Disaggregation of End-Use Load from Whole House Interval Meter Data

Jeffrey Phung, ADM Associates, Inc., Sacramento, CA
Thomas Adkins, SMUD, Sacramento, CA
Dan Mort, ADM Associates, Inc., Sacramento, CA

ABSTRACT

End use monitoring provides evaluators with a detailed picture of the operating schedule and intensity of a given end use product in a building. However, end use monitoring can be difficult and costly to set up and secure. Whole building hourly interval meter data, as compared to monthly billing data, offers high resolution data. However, whole building meter data does not readily disaggregate to specific end uses.

We conducted a program evaluation on SMUD's 2013-2015 Residential Pool/Spa Program, which provided downstream rebates for variable speed drive (VSD) pool/spa pumps. ADM Associates utilized both an engineering analysis and an econometric analysis to evaluate the energy savings of this program.

Baseline data was not captured for the pumps prior to replacement. Single-speed pool pumps operate in a recurring pattern and controlled by a 24-hour timeclock. ADM theorized that this should generate a repeating, rectangular load, on a daily basis that could be observed in a customer's advanced meter infrastructure (AMI) whole house interval meter data. Armed with this theoretical precept, ADM created an approach to extract the baseline data to supplant the operation schedule and hourly kW from the pumps in the baseline period.

Introduction

End use monitoring provides evaluators with a detailed picture of the operation and load of a given end use. However, monitoring can be limited by a project's timeline and budget. Furthermore, the ability to conduct baseline end use monitoring can be outside of an evaluator's window of opportunity. Program evaluators may wish to explore baseline period behavior long after changes have taken place. While end use monitoring on similar buildings may be used to provide a "proxy" result, this type of analysis does not provide a substitute for a specific building's behavior. Rather, it is only interpretable to the extent that it is generalized and then applied to a representative group. AMI data is becoming ever-present in the marketplace and provides utility interval meter, customer-specific data. This data, however, is captured at a whole-building level and does not disaggregate to specific end uses. A significant amount of progress has been made in the extraction of HVAC loads from whole house data using weather data. However, little progress has been made in the disaggregation of other types of specific end use loads from interval meter data.

Beginning in 2013, Sacramento Municipal Utility District (SMUD) implemented a Residential Pool and Spa Program as part of their energy efficiency portfolio. This program provides downstream rebates to customers who purchase a high efficiency VSD pool or spa pump. Customers participating in this program fall under two scenarios: (1) replacing a less efficient existing pool/spa pump, or (2) installing a pool/spa pump on a newly installed pool/spa. Customers replacing an existing pool/spa pump may be replacing a pool pump due to failure (replace on burnout) or they may be replacing a functional pool pump prior to the end of its lifespan (early replacement).

ADM evaluated the gross energy savings of SMUD's Residential Pool and Spa Program for program years 2013-2015. ADM was required to establish 24-hour load profiles for the baseline pumps, the VSD pumps, and gross energy savings estimates and peak demand savings estimates.

There were specific limitations for conducting gross impact evaluation for this program. Customers who participated in the program were not required to submit information regarding existing equipment; therefore nameplate information and operating schedules for equipment in the pre-installation (baseline) period was not available. ADM was only able to collect primary data for the post-installation equipment and schedule of operation.

Despite these limitations, Advanced Metering Infrastructure (AMI) had been implemented throughout the SMUD service territory starting in early 2012. AMI data was available for all participants at an hourly resolution. ADM was tasked with extracting pre-installation load profile information from pre-installation hourly interval meter data.

Residential Pool/Spa Pumps

Pool/spa motors are used primarily for running pool/spa filtration systems. A pool/spa is filtered by pumping water through a filter. Filtration systems are typically controlled by a control unit, which automates the start and stop times for pool/spa pumps on a daily basis. Current regulations do not dictate the frequency or duration that pool/spa pumps need to operate in residential pools (Worth 2013). The run-time and operating schedule for most residential pool or spa pumps is determined by a third-party pool maintenance company employed by the property owner or the property owner themselves, and may not be optimized to the volume of the pool or the flowrate of its corresponding pool pump.

There are three different types of pool pump motors: single-speed, two-speed, and VSD pumps. The nomenclature for these pumps refer to the difference in the ability to adjust the speed at which the motor can be set. Single-speed pumps are designed to operate at a single, non-adjustable full-speed and represent the historical tradition for pool pumps. Two-speed pumps can be operated at full-speed or half-speed. VSD pumps can be operated at a multitude of different speed settings, with a full range of variability. VSD and two-speed pumps provide energy savings over single-speed pool pumps because the energy to pump a gallon of water is lower as the speed of the pump is reduced. A control unit on the VSD allows the operator to not only program the start-time and duration that the pump is operated, but also the speeds at which the pump is set during its scheduled run-time. There is no automatic feedback to control the speed of either the two-speed or variable speed pumps.

A baseline study was conducted by ADM for SMUD in 2008 for residential pool/spa pumps. That baseline study found that the majority of homes sampled and visited operated their pool/spa pump at a single speed for a single, multi-hour interval per day (91%), with the hours of averaging 6.6 hours with a standard deviation of 2.6 hours. ADM expected homes participating in SMUD's rebate program to replace pool/spa pumps of similar nature to those found in the baseline study.

The previous baseline study shows that pool/spa pumps have a scheduled onset and offset that repeats on a daily basis, creating a rectangular load profile that occurs at the same time daily. Using this characteristic, ADM developed a method for extracting the end use load shape from hourly AMI data for the pre-installation condition. Here we discuss ADM's approach to extracting the pre-installation, single-speed load shape, and the broader uses that may be explored in the future.

Methodology

ADM employed two different techniques to evaluate the program energy and demand savings. The goal in utilizing the two techniques was to assess the appropriateness of both techniques and compare the results to those in previously established sources. The two approaches can be described as follows:

1. An engineering analysis – Comparing the operating conditions from the pre- and post-installation periods. The engineering approach compares the hours of operation and kW of the post-

installation period to the hours of operation and kW of the pre-installation period. We developed an approach to estimate both the hours of operation and the kW load of the baseline pumps using AMI data. This approach is detailed in the section “Baseline load profile development.” Hours of operation and kW for pumps in the post-installation period were gathered from customer on-site visits, as detailed in the section “Efficient load profile development.”

2. An econometric analysis – Comparing a participants’ whole building consumption data for the twelve months after installing the new pump to the twelve months prior to installing the new pump. This approach uses a weather-normalized fixed effects regression to compare whole-building consumption before and after installation of the measure for each participant. This analysis represents a typical approach and is detailed in the section “Econometric Approach.”

For both approaches, ADM also asked sampled participants to complete a questionnaire, which was used to capture participant profile information that may be relevant for the study. Additionally, using the information obtained in the engineering analysis, 8,760 hour load profiles were generated both for the pre-installation period and for the post-installation period.

Data Sources

Data sources and how they were used in the analysis are documented in Table 1.

Table 1. Description of data sources

Data source	Purpose/Application
Program tracking data	Used to generate samples and invite customers to participate in the program evaluation.
Hourly AMI data	Used in the generation of pre-installation operating schedules for the engineering analysis. Also used for the fixed effects regression used for the econometric analysis.
Sacramento County Assessor data & Google Earth™	Used to determine the installation dates of pools and determine whether customers installed pumps on a new construction pool or existing pool.
Weather Underground	Historical weather data used to weather-normalize the fixed effects regression model used in the econometric analysis.

Sampling

A stratified random sample design was used to represent the three program years in the final sample. The total sample size was designed to be sufficient across the three program years to obtain a relative precision of ±10% at the 90% level of confidence. A total sample size of 67 participants was determined to be sufficient to reach this level of precision. For design purposes, an initial target sample size was set to 75.

The sample was stratified by year based on the proportion of participants in each program year to the total program participation in all three years. This resulted in target sample sizes of 15, 25, and 35 for 2013, 2014, and 2015, respectively.

A random sample of participants for each program year were contacted for participation in an on-site visit and complementary web survey. Table 2 provides the number of sampled participants by program year. The number of surveys is a subset of the number of site visits. Although all site visit participants were asked to complete the survey, responses could not be collected for ten participants.

Table 2. Sampled participants by program year

Program year	Number of installed units	Number of site visits	Number of surveys
2013	572	15	14
2014	781	25	21
2015	1,077	40	35
Total	2,430	80	70

Sampled participants participated in a site visit during which photographs were taken of the installed equipment including nameplates, the control unit, the VSD pump, and the pool sweeper motor where applicable. Additionally, the field technician recorded the pool/spa pump run-time schedule as listed in the control unit, as well as the corresponding speeds for each of the scheduled run-times. Using a True RMS[1] power meter, the ADM field technician recorded spot measurements at each of the scheduled speed settings. Power measurements included volts, amps, power, and power factor.

All sampled participants were also asked to complete a survey about program participation, their previously and newly installed equipment, their motivations for participation, and their overall satisfaction with the program.

Engineering Analysis

The following section will detail the engineering analysis including: using AMI data to generate the baseline load profile, using spot-measurements to generate the efficient load profile, and the calculation of unit savings.

Baseline load profile development. As noted previously, the properties needed to compare the post-installation pool pumps to the pre-installation pool pumps were not captured during the administration of the program. In order to capture the operating schedule and kW of the pool pump in the pre-installation period, ADM opportunistically utilized known characteristics about pool pump operations and other end uses in the residential sector. Baseline pool/spa pumps are controlled by 24-hour timeclocks, which run on a repeating, daily schedule. Other end uses, such as lighting, refrigeration, and HVAC are dynamic and unscheduled in their behavior. Therefore, the pool/spa pumps have a higher probability of appearing at a consistent time day after day with the same kW load.

Figure 1 provides an example of a customer’s whole building hourly kW over the course of five sequential days. As can be seen in the first day, a rectangular load turns on around 4 a.m. and turns off around 11 a.m. This load continues to be present in the sequential days, but scatter begins to be introduced in the following days due to other, unscheduled loads. From the first two days, a visual observation of the data shows the pump is just under 2 kW and operates from 4 a.m. to 11 a.m. Next, we develop a method to automate the extraction of this information so large volumes of data can be processed.

---

[1] Root Mean Squared

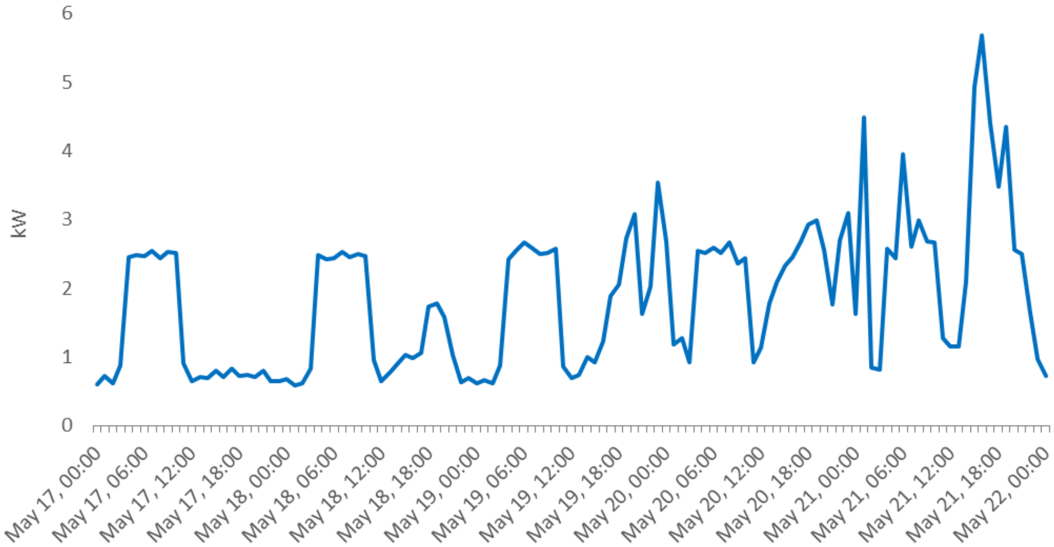

Figure 1. Sample of a customer's interval meter data for five sequential days from May 17-22, 2013

To capture the kW load of the pumps and the operating schedule, ADM transformed the data by looking at the differential between a given hour's kW and the hourly kW from two hours before. Because pool pumps could be turning on during a partial-hour, ADM used the difference between two periods to represent the full hourly load represented by the end use. Figure 2 presents this transformed data. Positive values represent times in which loads turn on, flat portions of time represent times in which loads do not change, and negative values represent times in which loads turn off. The peaks in relation to the start and stop time of the rectangular profile seen in Figure 1 continue to re-occur over the course of multiple days, while unscheduled loads appear sporadically. Therefore, one should be able to detect the start and stop time of the rectangular load profile by these times appearing naturally as the most frequent values over the course of multiple days.

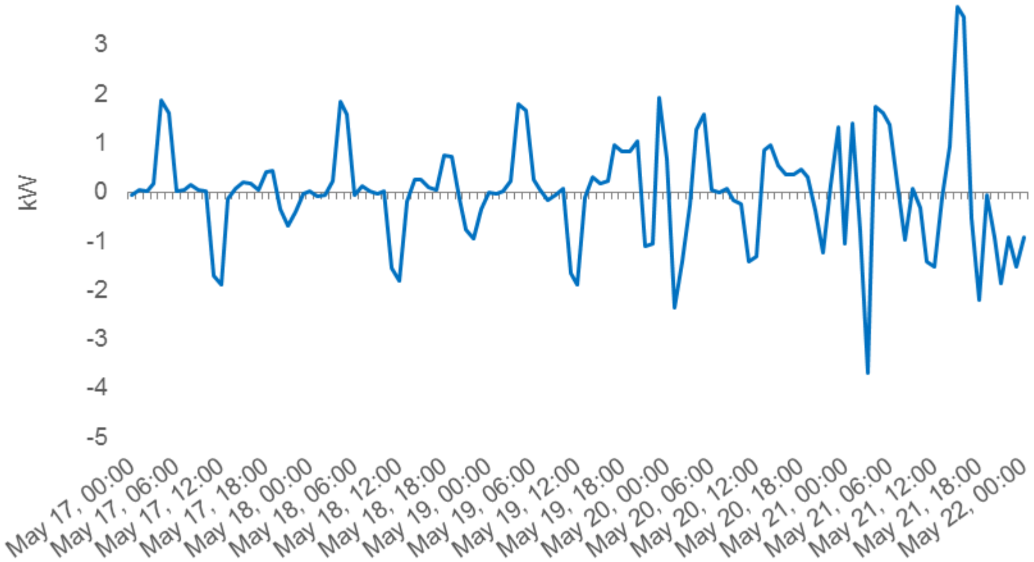

Figure 2. Sample of interval meter data after undergoing the two-hour-prior kW subtraction

In order to reduce the amount to which unscheduled loads contribute to the hourly pump load profiles, ADM first isolated the three days per month with the lowest daily energy use in kWh. These days provide good candidates in which the peaks and valleys associated with the rectangular end use profile can be detected even when the on-boarding and off-loading of other end use loads are superimposed on top of them. An average daily profile per month was generated using these three days. By averaging the three days, the rectangular signal is retained while minimizing contribution of the other end uses. Figure 3 provides an example of this averaging process using the first three days presented in Figure 1. As evidenced by the red-dashed line, the rectangular profile at hour 4 becomes much sharper and stronger while other end use profiles become flattened by this process.

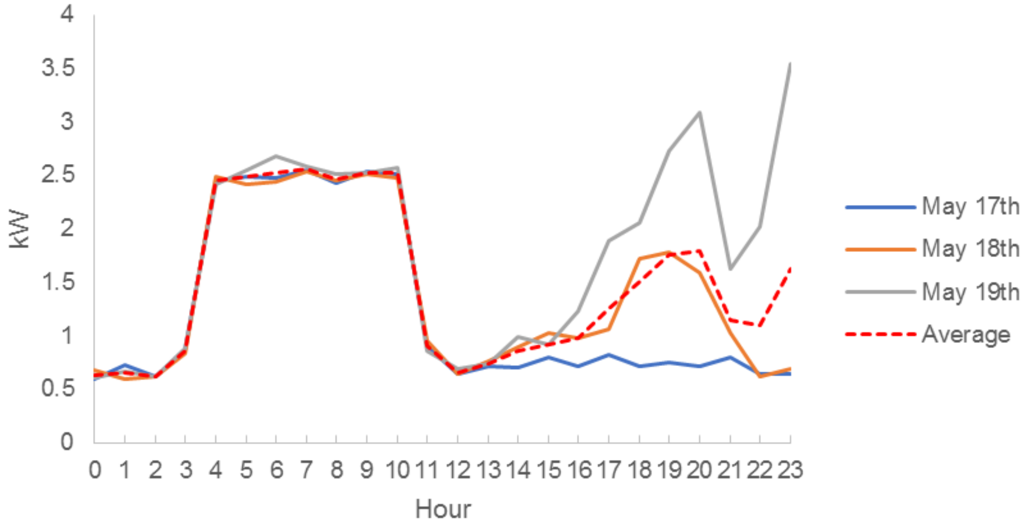

Figure 3. Sample of interval meter data after undergoing the averaging process

After obtaining these average lowest daily profiles per month, the 2-hour kW transformation is then applied by subtracting the kW from two hours prior to a given hour from the current hour's kW. This provides the size of the load in kW.

As part of the customer's survey, customers are asked when they would have changed their pool pumps from the summer to winter schedules and vice versa. These dates are used to segregate the year into two distinct time periods. For each time period, the mode of the positive time-points and the mode of the negative time-points is assumed to be the pump's "on-time" and "off-time", respectively. The mode of the absolute value of the peaks at the "on-time" and "off-time" peaks is determined to be the hourly kW of the pump.

While the above method is able to determine the kW when the pump is operating for the full hour, the pool pump may be turning on part-way through the hour. Next, the partial hour start and stop times are determined to fine tune the duration of pump operation. The daily profile was transformed again, but this time by taking the difference between the kW in a given hour and the kW in the previous hour. To determine the number of minutes the pump ran in the hour before the pump was fully on is calculated by taking the ratio of the kW difference between the first full hour and the previous hour and the full kW of the load then multiplying by 60 minutes. A similar process is done for the turn off transition. A 24 hour load profile for the pre-installation pool pump was then generated by taking the operating schedule, the partial hourly kW at the start and end hours, and the full hourly kW for the hours between the start and end time.

Efficient load profile development. Unlike the baseline load profiles, which were developed using the whole house hourly interval meter data, efficient load profiles (VSD profiles), were developed using a combination of data collected on-site and participant survey data. Table 3 details the data used.

Table 3. Description of data sources for the VSD profile development

Data type	Data source	Purpose/Application
Pool/Spa Pump Control Unit Data	Site-visit	Control unit with automated schedules of operation for the VSD pump. Includes the start/stop time of each run-time interval and the speed (in RPMs).
True RMS Multimeter Data	Site-visit	Used to record spot measurements of the voltage, current, wattage, and power factor for the pool/spa pump. A technician switched the pool/spa pump control unit into manual mode and ran the motor at each of the specified speeds listed in the control unit schedule to obtain the spot measurements.
Operating Schedule Change Dates	Survey	Used to determine if and when the customer switched from the winter operating schedule to summer operating schedule and when the customer intended to switch from summer operating schedule to winter operating schedule.[2]

Using the operating schedule observed by the field technician and the one-time measurements, ADM constructed a 24-hour summer load profile was constructed for each of the participants.

A winter schedule was not available at the time of field site visits. The information was not captured in the participant survey due to the likelihood of recall error and likelihood that customers were relying on a pool service to program the operating schedule of their pumps. ADM estimated a winter run-time schedule by taking the ratio of the hours of operation during the winter for a participant to the hours of operation during the summer for that same participant in the single-speed baseline condition.

ADM then utilized the dates that participants reported changing their pool/spa pumps from their winter to summer schedule and summer to winter schedule to generate an 8,760-hourly profile per participant. For instances in which the participant did not provide a response regarding the estimated operating schedule change dates, the mode was substituted for the missing data.[3] We then created an average 8,760-hourly profile by averaging the kW per hour across all participants.

Savings calculations. Annual energy savings for the engineering analysis were calculated by taking the 8,760 hour load profile energy use for the efficient pump and subtracting it from the 8,760 hour load profile energy use for the baseline period. Peak demand savings were calculated by taking the average per-hour savings for the peak demand window, as defined by SMUD.[4]

---

[2] The 2008 baseline study found that the majority of participants in this territory switched operating schedules for pools twice a year—once in mid-spring and once in mid-fall. From the survey responses, 8.5% of participants responded that they did not change their operating schedules, 23% did not know, and 68.5% provided a response.

[3] Several participants only operated their pumps at a single schedule across the year. Imputing missing data using the mode allows for the chance that a customer would have been an "n/a" if they had responded to the question.

[4] The peak demand window is defined as 4 p.m. to 7 p.m. on non-holiday weekdays for the months of June through September.

Econometric Approach

In contrast to the engineering analysis, ADM also conducted an econometric analysis of whole-house consumption data on the sampled participants. The econometric analysis utilizes weather-normalized fixed effects regression modeling, which represents the common approach for calculating gross impact savings in energy efficiency programs when using whole house meter data.

The regression model used in estimation of average annual energy savings is specified in the following equation:

$$AECit = \alpha_{ki} + \alpha_{02}POST_{it} + \alpha_{11}CDD_{it} + \alpha_{12}CDD_{it} * POST_{it} + \alpha_{21}HDD_{it} + \alpha_{22}POST_{it} * HDD_{it} + E_{it}$$

Where:

- α_{ki} is a customer-specific intercept term that adjusts for each individual customer's base usage as differentiated from the usage of the group as a whole.
- α_{11} and α_{12} are coefficients that adjust for the customer's cooling season weather-sensitive usage, as differentiated from the usage of the group as a whole.
- α_{21} and α_{22} are coefficients that adjust for the customer's heating season weather-sensitive usage, as differentiated from the usage of the group as a whole.

The average per home annual savings value was then calculated using the following equation:

$$Average\ Annual\ Savings = (\alpha_{02} * Post + \alpha_{12} * Average\ Daily\ CDD + \alpha_{22} * Average\ Daily\ HDD) * 365\ days$$

Average Daily CDD and Average Daily HDD were weighted according to the number of customers corresponding to each weather station[5].

In order to estimate the impact of the program during peak hours, we analyzed hourly consumption data. The model is defined by the following:

$$AECit = \alpha_{01} + \alpha_{04}PEAK_t + \alpha_{05}POST_{it} + \alpha_{08}PEAK_t * POST_{it} + E_{it}$$

Where:

- AEC_t is average electricity use during hour t for a participant i.
- $PEAK_t$ is an indicator variable (1 if during a defined peak hour or 0 otherwise) identifying whether or not hour t is one of the defined peak hours.
- $POST_t$ is an indicator variable (1 if during the post program period, 0 if prior to the program period) identifying whether or not the hour was during the post program period.
- E_t is an error term.

The coefficient of interest is α_{08}. The value of this coefficient estimates the average hourly demand reduction associated with the program

---

[5] The differences in weather stations is minor since all of SMUD is in one climate zone.

Results

The results of the study are presented as follows:

Baseline and Efficient Load Profile

An average hour usage curve for the pre-installation baseline and VSD pump conditions using the engineering approach are provided in Figure 4. As can be seen by the curve, on average the VSD pumps show a tighter distribution around their peak run time when compared to the baseline condition. This distribution is expected and supports a pattern of pumps running at a high speed for a short period of time with low speed usage for the remainder of operation. Savings profiles are presented in Figure 5.

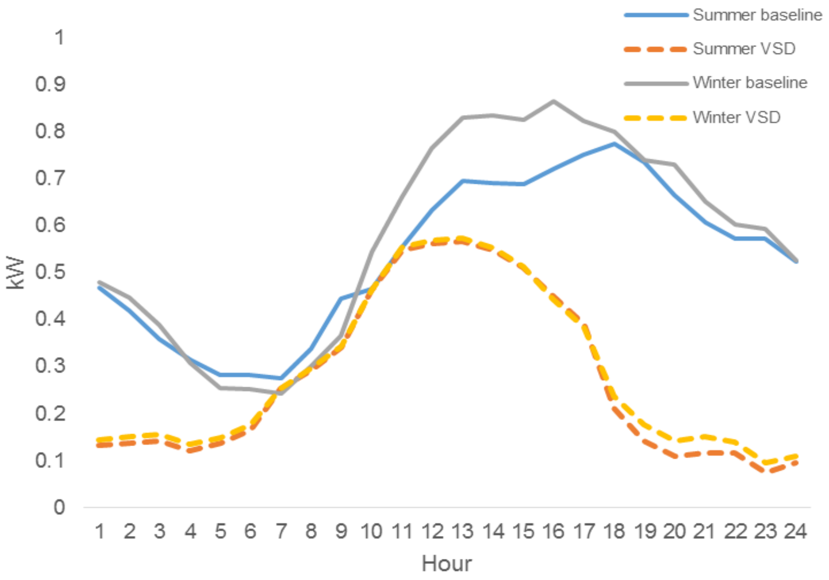

Figure 4. 24-hour average load profile for baseline and VSD pumps in summer and winter seasons

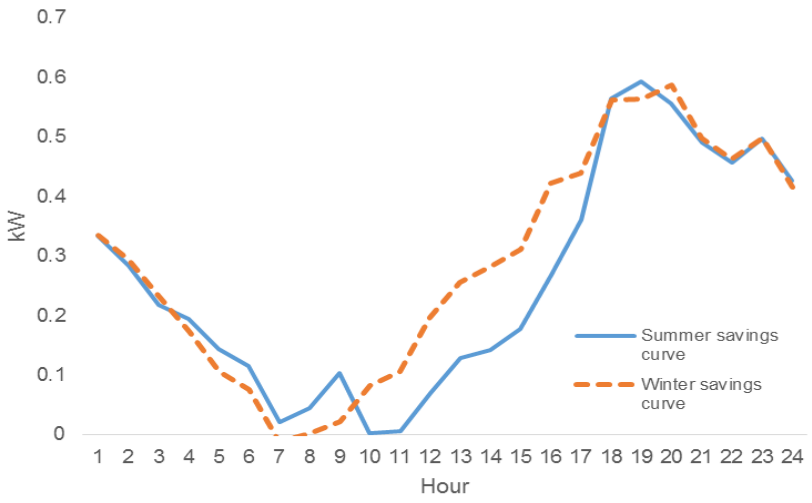

Figure 5. Average daily savings profile for VSD pumps versus baseline condition in summer and winter seasons

Gross Impact Savings

The average annual savings for single-speed pumps from the engineering analysis was 2,463 kWh and 0.51 kW. Savings calculated through the engineering approach show external validity when compared to findings from the *Savings Estimation Technical Reference Manual for the California Municipal Utilities Association* (ERS 2014).[6] The 2014 CA TRM provides estimated savings of 2,456 kWh/yr and 0.21 kW.

Table 4 provides calculated energy savings as generated through the econometric analysis. Unlike the engineering approach, the econometric analysis provided savings that were inconsistent and significantly lower than the 2014 CA TRM.

Table 4. Annual average energy and peak demand savings – econometric approach

Model	Savings	Std. error	95% CI	R-squared	P-value
Annual energy savings – econometric approach (kWh)	1,345	221	442	0.22	<0.01
Peak demand savings – econometric approach (kW)	0.102	0.017	0.033	0.03	<0.01

Participant Profile

Given the moderate to low fit of the regression analysis, ADM also examined whether some of the impacts could be attributed to other behavioral or qualitative changes occurred in the home that would impact the fit of the model. Sampled customers were asked if any significant changes to home energy use had occurred since participating in the program (e.g., changes to the home's structure, changes in the number of occupants, remodeling to the home, etc.). Participants described the types of changes that were made. Changes were then reviewed and grouped them based on similarity of response. Figure 6 provides a visualization of the results.

---

[6] The 2014 TRM provides generalizable estimated savings values for California's publicly owned utilities based on work papers and research throughout California. Guidance for VSD pool pumps is derived from work papers generated by SCE (Southern California Edison). Workbooks are provides by ERS which can be modified to accept location-specific inputs. The modification to Sacramento-specific inputs generate the savings referenced here.

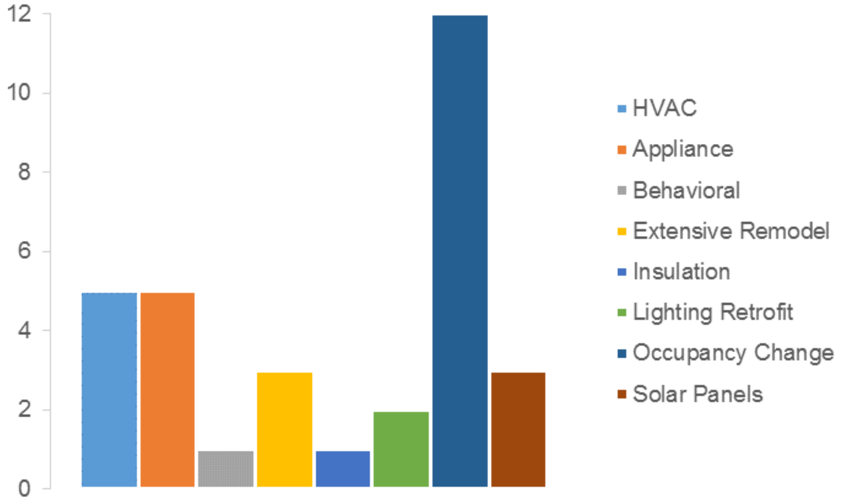

Figure 6. Reported changes to home energy use post-program participation

Thirty-two respondents (46%) reported some form of significant change to home energy use since participating in the program. The customer surveys provide evidence that suggests that changes to a customer's building profile occur in heavy succession to participation in the program. While some changes, such as cross-program participation, may increase savings predicted through regression, others may significantly mask savings, such as extensive home remodeling and changes in occupancy. These qualitative shifts will bias the explanatory power of a regression model and result in higher levels of error that will ultimately fail to represent savings attributable strictly to the program of interest.

Discussion

The study utilized a theoretical approach to supplement a simple engineering analysis of pool pumps. This theoretical approach for supplementing the engineering analysis was based on the idea that the end use of interest, single-speed pool pumps, run for an extended period of time with a consistent kW load. Because of the nature of residential loads, ADM was able to identify the rectangular load profile associated with pool pumps and extract it from the hourly interval meter data. The savings generated by an external data source was used to validate the end product of using the approach itself. Future research should specifically target how well these types of extractions compare sub-meter data, as this would further qualify when it is appropriate to utilize this method.

There are a multitude of different avenues for extrapolating the approach detailed in this paper to other applications. While this methodology is not a replacement for sub-metering or on-site monitoring, the approach can provide information in instances when traditional methods are either not cost effective or inappropriate for a given scenario. The ability to capture the energy intensity of the end use can only be accurately predicted if the energy intensity can be assumed to be "on" or "off" at a given time as opposed to partial loading. Furthermore, the building should exhibit some opportunity for other end uses to be off or operate on an inconsistent schedule. This study made use of low energy use days to accentuate the pump operation. By taking the lowest three usage days out of every month and averaging the 24-hour profile for these days together, the unscheduled loads were minimized while highlighting loads that were consistent and repetitive. The types of end uses that this method can be used on is specific to the resolution of the data. For this particular study, pool pumps are characterized by multiple hours of on-time with a flat load during that period of on-time. Other loads that are characterized by shorter periods of on-time would require finer resolution data for detection.

A potential avenue of exploration would be using this methodology for the disaggregation of commercial end uses. For example, a building that shuts down during the weekend would be a good candidate for the disaggregation of exterior lighting using this method. This approach would be able to accurately identify the energy intensity of exterior lighting for that building and, when mapped to sunrise and sunset information, determine whether the exterior lighting schedule corresponds to daylight hours or operates on a fixed schedule. Other residential appliances could be identified with finer time resolution interval meter data. The finer the time resolution the lower the energy intensity loads can be quantified and characterized.

Furthermore, the approach proposed in this paper may provide a worthwhile alternative to traditional regression analyses. Although the econometric analysis made use of an industry-standard weather-normalized model to predict average household savings, the findings from this approach were inconsistent both with the engineering analysis and the findings from previous work. Pre-/post- fixed effects regression designs are of particular use when the dependent variable's behavior is determined to be unvarying aside from the variable of interest. In this case, the regression analysis used for this evaluation is predicated on the idea that changes in a customer's whole building load are solely attributable to the implementation of the measure. The survey data showed significant correspondence between VSD pool/spa pump retrofits and large qualitative changes to the home. Unless this type of information is collected at a census-level resolution or enough participants are sampled to restrict statistical error at the independent variable level, the impact of these qualitative shifts on predicted savings remains inconclusive.

The proposed method for end use load disaggregation focuses on hourly interval meter data. Given this level of resolution, two-hour differences are used to isolate the full on-boarding and off-loading of end use loads instead of one-hour differences. This is because it cannot be assumed that loads on-board and off-load at the hour mark. Rather, one can anticipate that loads turn on and turn off during partial hours. If finer interval data were available, two-interval periods should be used to quantify the load change. The two-hour kW deltas were supplemented with single-hour kW deltas to attempt to reconstruct partial-hour schedules. As evaluators and utilities gain access to interval meter data at more granular levels of resolution, schedules detected using this approach can continue to be fine-tuned. We may also expect repeating, scheduled loads to be more prominent and easily distinguishable as they continue to reappear in interval meter data.

Conclusion

The method proposed in this study provides a starting point that is worthy of further research. Although there were limitations in the data available to conduct this study, the external corroboration of the results of the applied theoretical approach to the 2014 CA TRM provides evidence to suggest that the proposed method has merit and potential application. The findings also support the savings presented in the TRM. Traditional fixed effects regression provides evaluators insight into energy efficiency programs when pre-installation and post-installation information is not available. However, these methods can suffer from lack of interpretability when evaluators are unable to specify models that are reflective of the whole building's load profile. The method proposed in this study provides an intermediate solution that attempts to create load profiles when data is readily available to evaluators, but either budgetary or environmental constraints prevent evaluators from obtaining site-specific information either via on-site visit or sub-metering.

References

ERS (Energy & Resources Solutions). 2014. Savings Estimation Technical Reference Manual for the California Municipal Utilities Association. San Jose, CA: CMUA.

Worth, Ludovici, Joyce & Fernstrom. 2014. Pools & Spas: Codes and Standards Enhancement (CASE) Initiative For PY 2013: Title 20 Standards Development. Sacramento, CA: CEC.