ESL-IC-11-10-15

CASE STUDIES IN USING INTERVAL DATA ENERGY MODELS FOR SAVINGS VERIFICATION: LESSONS FROM THE GROCERY SECTOR.

Joan Effinger
Engineering Manager

Mark Effinger, P.E.
Engineer

Hannah Friedman, P.E.
Director

Portland Energy Conservation, Inc
Portland, OR USA

ABSTRACT

The use of whole building utility interval data for verifying energy savings from energy efficiency projects has become an attractive option as this data is increasingly available. Formal protocols, such as IPMVP Option C and ASHRAE Guideline 14, describe a whole building savings approach, but may require up to one full year of post-implementation data in order to claim annual energy savings. Many projects cannot absorb this long timeline.

This paper builds on previous research and investigates strategies to reduce the required post-implementation monitoring time. Five grocery energy efficiency projects were evaluated using whole building electric interval data to investigate how data resolution, monitoring period length and timing of the post-implementation monitoring period impact the accuracy of annualized savings estimates.

INTRODUCTION

The increasing availability of whole building utility interval data, through smart grid infrastructure or energy management systems, has made using this data an attractive option for verifying energy savings from energy efficiency programs. However, the approach for using interval meter data is not widely utilized, as timing constraints pose significant barriers for many projects. Strategies that reduce the required post-implementation monitoring length as well as evidence that demonstrates the impact on estimated savings are increasingly valuable.

Due to timing constraints, many energy efficiency projects apply a normalized savings approach using less than a full year of data. The normalized approach involves the creation of baseline and post-installation energy regressions from measured data. Both regressions are then driven by a common data set, such as TMY temperature (Reddy 2000). While there have been previous studies investigating the effect of data resolution and the impact of short term monitoring periods on savings accuracy, consensus on the "best" approach has not yet been achieved.

Previous studies, which focused on large commercial buildings, have suggested that whole building regressions created with daily data provided better predictions of annual energy consumption than hourly or hour-of-day (HOD) models (Katipamula, 1994). The validity of these models are often prescribed by presenting typical statistical indices such as the coefficient of determination (R^2), coefficient of variation of the root mean square error (CVRMSE) and mean bias error (MBE). However, when regressions are developed using less than one year of data, extrapolation error can be significant and is not captured by these indices (Haberl, 1997). Hence, a model with great "goodness of fit" may not accurately predict the desired annual energy consumption (Reddy, 2000).

One study has shown that models developed during the swing seasons produced the lowest average bias errors (Kissock, 1993). The conclusions from the Kissock study, based on a mix of three office and university buildings, suggest that proximity of the swing season's average temperature to the average annual temperature appears to influence the model's predictive capability. Other studies suggest that the most reliable results occur when regressions are developed using data that includes as much of the annual temperature range as possible (Montgomery, 1991). ASHRAE is currently working on RP-1404, which investigates the impact of short-term monitored data, so additional guidance should be available soon.

In the meantime, this paper describes the results from five grocery case studies using whole building interval data to determine electric savings for existing building commissioning (EBCx) projects. The case studies are used to examine how the duration of the post-installation monitoring period affects the accuracy of annualized savings estimates. A comparison of daily, hourly and hour of day models was conducted to evaluate the impact of data resolution on model quality and accuracy. The influence of seasonality during the monitoring period was also investigated in an attempt to identify an optimal timeframe when the monitoring period is less than one complete year. Uncertainty metrics, such as confidence intervals, are used to establish a framework for determining optimal monitoring timeframes to achieve acceptable saving estimates. The percent savings range that can be accurately verified using interval meter data is also discussed.

STATISTICAL METRICS

The best way to evaluate a whole building energy savings approach has been a topic of research for some time. In regression analysis, which is a major component of the whole building approach, a few standard statistical metrics are often used:

1. The coefficient of determination -R^2
2. The coefficient of variation of the root mean square error - CV(RMSE)
3. Mean bias error – MBE

These metrics can be found in any statistics text book, and there is ample discussion related to their use in a whole building savings approach in guidelines such as IPMVP-Volume 1 (2010) and in research papers such as Reddy (2000).

In general, these metrics are heavily used to evaluate the "goodness of fit" of a particular regression. While these metrics are used to evaluate the quality of regressions, they do not necessarily provide an indicator of accuracy when the regression is used to extrapolate beyond the range of data collected in the monitoring period (Haberl, et al, 1997).

In this paper, these statistical metrics are used to initially evaluate the quality of the energy regressions and identify the parameters most responsible for driving energy use at the five stores. However, it's recognized that poor regressions, defined by low R^2 and high CV(RMSE), may still produce "accurate" savings or at least savings that are good enough for some situations. Since many stakeholders may be concerned with accuracy more than precision or statements of statistical confidence, savings were calculated using a full year of baseline and post-installation monitoring. The full year savings was used as a basis of comparison to evaluate accuracy throughout this research.

SAVINGS ANALYSIS

This section outlines the process for determining the regression model specifications. Included is an analysis of possible driving variables and data resolution as well as an evaluation of what percent whole building savings can be validated using this approach. Annual electric savings are calculated using the avoided energy use method and are established as the "actual" savings throughout the paper. The normalized method is used to create annualized savings forming a comparison for shorter duration analysis in the next section.

The five stores included in this study were located in California. Two different grocery chains were represented and the store locations were split between coastal and inland climates.

Table 1 gives additional detail about each store including hours of operation, location, original project savings, and the percent whole building savings. It should be noted that the original project savings were calculated using a deemed savings approach based on DEER[1] (Database for Energy Efficient Resources). The original deemed savings ranged from 1.25% to 17.8%.

Driving Variables and Data Resolution

Regressions for total energy use prior to project implementation were developed from one full year of baseline data using both hourly and daily data. Separate regressions were created for dry bulb temperature, wet bulb temperature, and relative humidity to determine the parameter most responsible for driving energy use in the grocery stores included in this research. Each regression was evaluated using three standard statistical indices: coefficient of determination - R^2, coefficient of variation of the root mean square error - CV(RMSE) and mean bias error - MBE. A "best fit" regression is one that maximizes R^2 while minimizing both CV(RMSE) and MBE.

Relative humidity resulted in highly scattered regressions which indicate humidity is not a main

Table 1. Grocery Store Classification

Grocery Store	Chain 1			Chain 2	
	Store 1	Store 2	Store 3	Store 4	Store 5
Hours of operation	6 am – 11 pm	6 am – 11 pm	6 am – 11 pm	8 am – 10 pm	8 am – 10 pm
Location	Los Banos	Fresno	San Francisco	San Mateo	San Francisco
	Inland		Coastal		
Implementation date	June 2009	March 2009	October 2008	March 2009	January 2010
Deemed Savings (kWh)	357,750	113,893	199,686	38,000	30,400
% whole building savings	17.8%	4.5%	9.4%	1.9%	1.25%

[1] DEER info can be found at www.energy.ca.gov/deer.

driver for grocery store energy use. Dry bulb and wet bulb temperatures produced better fits which indicate they could be a main driver of energy use for grocery stores. There was no significant statistical difference between the regressions created with dry bulb or wet bulb temperatures. Since dry bulb temperatures are more easily obtained, they are used throughout this paper.

Store 4 and Store 5 had a significant difference between the occupied (8 am – 10 pm) and unoccupied (10 pm – 8 am) energy use (Figure 1). The red circle in Figure 1 is the unoccupied energy use and the black occupied energy use. For these stores, two hourly regressions (occupied & unoccupied) were developed using dry bulb as the driving variable.

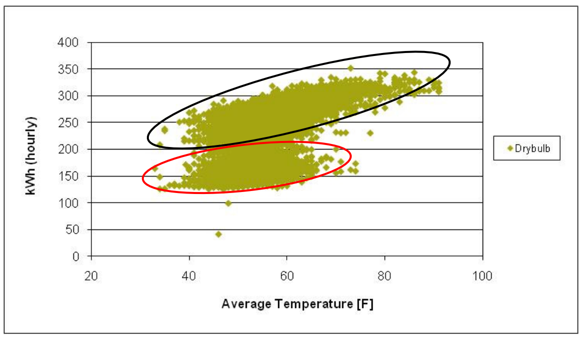

Figure 1. Store 4 Hourly Energy Use

The energy regressions of Store 1 and Store 2 were linear change-point models, as indicated by Figure 2 and described by ASHRAE Research Project 1050 (Kissock, 2003). Energy Explorer, a software tool developed by Kelly Kissock from the University of Dayton, Ohio was used to develop and analyze the hourly and daily change-point regressions for these stores.

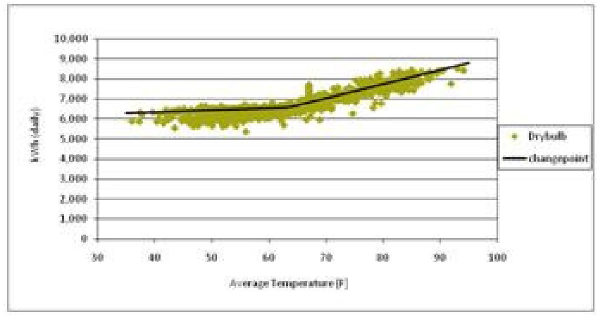

Figure 2. Example of Change Point

Table 2 and Table 3 summarize the statistical indices of the regressions for hourly and daily data, respectively, using dry bulb temperatures only. Since the variation in energy use experienced from hour to hour is averaged across the day, the daily regressions have significantly higher R^2 while CV(RMSE) is reduced by about half. Thus for all stores, regressions created with daily average data result in a better "fit" than regressions made with hourly data. This aligns with the International Performance Measurement and Verification Protocol (IPMVP) recommendation that hourly data be rolled into daily values and asserts the loss of resolution should not significantly increase the uncertainty of the results (IMPVP, 2009).

<u>Avoided Energy Use</u>

The avoided energy use method involves the creation of a baseline regression using a full year of monitored data prior to implementation. Then the actual outside air temperatures recorded during the

Table 2. Statistical Metrics for the Hourly Regression Models

Statistical Index	Store 1	Store 2	Store 3	Store 4: Unoccupied	Store 4: Occupied	Store 5: Unoccupied	Store 5: Occupied
Model used	Linear Change point	Linear Change point	Linear regression	Linear regression	Linear regression	Linear regression	Linear regression
R2	0.79	0.75	0.24	0.03	0.41	0.019	0.18
CV(RMSE)	5.3%	6.1%	4.6%	25.1%	8.6%	25.71%	9.02%
MBE	0%	0%	0%	0%	0%	0%	0%

Table 3. Statistical Metrics for the Daily Regression Models

Statistical Index	Store 1	Store 2	Store 3	Store 4	Store 5
Model Used	Linear change point	Linear change point	Linear regression	Linear regression	Linear regression
R2	0.92	0.94	0.25	0.6	0.35
CV(RMSE)	2.8%	2.7%	3.1%	3.9%	3.15%
MBE	-0.005%	0.1%	0%	0%	0%

year following the conclusion of energy efficiency implementation were used in the regressions to create an adjusted baseline for each store. The adjusted baseline is a prediction of how the building would have operated if the energy efficient change was not implemented. The difference between the adjusted baseline and measured post-installation energy use is the avoided energy use, or the closest attempt to "measure" energy savings (Equation 1). The avoided energy use is used as the basis of comparison when evaluating the predictive accuracy of regressions developed with less than 1 year of data later in the research.

Electric savings for each store was calculated using both hourly and daily data driven by dry bulb temperatures, but as the daily data was previously shown to be a better fit, only daily data is reported in Table 4.

Equation 1

Energy Savings =
Adjusted baseline – Post installation Energy Use

Comparing the savings for each store in Table 4 to the deemed project savings shows the three stores in Chain 1 used more energy (less savings) and the two stores in Chain 2 used less energy (more savings) than originally calculated. It is important to note that the original deemed savings were determined using a measure by measure approach. The avoided energy use is a whole building approach that captures everything occurring downstream from the main utility meter. All interactions between individual measures as well as other energy influencing changes to the building or its operation during the monitoring period will influence the savings determined using this approach. The differences between the deemed savings and the whole building approach illustrate the importance of tracking any changes that occur in the building during the monitoring timeframe (both pre and post). This is especially true when the attribution of savings to a particular measure, project or program is required by the stakeholders.

Smallest Percent Savings Detectable Using a Whole-Building Approach

One of the main barriers preventing the adoption of a whole building approach is the understanding of what percent whole building savings is required in order to accurately verify savings. While there has been discussion in the industry, consensus has not been reached. ASHRAE Guideline 14 states savings greater than 10% of whole building consumption are required, but this number is based on the historical approach of using monthly utility data. Other research implies that savings of at least 5% of building consumption can be detected when greater resolution data, such as interval meter data, is used to create the energy regressions (Katipamula, 1994).

In order to have high confidence that a project has achieved the estimated energy savings, the actual energy use in the post-implementation period should be statistically different from the adjusted baseline. In this application, this means that the actual energy

Table 4. Avoided Energy Use Results*

Store	Avoided Energy Use (kWh)	% whole building savings	Deemed Savings
Store 1	285,630	14.3%	17.8%
Store 2	94,597	3.8%	4.5%
Store 3	106,747	5.3%	9.4%
Store 4	101,908	5.0%	1.9%
Store 5	166,374	6.9%	1.25%

*The whole building results are used as the basis of comparison for the remainder of this paper.

use in the post-implementation period should be outside the uncertainty bands of the adjusted baseline. A visual way to depict this is shown in Figure 3, where the uncertainty in the adjusted baseline is depicted with dotted lines. The uncertainty was calculated according to Equation 2 and at the 90% confidence level.

Equation 2

$$avoided\ energy\ use\ \pm SE * z\ score$$

Where,
SE = Standard Error of the regression
z score= The critical value of a distribution based on the degrees of freedom and significance level.

In Figure 3, the actual post-implementation energy use, shown in green, is clearly outside the uncertainty band of the adjusted baseline. Thus, the avoided energy use method can be used with statistical confidence in this case. However, in Figure 4 the actual energy use overlaps with the uncertainty in the adjusted baseline. In this case, the savings is not statistically different from the adjusted energy use and our confidence that the estimated energy savings was achieved is low.

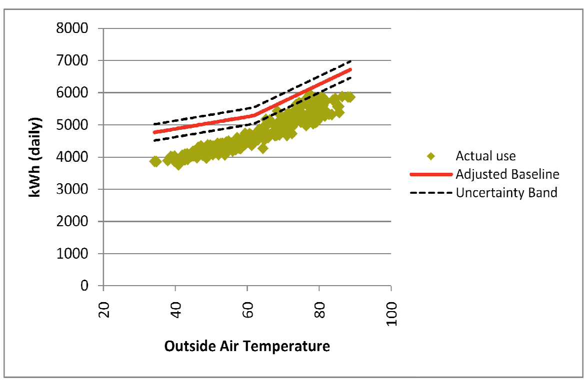

Figure 3. Store with a statistical difference between the adjusted baseline and post-installation data

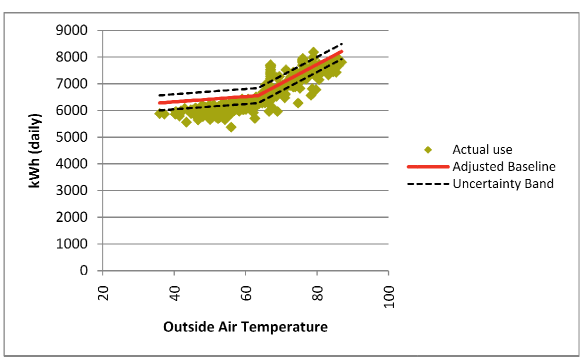

Figure 4. Store without a statistical difference between the adjusted baseline and post-installation data

Figures like these were created for each store but are not shown. Store 1 and Store 5 savings are outside the uncertainty of the adjusted baseline, thus can be confidently verified using this approach. When calculated using a 90% confidence interval, Store 3 and Store 4 has some overlap with the uncertainty of the adjusted baseline.

Not all projects or programs require such a high level of statistical confidence in the final savings. If a lower confidence interval is acceptable, such as 80%, the uncertainty band around the baseline narrows. Narrow uncertainty bands make it easier for the actual post-installation energy use to fall outside the bands, thus making the adjusted baseline and post-installation data statistically different. At the 80% confidence level, Store 3 and Store 4 can be confidently verified.

It's important to note that the statistical difference is a statement of confidence, not necessarily an indicator of the accuracy of the final savings estimate. Store 2 provides a fairly clear example related to the confidence versus accuracy issue. Note that the variation, CV(RMSE), of Store 2's daily baseline model is approximately 3% (Table 3). The savings from this particular store using the avoided energy use method falls just under 4%, which is barely larger than the variation of the model. As shown by Figure 4, the post-implementation energy use falls within the uncertainty bands. Therefore, there is no statistically significant difference between the adjusted baseline and post-implementation and no statistical confidence in the whole building savings for this project. However, the whole building savings approach estimated 94,597 kWh of savings which is fairly close to the original measure by measure deemed savings of 113,893 kWh. While we can't state with certainty Store 2's savings are statistically valid, the savings estimates may be "good enough" for many project stakeholders.

Specific project requirements will vary greatly by stakeholders, and many stakeholders will not require the high level of certainty presented in the preceding section. For example, ASHRAE Guideline 14 states for whole building savings greater than 10% the baseline model should have a CV(RMSE) of less than 20% (ASHRAE, 2002). While only one of these stores had savings above 10%, the CV(RMSE) for all stores are considerably lower than 20% (Table 3). Furthermore, the Regional Technical Forum (RTF)[2] stipulates the estimated energy savings must be ± 20% of the actual savings (Regional Technical

[2] The RTF is an advisory committee in the Pacific Northwest charged with developing standards to verify and evaluate conservation savings.

Forum, 2011). Table 5 shows that the uncertainty of the savings for both the hourly and daily models is significantly below this threshold for these five case studies when using a full year of data.

Therefore, when adopting a whole building approach, the expected percent whole building savings, the amount of variation present in the baseline, and the desired accuracy of the results should be considered. Stores with higher (>10%) percent whole building savings can achieve relatively high certainty with either hourly or daily data. Stores with lower percent whole building savings that require high certainty may need another approach. Hour of Day (HOD) models are one possible strategy to improve the certainty of lower savings projects.

Hour-of-Day Models

Hour-of-day models involve sorting the data from each monitoring period by the specific hour, then creating a separate energy regression for each hour. This process can quickly become quite cumbersome as it results in a minimum of 48 separate regressions when comparing baseline and post-installation operation. However, HOD models have the potential to account for variability in building loads, such as those created by occupancy (Katipamula, 1994).

HOD models were created for Stores 3, 4, and 5 using linear regressions. The combined standard error was calculated according to procedures described in IPMVP Appendix B (2009) and Equation 3.

Equation 3

$$SE_{total} = \sqrt{N} * \sqrt{SE_{1:00}{}^{2} + SE_{2:00}{}^{2} + \ldots + SE_{24:00}{}^{2}}$$

Where:

$$SE = \sqrt{\frac{\sum_{n} \left(Y_i - \hat{Y}\right)^2}{n - p}}$$

N = the number of savings results with the same Standard Error that are added together

p = number of independent variables in the regression equation

n = the number of samples

Table 6 illustrates the significant reduction in the uncertainty achieved by using HOD models. This reduction in uncertainty tightens the error bands around the adjusted baseline, further reducing the amount of overlap between the actual and adjusted baseline energy use. Using HOD models in Store 3 and Store 4 allows the savings to be more confidently verified.

Table 6. Hourly, Daily, and HOD error

	Store 3	Store 4	Store 5
Hourly Error (kWh)	1,045	3,262	4,018
Daily Error (kWh)	3,408	4,222	3,999
HOD Error (kWh)	848	1,325	1,490

While hour-of-day models improved the precision of the regressions, HOD models did not improve the accuracy of the savings. Table 7 shows that the HOD models predicted roughly the same avoided energy use as both the hourly and daily models. The CV(RMSE) was calculated for all 24 regressions for each store, but are not shown. In all five cases, the CV(RMSE) for the HOD models was better than hourly models and close to the CV(RMSE) for the daily models. In these case studies, the HOD models improve the confidence in the savings but do not ensure more accurate results.

Table 7. Comparison of Annual Savings Calculated from Hourly, Daily, and HOD Models

	Store 3	Store 4	Store 5
Hourly Avoided Energy Use (kWh)	107,018	102,038	169,112
Daily Avoided Energy Use (kWh)	106,747	101,908	166,374
HOD Avoided Energy Use (kWh)	106,680	101,576	166,167

Table 5. Uncertainty in the Avoided Savings for each Store

	Store 1	Store 2	Store 3	Store 4	Store 5
Avoided Savings (kWh)	285,630	94,597	106,747	101,908	166,374
Hourly Error (kWh)	1,145	1,655	1,045	3,262	4,018
Daily Error (kWh)	3,164	3,321	3,408	4,222	3,999

Normalized Savings

Normalized savings are calculated using separate regressions for the baseline and post-installation periods. Each regression is then driven with a common dataset, such as TMY temperature data. Formal procedures, such as IPMVP, require monitoring for the duration of claimed savings. As such, annual savings require one year of post-installation monitoring.

The intent of this study is to create normalized savings using less than one year of data to ascertain how much the monitoring time can be reduced while still achieving accurate savings.

For each store, both hourly and daily regressions using one full year of post implementation data were developed for total energy use based on dry bulb temperatures. Daily average dry bulb temperatures obtained from TMY3 were used to drive both the baseline and the post implementation models. The results are shown in Table 8. Uncertainty resulting from both the baseline and post-implementation regressions was combined in the savings analysis (Effinger, et al, 2008).

The avoided energy use method is driven by actual measured values during the post monitoring period while the normalized method uses TMY3 temperatures, which are averaged over several years. The slight difference observed between the final savings values of the normalized method and avoided energy use method is expected.

Table 8. Normalized Savings Results

Store	Savings (kWh)	% whole building savings
Store 1	278,843	14.0%
Store 2	96,648	3.8%
Store 3	105,989	5.0%
Store 4	101,873	5.0%
Store 5	168,294	6.9%

OBSERVATIONS

Shorter Duration Monitoring Period

Regressions using 9 months, 6 months, and 3 months of data after each store's implementation date were developed. Annualized savings for each store were calculated using the regressions created with the shorter duration post-installation data. The baseline regressions all used the full 12 months of pre-EBCx data.

Figure 5 shows that as the duration of the monitoring period decreases, in general, the accuracy of the savings also decreases. The stores with higher percent whole building savings (Store 1 and Store 5) can have shorter monitoring periods and still produce savings relatively close to the avoided energy use benchmark. However, as the percent whole building savings decreases (Store 3 and Store 4) a longer monitoring period is required to produce accurate savings. Thus, depending on the desired level of accuracy and the amount of whole building savings, 9 months and 6 months of monitored data could suffice. In most cases, the accuracy in the savings for 3 months of monitoring would be too low to produce acceptable savings.

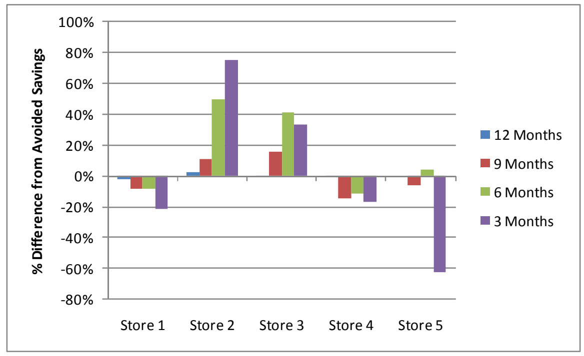

Figure 5. Shorter Duration Savings Results.

The implementation date appeared to affect the accuracy of annualized savings. For instance, Store 3 had an implementation date in October. The shorter duration monitoring, in this case, occurred over the winter and consistently over predicted energy use as compared to avoided savings. However, Store 4 had an implementation date in March with shorter duration monitoring occurring over the summer. In this case, energy use was under predicted as compared to the avoided savings.

As shown in Figure 5, the savings for Store 2 appeared significantly different than the avoided energy use method when monitoring periods were reduced to six months or less. Investigation into these differences yielded an interesting finding. The baseline data for this store was represented by a 4-parameter change point model. Neither the 6-month nor the 3-month regressions had an identifiable change-point and were instead described by a basic linear regression. The comparison of a linear post-model with a change-point baseline model artificially inflated the savings and the amount of error. An important conclusion from the Store 2 analysis is when a building operating profile follows a change-point in the baseline, enough data should be collected in the post-installation period to identify the post implementation change-point.

Framework for Optimal Monitoring Period

To establish a framework for determining optimal monitoring timeframes, post-installation regression models using the data from three month seasons and six month combination of seasons (Table 9) were developed for each store. With each of these eight season's regression models, annualized savings were calculated using the normalized savings method. The annualized savings were then compared to the avoided savings according to Equation 4.

Table 9. Definition of Seasons

Season	Months
Summer	June – August
Winter	December – February
Fall	September - November
Spring	March – May
Spring - Summer	March – August
Summer – Fall	June – November
Fall – Winter	September – February
Winter – Spring	December - May

Equation 4

$$\%\ difference = \frac{normalized - avoided}{avoided}$$

We evaluated the seasonality regressions using low desired accuracy, defined as ± 20% difference in Equation 4, and high desired accuracy, defined as ±10% difference in Equation 4. The results are shown in Figure 6 through Figure 10. In these Figures, the avoided energy use is the red square on the left and the green dots represent each season's annualized savings. The black horizontal lines are 20% from the avoided energy use and the yellow horizontal lines are 10% from the avoided energy use. Even though Store 2 did not have statistically significant savings, there were several monitoring periods where the results could be considered "good enough" based on this analysis.

Seasons whose dots fall within the horizontal lines correspond to seasons where the annualized savings produce acceptable results. As expected, fewer seasons' annualized savings fall within the low accuracy (±10%) horizontal bars.

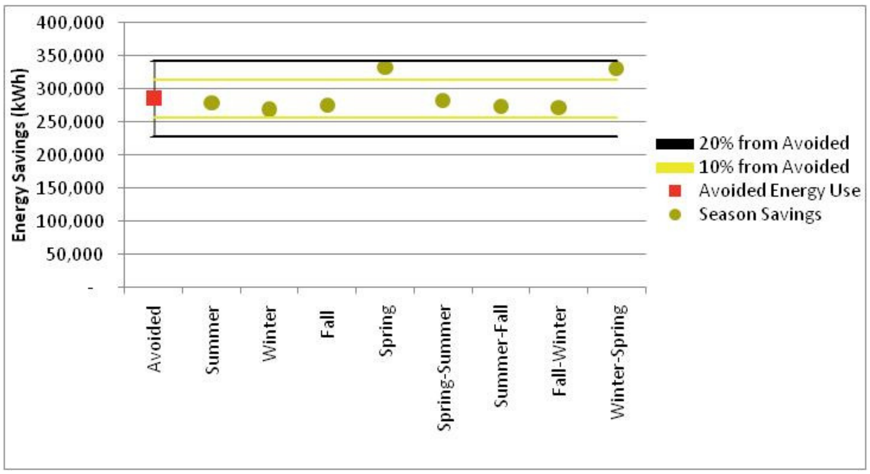

Figure 6. Store 1 Seasonality Results

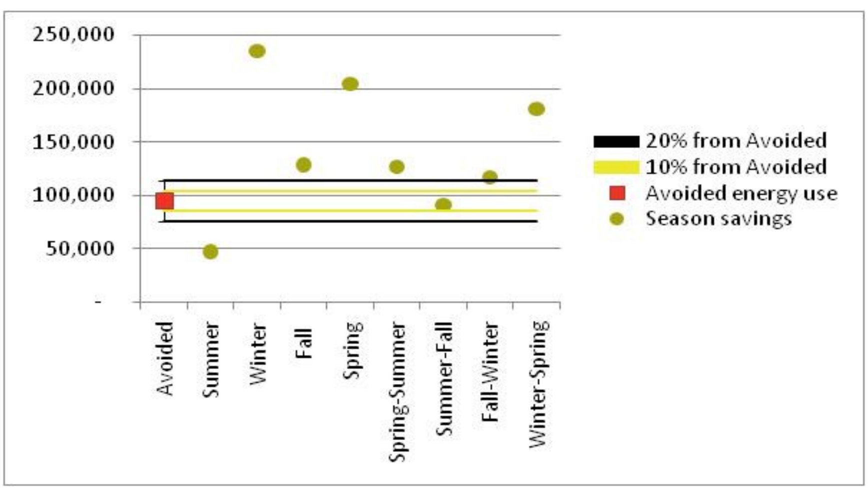

Figure 7. Store 2 Seasonality Results

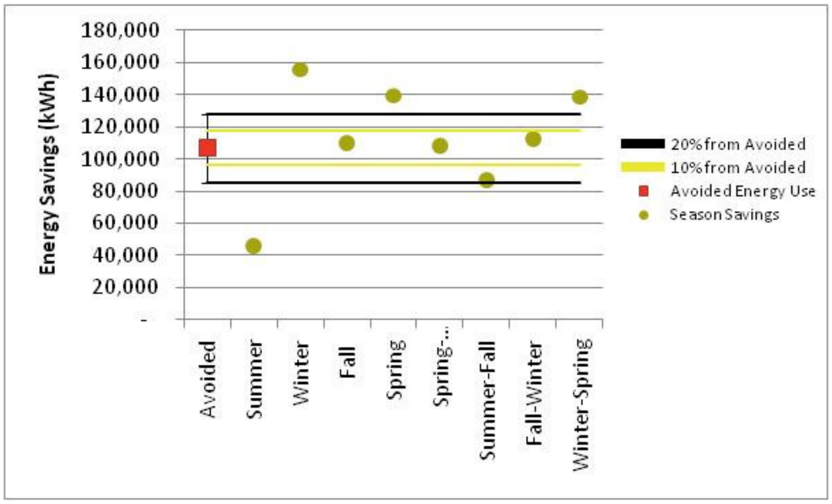

Figure 8. Store 3 Seasonality Results

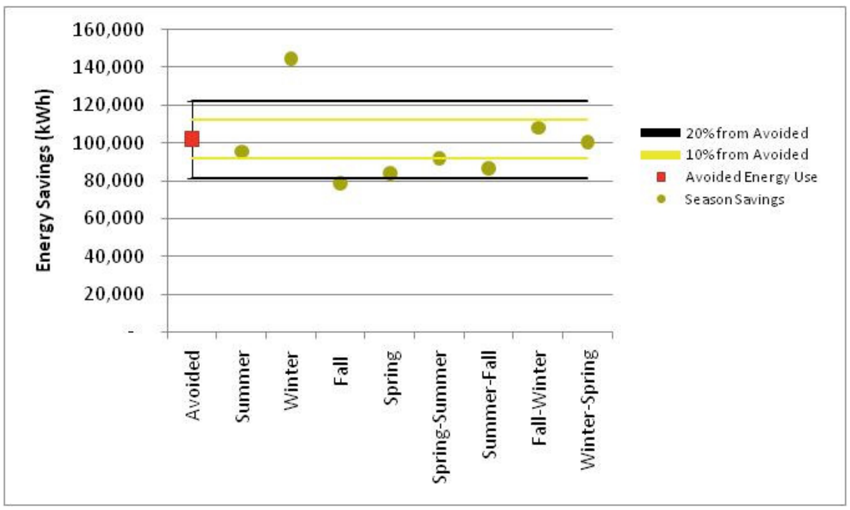

Figure 9. Store 4 Seasonality Results

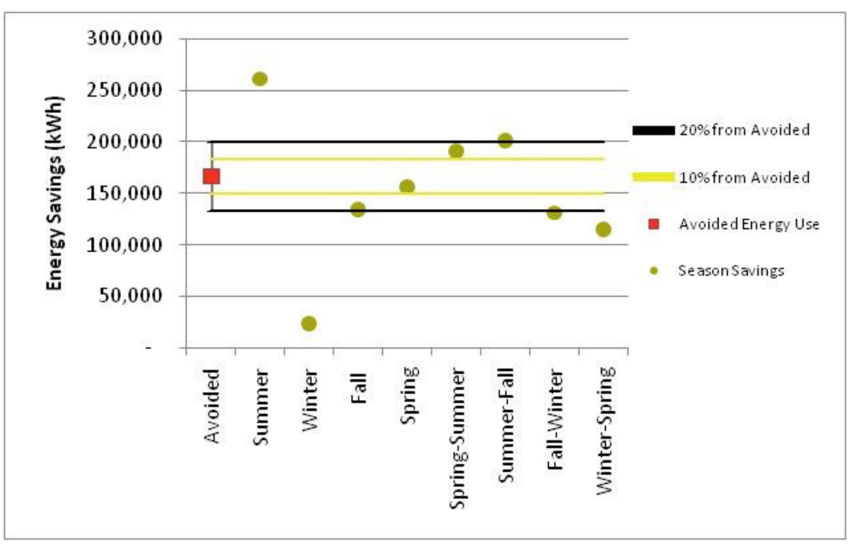

Figure 10. Store 5 Seasonality Results

To compare results across stores, Table 10 is a matrix of the stores and seasons. Seasons that fall within 20% of the stores' avoided savings are shaded green, while seasons that are marginal are shaded yellow. The first thing to note is that for Store 2 virtually no three or six month monitoring periods produced acceptable results. In cases like this, where savings are low, nine months or a full year of monitoring would be recommended. Looking across the other stores, the sixth month period of Summer – Fall produced the best results. The six month periods of Spring – Summer and Fall – Winter also produced acceptable results. Several three month monitoring periods occasionally produce acceptable accuracy but the extreme seasons of Summer and Winter were not as likely as swing seasons to meet the accuracy requirements.

Table 11 is the same matrix for higher accuracy. Seasons that fall within 10% of the stores' avoided savings are shaded green, while seasons that are marginal are shaded yellow. To meet this higher accuracy requirement most of these stores need six months of post monitored data. The six month period of Spring – Summer produced the most accurate results.

No matter what accuracy is desired, the higher the actual savings resulting from the project, the less impact duration and timing of the monitoring period appears to have on the optimal monitoring timeframe. For the projects with lower percent whole building savings, if higher accuracy is desired, a longer monitoring period may be required to produce acceptable results unless more energy savings can be achieved through additional implemented measures.

Temperature Analysis

The results from the seasonality investigation were analyzed in an attempt to identify measureable parameters that clearly indicate how much data is required to produce accurate results. First, we tested the suggestion from Kissock that proximity of the data set's average temperature to the average annual temperature appears to influence the model's predictive capability. For each store, the Spring and the combination of Spring – Summer seasons had average temperatures closest to the average annual temperatures. However, for the five case studies in this analysis, the Fall and Spring – Summer seasons predicted savings that were closest to the actual avoided energy use. Thus, the proximity of the data set's average temperature to the average annual temperature does not appear to predict the best results

Table 10. Lower accuracy (± 20 %) seasonality results

	Store 1: Los Banos 14.3%	Store 2: Fresno 3.8%	Store 3: San Francisco 5.3%	Store 4: San Mateo 5.0%	Store 5: San Francisco 6.9%
Summer	2.4%	-50.0%	56.8%	6.3%	-56.7%
Winter	5.9%	148.6%	-45.8%	-41.8%	85.8%
Fall	3.6%	35.6%	-2.9%	23.0%	19.5%
Spring	-16.3%	116.0%	-30.6%	17.4%	6.2%
Spring-Summer	1.0%	34.0%	-1.4%	9.9%	-14.9%
Summer-Fall	4.1%	-3.8%	18.6%	14.8%	-20.7%
Fall-Winter	5.0%	23.5%	-5.6%	-6.0%	21.3%
Winter-Spring	-15.8%	91.1%	-29.8%	1.4%	30.6%

Table 11. Higher accuracy (± 10%) seasonality results

	Store 1: Los Banos 14.3%	Store 2: Fresno 3.8%	Store 3: San Francisco 5.3%	Store 4: San Mateo 5.0%	Store 5: San Francisco 6.9%
Summer	2.4%	-50.0%	56.8%	6.3%	-56.7%
Winter	5.9%	148.6%	-45.8%	-41.8%	85.8%
Fall	3.6%	35.6%	-2.9%	23.0%	19.5%
Spring	-16.3%	116.0%	-30.6%	17.4%	6.2%
Spring-Summer	1.0%	34.0%	-1.4%	9.9%	-14.9%
Summer-Fall	4.1%	-3.8%	18.6%	14.8%	-20.7%
Fall-Winter	5.0%	23.5%	-5.6%	-6.0%	21.3%
Winter-Spring	-15.8%	91.1%	-29.8%	1.4%	30.6%

on its own.

Next, the range of temperatures experienced in each season was evaluated, as suggested by Montgomery (1991). Fall and all of the six month periods captured a majority (>75%) of the annual temperature range. This does have some correlation with Fall and Spring – Summer seasons producing results closest to the actual avoided energy savings. However, the range of temperatures alone also does not appear to predict the best results, as savings predicted by Winter – Spring were often very different from actual avoided energy savings.

While definitive conclusions should not be drawn from such a small sample set, it appears within these case studies that a combination of the proximity of average temperature and the range of temperature experienced in the monitoring period might predict the best results. For instance, while the average Spring temperatures were always close to the average TMY temperatures, Spring almost never covers the majority of the typical annual temperature range. The converse is true with Winter-Spring. The average period is never closest to the average TMY, but always included the majority of the range. The best predictive season, Spring – Summer, has average temperatures closest to the average TMY and includes the majority of the range of temperatures.

CONCLUSIONS

Based on the research from these five case studies, we were able to make general conclusions for the grocery sector in the following areas:

- Key driving variables for regression analysis
- Best fit model types
- Percent savings that are statistically valid to use a whole building analysis approach
- Determining post-period duration and timing

The analysis of dry bulb temperature, wet bulb temperature, and relative humidity as possible driving variables for grocery stores showed that either wet bulb or dry bulb temperatures could be used as the primary driving variable. Since it's easier to obtain dry bulb than wet bulb temperature, dry bulb temperature was used in this analysis. Wet-bulb may play a greater roll in more humid climates, or when the refrigeration equipment is water-cooled instead of air-cooled.

The evaluation of hourly, daily, and HOD models showed daily models produce better regressions than hourly models without a significant increase in uncertainty of the savings. Compared with hourly and daily models, HOD models were shown to reduce the uncertainty in the savings but did not result in higher accuracy.

The proximity of the average temperature of the shorter data set was compared to the average temperature of TMY data to further evaluate the theories that close proximity would produce the best predictors. However, for these case studies, that was not the case. It appears that some interaction between being near the average annual temperature and capturing a full range of TMY temperature is important, although more case studies should be conducted to make definitive conclusions.

Previous guidelines and existing research generally stipulate a whole building approach be used only in projects where percent savings are greater than 10%. One store had 14% savings, which were statistically significant at 90% confidence. Three stores had savings around 5%, which were statistically significant at 80% confidence. One store had lower savings (3.8%) which did not meet statistical confidence levels. However, as compared to original deemed savings, the savings estimated by the whole building approach were fairly accurate.

The case studies analyzed in this paper showed a strong interdependency between the percent whole building savings and desired accuracy on the optimal monitoring timeframes to achieve acceptable savings estimates. For instance, for projects with higher percent whole building savings (~10%) that desire lower accuracy (±20%), any season produced acceptable savings. However, if higher accuracy (±10%) is desired the best monitoring period is the six month period between March and August. Some three month seasons (Summer, Winter, and Fall) and other six month seasons (Summer – Fall and Fall – Winter) produce acceptable savings.

Considering monitoring duration and timing is more important when analyzing projects with lower percent whole building savings (~5%). Shorter duration monitoring is possible if lower accuracy is acceptable. The best monitoring period is the six month period between March and August. Swing seasons (Spring and Fall) also produced acceptable results. However, if higher accuracy is required, the six month period of Spring – Summer (March to August) is the only time that consistently produced acceptable savings across all stores. If shorter monitoring duration and higher accuracy is desired, it is recommended that additional energy efficiency measures be installed to increase the percent whole building savings so the results can be statistically valid.

This study made progress towards examining the feasibility of whole building energy savings verification using interval meter data for the grocery sector. For the grocery sector, as well as other building sectors, wide-spread adoption of the whole building approach for savings verification will require more evidence that demonstrates how much data is required and when the data should be collected to produce sufficiently accurate results.

More data from additional climate zones would be required to make universal conclusions. The climate zones captured in this study included mild coastal regions and hotter inland regions of California. It's unknown whether the observations in this research would apply to climates with more extreme conditions.

The case studies and overarching analysis presented here gives reason to believe that, with proper model specification and statistical analysis, this approach can gain significant traction for cost-effectively verifying savings in commercial buildings.

REFERENCES

ASHRAE 2002. ASHRAE Guideline 14-2002. Measurement of Energy and Demand Savings. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

California Commissioning Collaborative (CCC), 2008. Using Interval Data Energy Models: IPMVP Options B and C. Guidelines for Verifying Existing Building Commissioning Project Savings.

Effinger, M., Anthony, J., Webster, L., 2009. Case Stuides in Using Whole Building Interval Data to Determine Annualized Electrical Savings. ICEBO Conference, 2009.

Fels, M.F. (Ed.) 1986. Special Issue devoted to Measure Engergy Savings, The Princeton Scorkeeping Method (PRISM). Energy and Buildings 9(1) and 9(2).

Haberl, J.A., Reddy, A., Elleson, J. 1997. Determining Long-Term Performance of Cool Storage Systems from Short-Term Tests. ASHRAE Reseasrch Project 1004.

IPMVP. 2010. International Performance Measurement and Verification Protocol, Volume 1. Efficiency Valuation Organization.

Katipamula, S., Reddy, T.A., Claridge, D.E., 1994. Effect of Time Resolution on Statistical Modeling of Cooling Energy Use in Large Commercial Buildings. ASHRAE Transactions.

Kissock, J.K., Reddy, T.A., Fletcher, D., Claridge, D.E., 1993. The Effect of Short Data Periods on the Annual Prediction Accuracy of Temperature-Dependent regression Models of Commercial Building Energy Use. Joint Solar Engineering Conference, ASME 1993.

Kissock, J.K., Haberl, J.K., Claridge, D.E., 2003. Inverse Modeling Toolkit: Numerical Algorithms. ASHRAE Transactions, Volume 109, Part 2.

Katipamula, S., Reddy, T.A., Claridge, D.E., 1995. Bias in Predicting Annual Energy Use in Commercial Buildings with Regression Models Developed from Short Data Sets. Joint Solar Engineering Conference, 1995.

Montgomery, D.C, 1991. Design and Analysis of Experiments, Third Edition, John Wiley and Sons, New York.

Reddy, T.A., Claridge, D.E., 2000. Uncertainty of "Measured" Energy Savings from Statistical Baseline Models. HVAC&R Research

Regional Technical Forum, 2011. Guidelines for the Development and Maintenance of RTF Savings Estimation Methods.