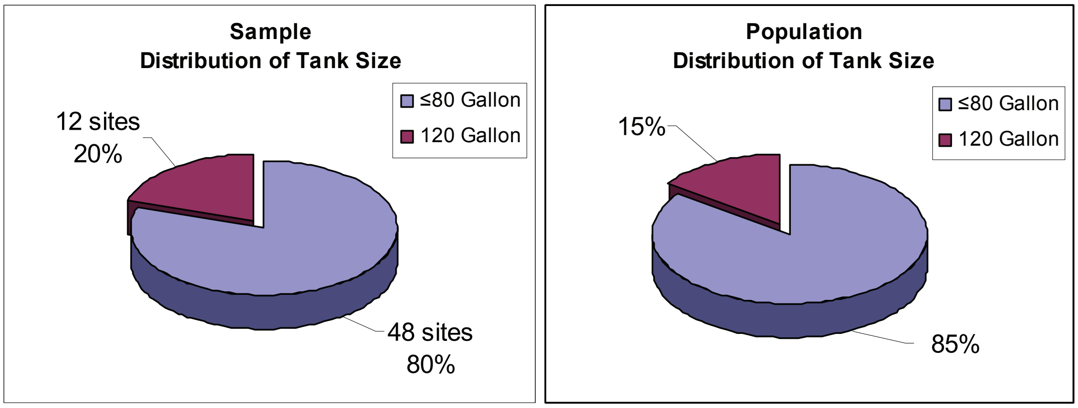

Figure 4: Distribution of tanks sizes in the sample and the total population

The data was also post-stratified based upon the number of occupants (less than or equal to 2 or greater than 2) within each tank size category, resulting in a total of four strata. As stated earlier, this relies upon the assumption that number of occupants observed in the sample was consistent with that of the population. This is represented in Figure 5, where green illustrates the four strata used in the stratification of the sample and weighting of the DRV. Population weights were calculated for each stratum, reflecting the frequency that each type of water heater scenario would occur in the population. By weighting the final results of the study the DRV changed from 0.24 kW to 0.23 kW. Therefore, the program evaluation of the time-clock-controlled water heater concluded that the final DRV was 0.23 kW.

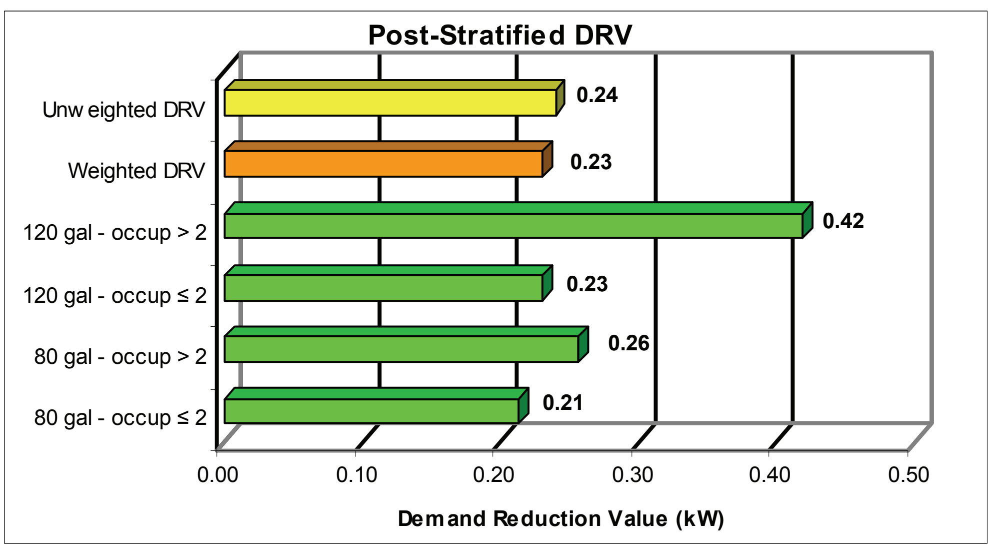

Figure 5: Unweighted, weighted, and strata demand reduction values

Conclusion

Site Characteristic Impact

Site characteristic diversity is a norm in the world of residential program evaluation. Within any given sample, whether it be randomly drawn or developed through stratification, it quickly becomes apparent once the on-site visits start that each location will introduce new obstacles and diverse lifestyle patterns directly affecting the data. The impact of site characteristics on a residential program evaluation is made evident through the post-stratification of the sample. The program's DRV changed as a direct result of organizing and grouping similar sites and developing strata. Although this change was not drastic, correcting the DRV from 0.24 kW to 0.23 kW, it ensured that the final result accounted for the diversity observed from site to site.

During the on-site visit much of the time is spent gathering detailed information from the homeowner. Much of this information is redundant, but it will enable the evaluator to triangulate in on useable data. This self-reported data is frequently skewed, but the evaluator is able to combine the gathered information to draw accurate conclusions about the household, and in this case the hot water consumption. The gathering of detailed site characteristics information may seem a nuisance during the on-site visit portion of a program evaluation, but as the evaluation process evolves and the data analysis phase begins it becomes clear that the additional information can be very helpful. Having detailed site specific information readily available can often ease the burden of troubleshooting when the road blocks of program evaluation emerge. Time and again discrepancies occur within a data set and assumptions must be made. Having such information on hand can aid in the correction of discrepancies and can also be used to justify assumptions made by the evaluator. When data appears to be flawed for no rhyme or reason detailed site information can prove to be the lifeline in understanding inconstancies within a data set.

A comparison of**Error! Not a valid bookmark self-reference.** and Figure 6 reveals how the on-site interviews can be used in conjunction with the demand data as a means of interpreting the water heater's energy use patterns. The pre period, represented by the blue line in the figure, shows a pattern of marginal demand fluctuation within a range of 0.1 to 0.8 kW with a peak usage between the hours of 8 PM and 9 PM. By assessing the data made available in the table, one can see that the majority of hot water consuming activities are performed after 8 PM, as is encouraged by the time-of-use rate. Therefore, the pre-period peak demand is validated, and a trivial amount of hot water is used until late afternoon. For this site it can be said that the demand in the first half of the day is most likely the result of cycling or reheating of the stored water due to standby heat loss. Interpretation of the data at this level can prove to be a useful tool when attempting to understand the effects of lifestyle patterns on energy demand.

Table 2: Site 1 Characteristics/On-site Collected Info

Site 1		
Dishwasher	Quantity	1
	per week	4
	Time	After 8 PM
Clothes Washer	Quantity	1
	per week	2
	Time	After 8 PM
Showers	Quantity	1
	per week	4
	Time	After 8 PM
Household Occupancy		1

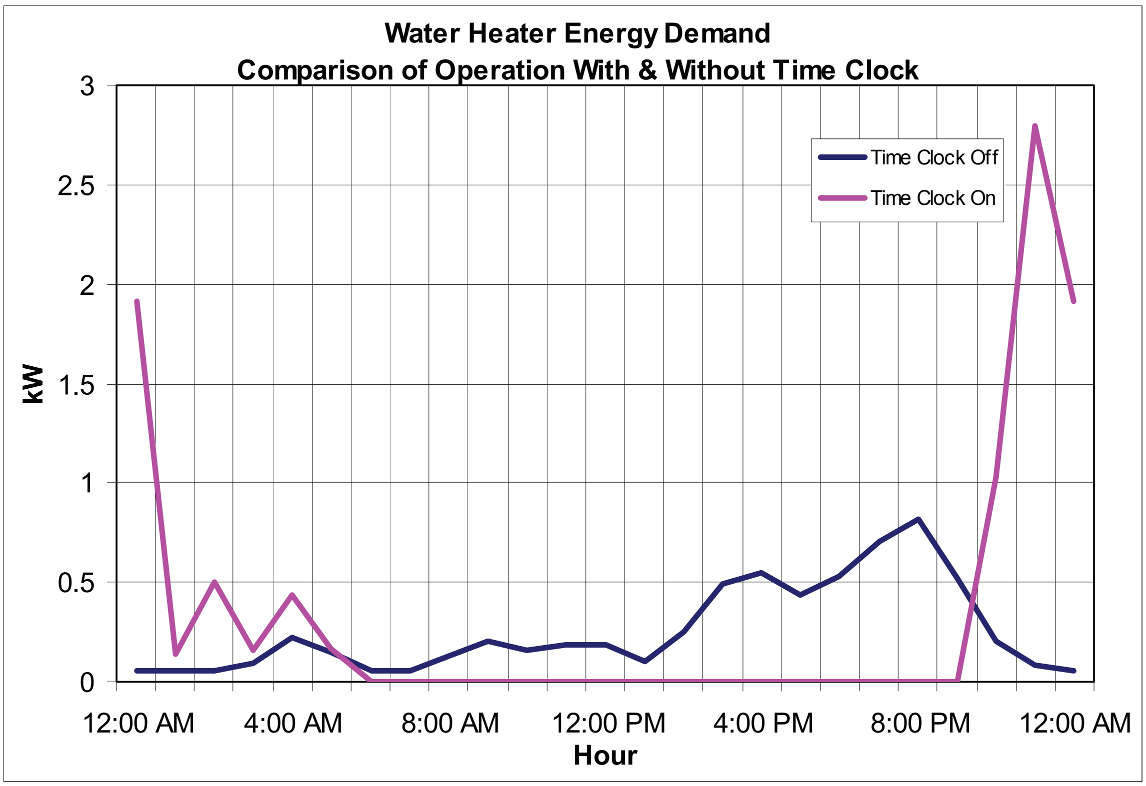

Figure 6: Site 1- Visualize-It® printout of pre and post water heater data (KEMA Inc.)

Customer Response

If a primary objective of the program was to make the owners of a time-clock-controlled water heater aware of their hot water based energy consumption then the program can be considered a success. This was accomplished by offering the customers time-of-use rates. Homeowners were encouraged to operate their water consuming appliances during off-peak hours, as this would incur the lowest operating cost. During the on-site interviews customers made it quite clear that they were eager to capitalize on this opportunity to save on energy cost. The majority of individuals had adjusted their daily schedules in order to insure that the washing machine and dishwasher were run during off-peak hours, and when possible showering was done during this lower-cost time period as well. One of the more common complaints was the occasional lack of hot water. For some of the customers if hot water consumption was slightly higher than typical the water tank would run low, resulting in a period of no hot water. This did not appear to be an issue at the majority of sites, and it seemed the problem could be avoided if the customers reworked their daily schedules in order to more efficiently use the available hot water. Some of the time-clock-controls, about 36%, were equipped with an override button. This feature gave the homeowner the option to override the time clock control with the push of a button. The override would allow the water heater to operate unregulated for two hours, but most homeowners with this option were not aware of it. Overall, customers were generally pleased with the program and willing to adjust their daily schedules in order to optimize their savings.

ISO – NE Standards

The ISO-NE Load Response Program (LRP) Manual and the ISO-NE M&V Manual for FCM are both relevant to studies of this nature. However, it became evident that a contradiction between the two exists. The ISO-NE LRP Manual relates to the study because of its mention of constant load devices. It states:

"A constant load device is one that operates at the same demand (kW) whenever it is on, such as a bank of fluorescent lights controlled by a single switch or a single-speed compressor in a packaged air conditioner unit. Since demand is rarely perfectly constant, a load can be considered as constant if it varies by no more than 5-10% from its average value during operation."

An electric water heater utilizes purely resistive heating elements as the source of heat output making it a prime example of a constant load device. The trademark of a true resistive load is that it is assumed to operate at constant amperage. An understanding of the demand (wattage) calculation in its simplest form, the product of amperage and voltage, illustrates the point that demand fluctuation is a result of voltage drift. The ISO-NE LRP Manual clearly states that "…A load can be considered constant if it varies by no more than 5-10% from its average value during operation." This statement could be interpreted as an allowance of acceptable voltage drift within the range of 5-10% because the demand fluctuation is driven by instantaneous voltage measurements.

On the other hand, ISO-NE M&V Manual for FCM addresses issues regarding to the monitoring of these types of project. The excerpt below specifically references the use of monitoring equipment which does not directly record electrical demand measures.

"Any measurement or monitoring equipment of proxy variables that do not directly measure electrical demand, including but not limited to voltage, current, temperature, flow rates and operating hours, must have an accuracy rating such that the overall accuracy of the calculated demand (kW) using the proxy variables is not less than ± 2%."

The TOU CT Loggers are a prime example of this type of monitoring equipment, as they strictly record transitional and time-of-use data. The data was then converted into a calculated demand (kW), using spot power measurements acquired from the on-site visits. This section of the manual requires that the overall accuracy of the calculated demand be within a range of ± 2%. It is difficult to assure that the fluctuation of the calculated demand is within 2% of the actual demand when referring to constant load devices like the electric water heater. This is because constant load devices, as mentioned in the ISO-NE LRP Manual, are prone to fluctuation of 5-10% due to voltage drift over the course of the study. The ISO-NE FCM M&V would have precedence over the LRP Manual, requiring the ± 2% standard for proxy variable measurements, but this standard is near impossible to achieve because of constant load device fluctuation. It can be concluded that this particular ISO-NE FCM M&V standard is non-essential for a program evaluation of this nature, specifically studies revolving around constant load devices.[2]

Voltage Drift Assessment

As previously stated, multiple metering scenarios were employed throughout the sample in order to assess the impact of voltage drift on the calculated DRV. The majority of sites were metered with TOU CT Loggers, but this method did not provide enough data for voltage drift analysis. In order to gather detailed interval data which could be used to quantify the effects of voltage drift, Elite Loggers were installed at several locations. These Elite Loggers made 15 minute voltage, amperage, kilowatt, and power factor data readily available.

For the purpose of voltage drift assessment in this application it was only necessary to analyze the voltage data between the hours of interest, 1 PM to 5 PM, because this is the time frame in which the DRV is calculated. After removing the excess voltage data from the set a statistical analysis was performed in order to calculate a relative precision. In the case of random sampling, the method used to develop the sample for this study, relative precision is a ratio of error across a sample based upon sample size, mean, and variance. Here the definition of relative precision can be interpreted as the average fluctuation seen in the voltage data at each specific site. An average relative precision was calculated as ±0.06 % among the sites metered with Elite Loggers. In order to calculate the total relative precision for the entire sample it was necessary to combine the uncertainty of the spot power measurement with the uncertainty due to the voltage drift. This incorporated the sites metered using spot power measurements and TOU CT loggers. The handheld meter used for the spot power measurements has a rated relative precision for demand measurements of ±1.5 %. A root-sum-square value, shown below, is used to calculate a total relative precision through the combination of the calculated relative precision of the Elite Logger voltage data and the rated relative precision of the handheld meter.

$$RP_{Total} = \sqrt{RP_a^2 + RP_b^2}$$

Where,

RP_{Total} = total relative precision of the voltage data within the entire sample (±1.501 %)
RP_a = relative precision calculated at sites metered with Elite Loggers (±0.06 %)
RP_a = relative precision of the handheld meter (±1.5 %)

This calculation yielded a total relative precision value of approximately ±1.5 %, highlighting the fact that the typical voltage drift observed in the data did not have a significant impact on the uncertainty of the overall power measurement. It can be stated the voltage drift within the entire sample is within ±1.5 % of the average voltage at each site. Therefore, the voltage drift does not give rise to any issues with respect to metering and verification compliance as explained in the ISO-NE FCM M&V, and the ISO-NE LRP Manual.

[2] This ±2% accuracy requirement for demand calculations using proxy variables has been widely criticized for overreaching the intent of the metering requirements section and touching on data analysis procedures. PJM Interconnection utilized the ISO-NE M&V manual as a starting point for their manual, but dropped this proxy variable accuracy requirement.