LBNL-6874E

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Review of Prior Commercial Building Energy Efficiency Retrofit Evaluation: A Report to Snohomish Public Utility District

Phillip N. Price

Energy Technologies Area

April, 2015

This work was funded by Snohomish Public Utility District and was supported by the U.S. Department of Energy under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231

Review of Prior Commercial Building Energy Efficiency Retrofit Evaluation: A Report to Snohomish Public Utility District

Phillip N Price, Lawrence Berkeley National Laboratory
November 1, 2014

I. Introduction

Snohomish County Public Utility District ("the District" or "Snohomish PUD") provides electricity to about 325,000 customers in Snohomish County, Washington. The District has an incentive programs to encourage commercial customers to improve energy efficiency: the District partially reimburses the cost of approved retrofits if they provide a level of energy performance improvement that is specified by contract. In 2013 the District contracted with Lawrence Berkeley National Laboratory to provide a third-party review of the Monitoring and Verification (M&V) practices the District uses to evaluate whether companies are meeting their contractual obligations. This work helps LBNL understand the challenges faced by real-world practitioners of M&V of energy savings, and builds on a body of related work such as Price et al. (2013).

The District selected a typical project for which they had already performed an evaluation. The present report includes the District's original evaluation as well as LBNL's review of their approach. The review is based on the document itself; on investigation of the load data and outdoor air temperature data from the building evaluated in the document; and on phone discussions with Bill Harris of the Snohomish County Public Utility District. We will call the building studied in the document the "subject building," the original Snohomish PUD report will be referred to as the "Evaluation," and this discussion by LBNL is called the "Review."

II. Snohomish County PUD Evaluation of an Energy Retrofit Project

This section contains the "Evaluation" that was prepared by Snohomish PUD, beginning on the following page. It has been reformatted to fit in this document, and some clarifying notes have been added [in brackets] but it is otherwise unchanged. In the Evaluation, "I" refers to Bill Harris of Snohomish PUD, and "we" refers to Snohomish PUD overall. The Evaluation is typeset in a sans serif font so that it is more easily distinguished from the rest of this report.

Canyon Park Heights Commercial Building Evaluation
Bill Harris
2012-04-06

1. Evaluation Request

The responsible energy efficiency engineer requested that I [Bill Harris of Snohomish PUD] evaluate the savings from a project in the Canyon Park Heights building in Bothell, Washington to determine if the incentive payment should be made:

From a statistical analysis, we would like to confirm that using an extrapolation of the data available from PUD load data that a minimum of 546,156 kwh/yr will likely be realized. Your independent assessment will be welcome and appreciated.

The referenced savings correlates roughly to an annualized 20% education in energy use from the baseline, or an annualized average building load reduction of just under 63 KW (24/7).

With the statistical analysis of the 15 min. interval data, we would like to confirm that contract satisfaction has been met sufficient to make full payment on this not-to-exceed contract:

Measurement & Verification (M&V)
This HVAC retrofit proposes to reduce the building annual energy use by 20% compared to annualized energy use in CY 2010. PUD will analyze MV 90 data and compare to pre-Baseline energy use.

Meter #	Annual kWh	Notes
129805	2,241,600	kwh/yr registered in 12 months ending Mar 2012 billing)
	2,393,100	kwh/yr registered in 12 months ending CY 2011 (Dec billing)
	2,626,500	kwh/yr registered in 12 months ending April 2011
BASELINE:	**2,730,900**	kwh/yr registered in 12 months ending CY 2010
not full year	2,302,000	kwh/yr registered in **10 months** ending June 2010
Previous ownership	3,042,000	kwh/yr registered in CY 2008 (previous ownership)

	Authorized
Date:	05/05/11
Signed Off by:	
Total Project Cost:	$273,029.00
Total Incentive Amount:	$163,846.80
Total kWh/Yr Savings:	546,156

2. Summary

Based on my analysis of the MV-90 data on this facility and subject to the limitations listed below, the data supports a best-estimate savings of ~22% or ~570,207 kWh, which meets and slightly exceeds the anticipated reduction. There is less than a 2% probability, conditional on the data and the statistical model, that we did not attain the contracted savings.

3. Analysis

3.1 Data

This analysis is based on 15-minute MV-90 interval data measured in kW. After discussion with the engineer, we decided to use post data from October 1, 2011, after the September date when the majority of the retrofit had been engaged, through March 16, 2012, the last day of the data at the time the data analyst pulled it.

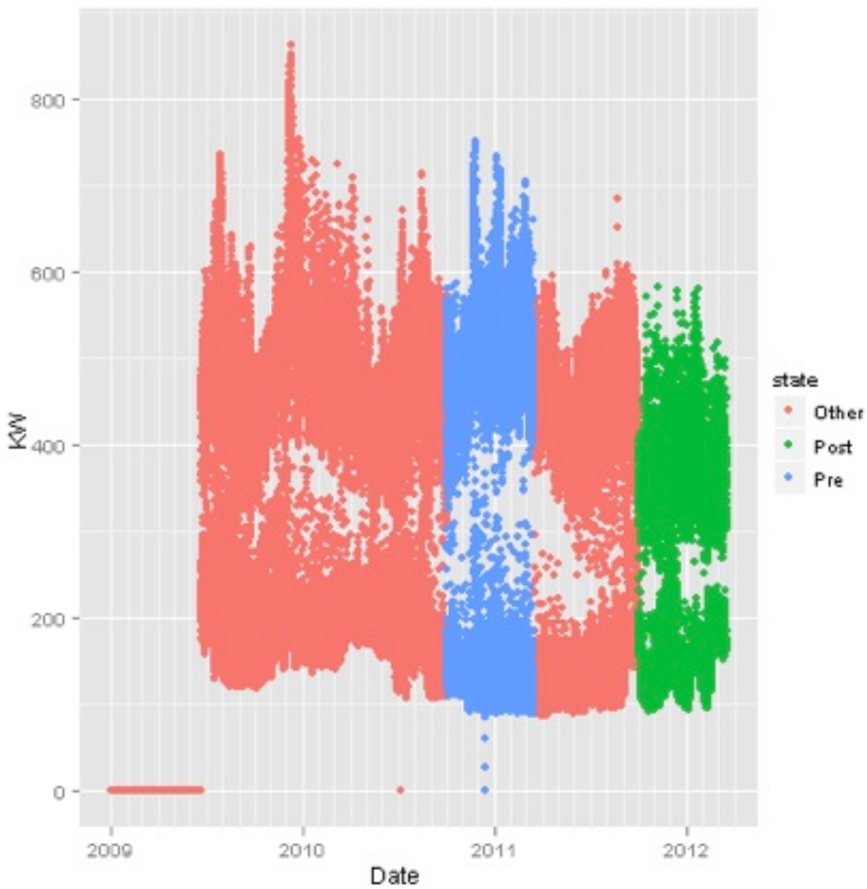
We decided to pull an equivalent period from twelve months earlier for the pre period, as that was before the contract was signed between the PUD and the customer, and it covered the same seasonal period.

To make it concrete, here are selected columns of the last six rows of the data:

RECORDER.ID	KW	datetime	weekday	fracthr	state
...	316	2012-03-16 13:15:00	TRUE	13.25	Post
...	303	2012-03-16 13:30:00	TRUE	13.5	Post
...	303	2012-03-16 13:45:00	TRUE	13.75	Post
...	301	2012-03-16 14:00:00	TRUE	14	Post
...	307	2012-03-16 14:15:00	TRUE	14.25	Post
...	310	2012-03-16 14:30:00	TRUE	14.5	Post

The graph shows the kW data selected for the pre and post periods.

```
qplot(datetime,KW,colour=state,data=cp)+
          scale_x_datetime("Date",major="years",format="%Y")
```



3.2 Analysis Process

This analysis is based on three key concepts: treating the data as a mixture distribution, treating the data as if it came from a lognormal distribution, and computing summary statistics through simulation.

Power data such as this can be seen as a mixture distribution, created by two or more processes that have different characteristics. In this case, the data likely has at least one distribution when the building is "on" and another when it is "off."

To simplify the analysis, I used version 0.4.4 of the mixtools package (Tatiana Benaglia, Didier Chauveau, David R. Hunter, Derek Young (2009). mixtools: An R Package for Analyzing Finite Mixture Models. Journal of Statistical Software, 32(6), 1-29.) on R version 2.14.1 for x86_64-pc-mingw32/x64 (64-bit). The normalmixEM() function from mixtools can be used to extract the components of a mixture distribution, if one specifies the number of expected components.

Based on looking at the data, each component of the mixture seems to be distributed lognormally. That's not surprising, given that the data is all positive, and so I modelled each mixture component as lognormal.

For both the pre and post periods, I fitted a mixture model, printed the summary data, produced the density graphs, and calculated the expected savings as the weighted (by lambda) average of the consumption (mu) in the various power states.

I used simulation to generate the distributions of the estimated savings and the estimated percentage savings. From those distributions, I calculated confidence intervals and the realization rate.

[This paragraph is rendered irrelevant by reformatting done by LBNL]. This report was generated with Emacs org-mode 7.8.03, manually converted from Open Document Format to Office Open XML, and reflowed with manually-inserted page breaks and graphics shifts. The source code and data have been attached to the PDF version of the report. Selected source code and results have been shown for reference in shaded blocks in the report.

4. Results

`modpre01` is a fitted model for the pre period, and `modpost01` is a fitted model for the post period. The code shows the model definitions, the number of iterations required by each to converge, and the summary of each model.

`lambda` is the fraction of time the data spends in each state (each component of the mixture). `mu` is the estimated mean of the distribution of the natural logarithm of kW, and `sigma` is the estimated standard deviation of that distribution.

The next code block shows the definition of the models for the pre and post periods, the calculation of the summary statistics for the lognormal models, and the calculation of the mean values of the mixture components in the pre and the post in kW.

```
modpre01 <- normalmixEM(log(cp[cp$state=="Pre"&cp$KW!=0,"KW"]),
    lambda=0.5,
    mu=c(log(100),log(400)),
    sigma=c(log(50),log(60)))
modpost01 <- normalmixEM(log(cp[cp$state=="Post"&cp$KW!=0,"KW"]),
     lambda=0.5,
     mu=c(4.7,5.1,5.8),
     sigma=c(0.3,0.3,0.5))
summary(modpre01)
summary(modpost01)
exp(modpre01$mu)
exp(modpost01$mu)
```

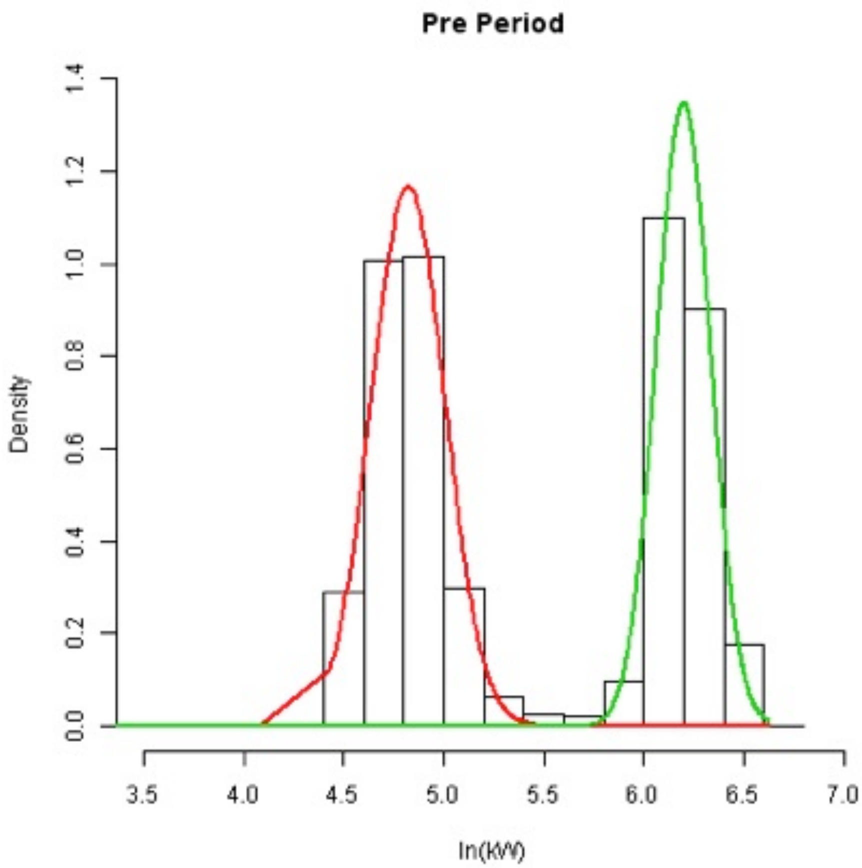
```
number of iterations= 210
number of iterations= 148
summary of normalmixEM object:
       comp 1     comp 2
lambda 0.539959   0.460041
mu     4.822532   6.200602
sigma  0.184532   0.136039
loglik at estimate:     -4444.196
summary of normalmixEM object:
       comp 1     comp 2     comp 3
lambda      0.313401   0.301415   0.385184
mu     4.806790   5.164625   5.889280
sigma  0.122051   0.119376   0.133375
loglik at estimate:     -5241.955
[1] 124.2794 493.0458
[1] 122.3382 174.9719 361.1451
```

The last two lines above show that the pre model has two components, one with 124 kW and the other with 493 kW. The building spent 54% of its time in the low-power mode.

The post period is better modelled as having three components with means of 122 kW, 175 kW, and 361 kW. The building spends 31% of its time in the low-power mode, 30% in the medium-power mode, and 39% in the high-power mode.

The graph shows the pre data distribution and the fitted model.

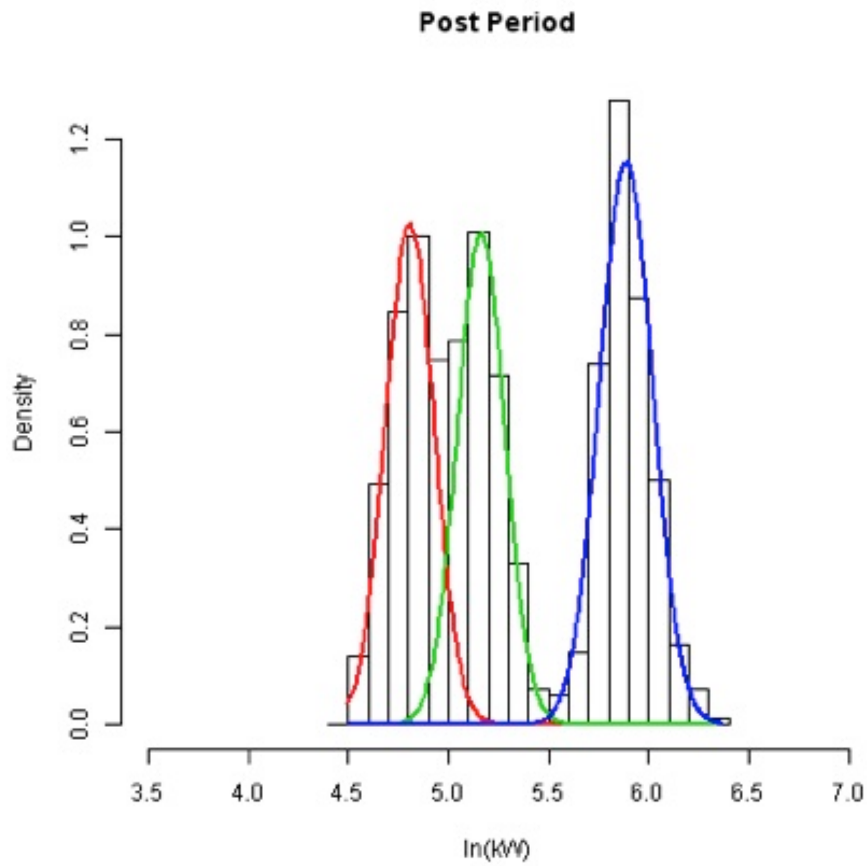
```
plot(modpre01,
     density=TRUE,
     whichplots=2,
     xlab2="ln(kW)",
     xlim=c(3.5,7),
     main2="Pre Period")
```



The graph matches the summary table, with component means at exp(4.82) = 124 kW and exp(6.20) = 493 kW.

Compare that to the post period graph:

```
plot(modpost01,
    density=TRUE,
    whichplots=2,
    xlab2="ln(kW)",
    xlim=c(3.5,7),
    main2="Post Period")
```



Clearly the high-power component now uses less power than in the pre. The low-power component has split in two, with one component somewhat higher than the previous low-power component.

The weighted (by lambda) average of the mean power levels gives the overall average power.

To calculate the distribution of the savings and savings percentage, I simulate 1,000 years of 35,040 quarter-hour data points using the lognormal mixture distribution previously estimated. From that distribution of 1,000 samples, I can calculate summary statistics by counting.

The next code block shows the simulation and the summary statistics for the percentage savings, the total kWh saved in a year, and the realization rate expressed as a percentage.

```
## Get lists of lognormal parameters:
lnpremu <- modpre01$mu
lnpostmu <- modpost01$mu
lnpresigma <- modpre01$sigma
lnpostsigma <- modpost01$sigma
prelambda <- modpre01$lambda
postlambda <- modpost01$lambda
nosims <- 1000

set.seed(1000)
```

```
prepwr <- rep(0,nosims)
postpwr <- rep(0,nosims)
for (i in 1:nosims) {
      prepwr[i] <-
            sum(exp(rnormmix(4*8760,lambda=prelambda,mu=lnpremu,sigma
      =lnpresigma)))/4 postpwr[i] <-
            sum(exp(rnormmix(4*8760,lambda=postlambda,mu=lnpostmu,sig
      ma=lnpostsigma)))/4}

savings <- prepwr-postpwr
savingspc <- 1-postpwr/prepwr

100*summary(savingspc,digits=3)
summary(-savings,digits=6)
100*summary(-savings/546156,digits=3)

  Min.      1st Qu.     Median      Mean        3rd Qu.     Max.
  20.6      21.7        21.9        21.9        22.1        23.0
  Min.      1st Qu.     Median      Mean        3rd Qu.     Max.
-603155     -577309     -570425     -570207     -562715     -528264
  Min.      1st Qu.     Median      Mean        3rd Qu.     Max.
-110.0     -106.0     -104.0        -104.0     -103.0       -96.7
```

We likely saved 570,200 kWh with a 50% confidence interval extending from 562,700 kWh to 577,300 kWh. The realization rate is 104%, and the percentage saved is 21.9%.

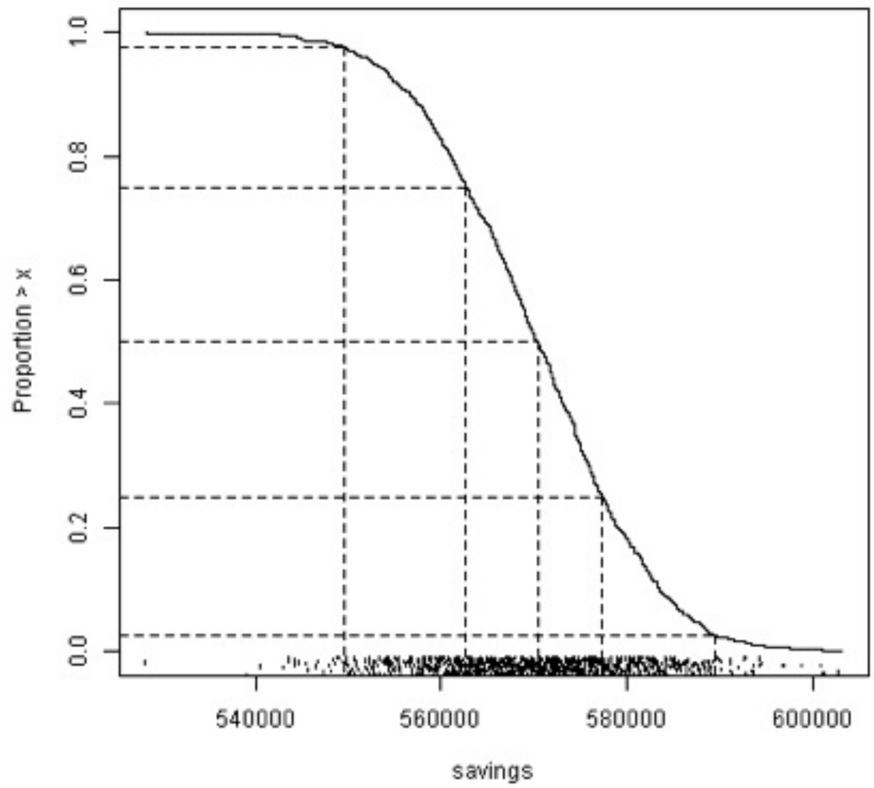
```
quantile(savings,c(0.025,0.975))
100*sum(savings>=546156)/1000

      2.5% 97.5%
549744.5 589599.5
[1] 98.7
```

The 95% confidence interval of savings, conditional on the model and the data, ranges from 549,745 kWh to 589,600 kWh. There is a 98.7% probability that the estimated savings from the measured data exceeds the contracted savings of 546,156 kWh.

The same result can be seen graphically in an empirical cumulative distribution function (ECDF).

```
Ecdf(savings,
      what="1-F",
      datadensity="rug",
      q=c(0.025,0.25,0.5,0.75,0.975),
      subtitles=FALSE)
```



That graph shows the probability (on the ordinate) that at least a particular kWh savings (on the abscissa) has been achieved. Lines on the graph show the 95% and 50% confidence intervals as well as the median savings.

The rug plot along the abscissa shows the distribution of the 1,000 sampled data points used as the basis of this inference.

5. Analysis Limitations

Any modeling has limitations. Here are key limitations I see:

- The data may be serially correlated, but the analysis assumes independent data.
    - It might be argued that this is conditionally independent, identically-distributed data, which would eliminate this concern.
    - If the data is serially correlated, that should only affect the error analysis, making the computed values of sigma too small but not affecting the best estimate of savings.

- The data does not include April through September data.
    - We do not yet have post data that extends for more than six months.
    - Data was picked from the same time of year in the pre and the post to reduce any seasonal effects.

- The match between the estimated savings and the measured savings in the table in the evaluation request suggests this is not likely to be a significant problem. The simulated consumption of 2.6 million kWh in the pre period from October 1, 2010 through March 16, 2011 aligns well with the 2.6 million kWh in the twelve months ending April 2011, and the simulated consumption of 2.0 million kWh in the post period from October 1, 2011 through March 16, 2012 aligns reasonably with the 2.2 million kWh in the twelve months ending March 2011, especially considering that the contractors were still installing and commissioning the system during half of that measured period.

```
Min.         1st Qu.        Median        Mean         3rd Qu.      Max.
2,566,000    2,597,000      2,604,000     2,603,000    2,609,000    2,632,000
Min.         1st Qu.        Median        Mean         3rd Qu.      Max.
2,016,000    2,029,000      2,033,000     2,033,000    2,037,000    2,049,000
```

- Seven data points are missing, because MV-90xi software handles the conversion to and from DST incorrectly.
    - That's less than 0.04% of the data in each of the pre and the post.

6. Acknowledgements
In this project, Snohomish PUD worked with property manager Lee Shephard of Shephard Commercial Real Estate LLC, Bellevue, Washington, Cli_ Chamberlain, PE, of Chamberlain Mechanical Corporation, Kirkland, Washington, and Al Cunningham of Cunningham Engineering, Bellevue, Washington.

III. LBNL Review of the Evaluation

The Evaluation was based on a comparison of load data from a five-and-a-half-month period starting October 1, 2011 to data from the same months a year earlier. Conceptually, the Evaluation takes a standard approach to performing Measurement and Verification (M&V) of retrofit savings: energy consumption over a period prior to the Energy Conservation Measures (ECMs), called the "pre-ECM" period or just "pre" period), is compared to the energy consumption following the ECMs (the "post" period). The energy savings is defined as a "baseline" energy use minus the energy used during the post period, plus or minus "adjustments" due to changes in energy consumption that are not associated with the ECMs. Typically, the baseline energy use is either assumed to be equal to the energy used during the pre period, or is determined from a simple statistical model that includes the effect of explanatory variables such as outdoor air temperature. The standard approach is formalized by ASHRAE Guideline 14 (ASHRAE, 2002).

The Evaluation applies the basic concept of Guideline 14, with a slight modification: in the conventional approach the baseline energy use is uncertain but the energy used during the post period is known, but the Evaluation treats the energy used in the post period as a single realization of a statistical model. In essence, the Evaluation acknowledges that the energy saved in this particular year (or this particular post period) is just one possible value that might over- or under-estimate the savings for other years. This is a correct or at least justifiable way of looking at the issue that is nonetheless non-standard.

Similarly, the Evaluation takes the view that the energy consumed in the pre period is just one possible value that could have been observed in those calendar months. The methodology used in the Evaluation recognizes (correctly) that there is stochastic variability in the energy consumed during the pre and post periods, and that this variability contributes to the uncertainty in the magnitude of savings that can be credited to the ECMs.

In the Evaluation, the "baseline model" is a statistical model in which the load at a given moment is a sample from a mixture distribution. The logarithms of the load data during the pre period are assumed to be draws from a statistical distribution that is the superposition of two Normal (Gaussian) distributions whose parameters are determined from a fit to the data. A similar mixture model, with three Normal distribtions whose parameter values differ from those in the pre period, is used for the post period. (See Figure 1, below).

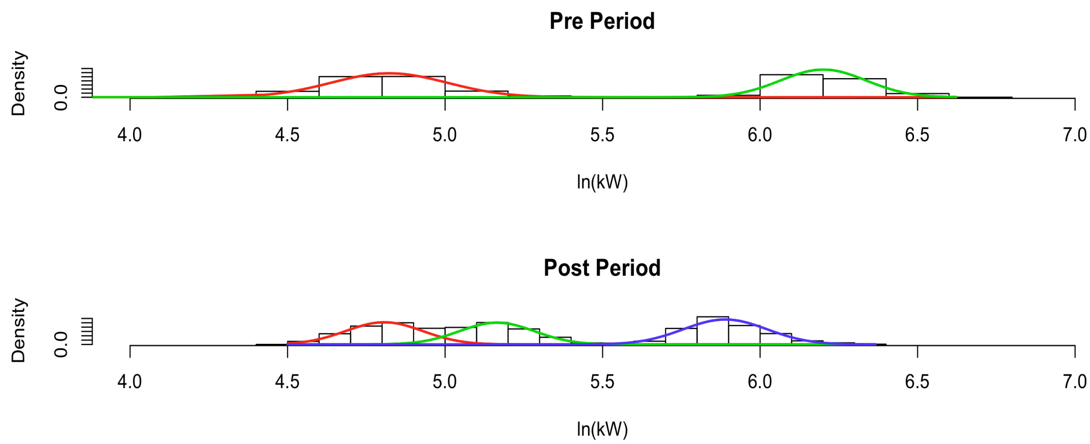


Figure 1: Mixture distributions for pre- and post-ECM period.

The model parameters are chosen to fit the pre and post data, respectively. The models are constructed so that a random sample of simulated data will have approximately the same statistical properties as the actual data from the building. A year of simulated data is drawn at random from the model that represents the pre period, another is drawn from the model for the post period, and the simulated savings are determined by subtracting the total simulated energy consumption in the post period from the total in the pre period. This procedure is followed many times (1000 times, in the Evaluation), yielding 1000 savings estimates. The statistical distribution of those estimates is then summarized.

This approach differs from the more conventional approach in that it does not compare the load in the post period to the baseline load, but rather it compares each simulated load value in the post period to each simulated baseline load.

III A. Comparison of results of the standard approach to the approach used in the Evaluation

The Evaluation reports estimated savings of 570,000 kWh, and a 50% confidence interval from 562,700 kWh to 577,300 kWh. (There is a somewhat complicated terminology surrounding the terms "confidence intervals," "credible intervals," "prediction intervals," and so on. The evaluation uses the term "confidence interval" to describe the uncertainty distribution, and we think this is in fact the correct term while acknowledging that we are not positive of this).

A more conventional approach to M&V would be based simply on the average load during the pre and post period. The average load during the pre period was 279 kW, and the average in the post period was 232 kW. Extrapolating to a year (8760 hours), this yields an estimated savings of 570,000 kWh (identical to that reported by the Evaluation).

ASHRAE Guideline 14 allows several methods of estimating the uncertainty in the savings, depending on the details of the baseline model used. The simplest baseline model would assume that the load is constant during the pre period and during the post period, and treat deviations from this model as errors (or "residuals"). If these errors are (incorrectly) treated as statistically independent, the uncertainty in the savings would be calculated as follows. The standard deviation of the load during the pre period is 191 kW, and there are 16,000 intervals in the pre period, so the uncertainty in the mean during the pre period is 191/sqrt(16000) = 1.5 kW. The same calculation holds for the post period. (There is no way to know the error of the baseline in the post period, so for this calculation we assume it is the same as the pre period). The total savings uncertainty according to this calculation method, expressed as a standard error of the mean, is $v = \sqrt{1.5^2 + 1.5^2} = 2.12$ kW. For a year with 8760 hours, this translates to a standard error of 18,600 kWh. To compare to the results from the Evaluation, we express this as a 50% confidence interval by using the fact that the central 50% of a normal distribution is contained within +/- 0.75 standard errors of the mean, which in this case is 14,000 kWh. Thus, this method would yield an estimated savings of 570,000 kWh, and a 50% confidence interval from 556,000 kWh to 584,000 kWh.

ASHRAE Guideline 14 notes the importance of adjusting for autocorrelation and proposes a method for doing so. The method relies on calculating an 'effective' number of independent observations, based on the lag-1 autocorrelation of the residuals. The ASHRAE formula is correct for a specific type of time-series model on the residuals, specifically an AR(1) model, but in fact most real building data are not well described by an AR(1) model and the ASHRAE formula may either over- or under-estimate the uncertainty. With the data from the subject building, the ASHRAE formula would estimate a 50% confidence interval about eight times wider than the interval based on an assumption of independent data points: the 50% confidence interval ranges from 458,000 kWh to 682,000 kWh.

We see that in this building, using these data, both the conventional ASHRAE approach and the approach used in the Evaluation yield very similar results for the central estimate of the annual savings: 570,000 kWh. They also yield similar results for the uncertainty in this number, if autocorrelation is ignored. This is not a coincidence, and can be expected to be the case in similar analyses of other datasets: the fact that the mixture models are a good statistical representation of the data from which they were constructed implies that both the mean and the standard deviation of the simulated data will be close to the mean and standard deviation of the actual data, which, through the Central Limit Theorem, in turn implies that the mean and the standard error of the mean will be similar in the final results. (This is true even though the mixture distribution assumed in the Evaluation is rather non-normal, being constructed of a superposition of three lognormal distributions.)

III B. Advantage of the Approach used in the Evaluation

The conventional approach of simply comparing load during the pre- and post-ECM period does not necessarily provide any insight or information about how the energy savings came about. In contrast, the mixture-model approach used in the Evaluation shows that in the post-ECM period there was a substantial reduction in the upper end of the statistical distribution of load: the daily peaks aren't nearly as high. If this was an expected or intended outcome of the specific ECMs implemented in this building, then this observation suggests that the observed energy savings are "real" and not the result of stochastic variability or of some change in the building that is unrelated to the ECMs. As a general principle, trying to identify the "fingerprint" of the ECMs is probably good practice: automatic timers that shut off lights and equipment after hours should show up in the after-hours load, not in the peak load; improvements to the HVAC system should show up when the HVAC system is operating; and so on. This is a nascent area of M&V practice and so far there are no formalized methods for applying these concepts., but the principle can still be used to guide M&V analysis and to suggest alternative analyses if a result seems unusual.

In the building in question, the main ECM was the installation of Variable Air Volume (VAV) units and the alteration of the building's HVAC controls to take advantage of these units. These improvements would be expected to affect the electric load only when the ventilation system is turned on, and to provide more benefit on average when the HVAC load is high. The high-load mixture component represents data from (roughly) business hours on weekdays, when occupancy is high and the HVAC load is highest, so the substantial downward shift in this mode is what is expected given this ECM. However, this observation begs the question of how the building should be evaluated if it had somehow attained the same estimated energy savings but not through a downward shift of that upper component.

Actually in this particular building there are only two ways such a large magnitude could show up: that high-consumption component of the mixture represents most of the energy used by the building, so any large savings would have to come from either (1) a downward shift, or at least a downward skew, to that component, or (2) from a shift of hours of operation out of the high-consumption component into one of the others. But suppose there were some other building in which VAV boxes were installed and the upper component of the mixture remained unchanged between the pre and post period, but there was a downward shift to a lower-energy component of the mixture. Would one then conclude that the VAV boxes were effective in reducing the energy use in the building, but for reasons not understood? Or would one conclude that the VAV boxes were not saving substantial energy but that some other energy-saving actions had taken place? (The latter is not very far-fetched: the attention paid to building systems and schedules during a retrofit might well lead to discovering and implementing energy saving opportunities in addition to the originally planned ECMs). In this Review we do not propose to answer these questions, we merely note that the ability to separate the energy savings into mixture components, or to otherwise characterize the ways in which the savings occur, is a potentially valuable improvement over other approaches but that we do not currently see how to exploit that potential.

III C. Shortcomings of both approaches

Neither the ASHRAE approach nor the approach taken in the Evaluation will yield accurate uncertainty estimates. The assumption of independent data points (as in the Evaluation) will lead to confidence intervals that are much too narrow. The assumption that autocorrelation of data points is described by an AR(1) time series model may lead to confidence intervals that are too narrow or too wide, depending on the details of how the AR(1) assumption is violated in the data from that particular building and time period; detailed analysis of this is outside the scope of this Review.

III D. Conclusions from Review of an Evaluation of a Prior Commercial Building Retrofit Project

The mixture-model approach used in the Evaluation has some advantages over more conventional approaches, and does not have substantial disadvantages for this building. Or, rather, it does not have *technical* disadvantages: it is undeniably somewhat more complicated to perform and to understand.

As a general rule – as opposed to the application to the specific building in this evaluation – the mixture model approach that was used by Snohomish PUD should give nearly the same central estimate and confidence intervals as a conventional approach based on the pre- and post-ECM mean load. In both cases the confidence intervals will be too narrow, since both approaches rely on an assumption of statistically independent data points that is strongly violated by actual data from most buildings.

The Snohomish approach based on mixture modeling has the advantage that it gives some insight into where the savings are coming from. Such insight is clearly valuable from the perspective of an engineer trying to understand what is happening. It's less clear whether this insight can turn into actionable assessment in the context of an either/or decision about whether an energy target has been met. In the present building the benefits of the ECM are large enough that there is little question that the target savings were attained, but one can imagine other buildings in which the situation is much less clear.

In short, the approach used in the Evaluation is a slight improvement over common industry practice. It can be further improved as discussed below.

IV. Suggestions for Future Improvements

1. Plot the load data time series in a way that allows important features to be seen.

The first step in any analysis of load data should be to plot the load data as a function of time. Plotting a full year of interval data – or even a few months – can yield plots that are hard to interpret if care isn't taken in selecting the aspect ratio and other parameters. The raw data plots in the Evaluation suffer from some problems in this regard. Price (2010) provides some advice for how to create plots.

Useful plots include (1) overlay data from one period over data from a comparison period (as in Figure 2); and (2) a single plot that shows a range of data starting a few weeks before the installation of an ECM and ending a few weeks after (Figure 3).

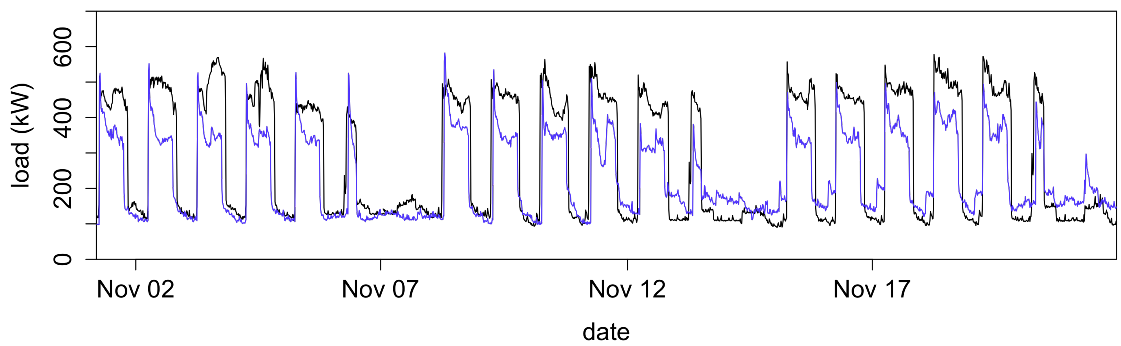


Figure 2: Load during pre period (black) and 52 weeks later during the post period (blue). Dates shown are for the Pre period; because the year is not a multiple of 7 days long, dates in the Post period are off by one day.

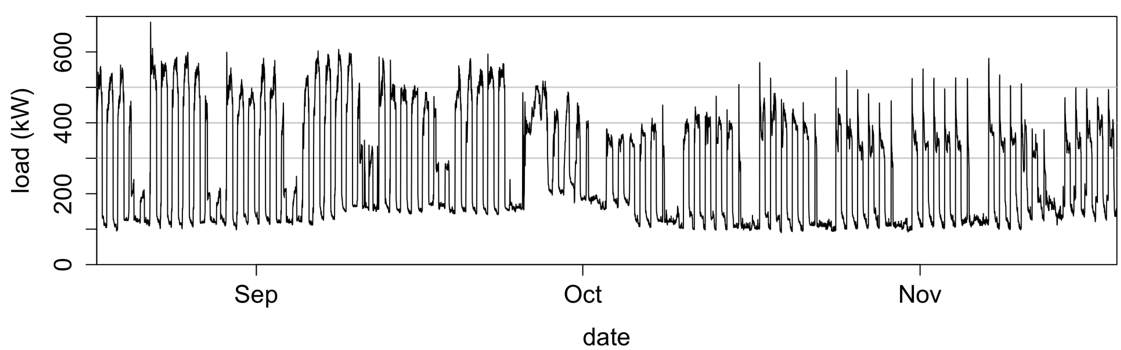


Figure 3: Load before, during, and after the installation of Energy Conservation Measures in late September. Peak load drops from around 500 kW in late September to around 400 kW in early October (except for one 15-minute interval each day).

The point of the plots is to look for irregularities or anomalies that could interfere with the ability to estimate the energy savings from the ECM. There are so many examples that it is impossible to list them all. Some examples of possible phenomena are:

1. The building's base load (minimum nighttime load) increases because the building stops shutting down at night or because changes are made to outdoor lighting or because of the addition of continuously operated equipment such as computer servers.
2. The building's base load decreases because scheduling errors are corrected, some night-time lighting is eliminated, or for other reasons.
3. The building begins using more or less energy during the day because of a change in occupancy levels or production levels.
4. The building begins using more or less energy on hot days because of a change in temperature setpoints.

There are many other examples as well.

There is no current computer technology that can reliably analyze an electric load time series and identify all of the behaviors that can occur. If computer tools are provided that make it easy to generate the required plots, it does not take a highly skilled statistician or data analyst to recognize many features that could lead to incorrect M&V assessments. For example, Figure 4 shows electric load data from a building that stops shutting down at night. If this behavior occurs in the "pre" period of an M&V assessment, the effectiveness of the ECM may be overestimated: if this behavior is remedied in (or before) the "post" period, the building's energy consumption will decrease whether or not the ECM is effective. Conversely, if behavior such as this were to commence during the "post" period, even the energy savings of a highly effective ECM would be masked by the energy wasted by operating the building at night.

Determining ways to handle these sorts of phenomena may require skilled engineers or data analysts who can figure out how to adjust for them, but much less skill is required in order to simply notice that something is anomalous. So, one approach to scaling up M&V practices to handle a large number of buildings without a large staff of skilled analysts would be to train some less-skilled analysts to visually scan load data plots for problematic behavior. Buildings that exhibit such behavior can be passed along to more highly trained or skilled analysts, while the rest of the buildings (hopefully the majority) can be processed using standard methods.

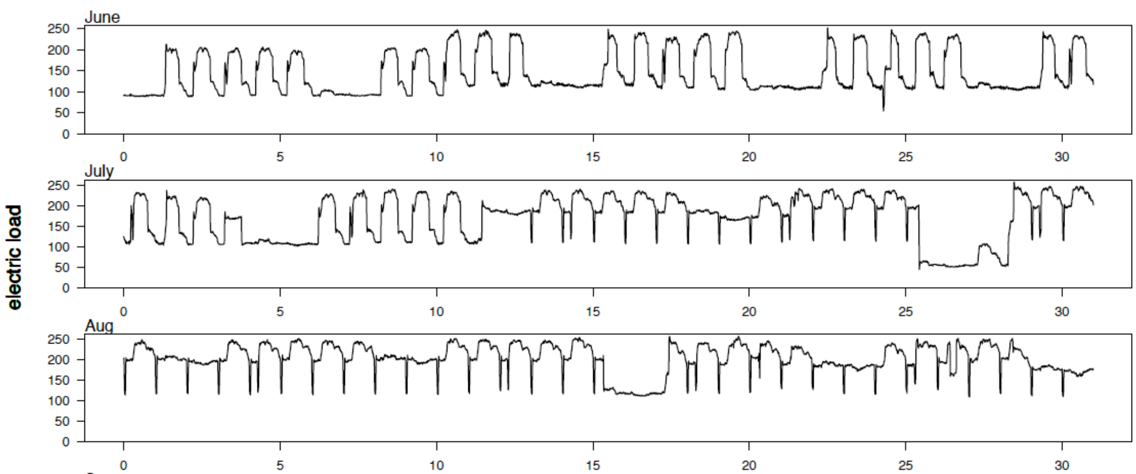


Figure 4: Three months of electric load data (load is in kW) for a courthouse building. The building stops shutting down at night, starting in mid-July.

2. Plot the load data vs outdoor air temperature (and other explanatory variables, if data are available).

Figure 5 shows the electric load as a function of the outdoor air temperature during the pre period (black) and the post period (blue). The data divide into an upper cloud of points that shows evident temperature-dependence of load, and a lower cloud that shows little or no temperature dependence. This is a common phenomenon among commercial buildings, many of which are temperature-controlled during the day and either uncontrolled or less tightly controlled at night.

The two or more mixture components that were identified in the Evaluation are clearly seen here. As with the mixture modeling, this plot invites the question of whether the observed changes in the building are consistent with what is expected from the ECM that was performed.

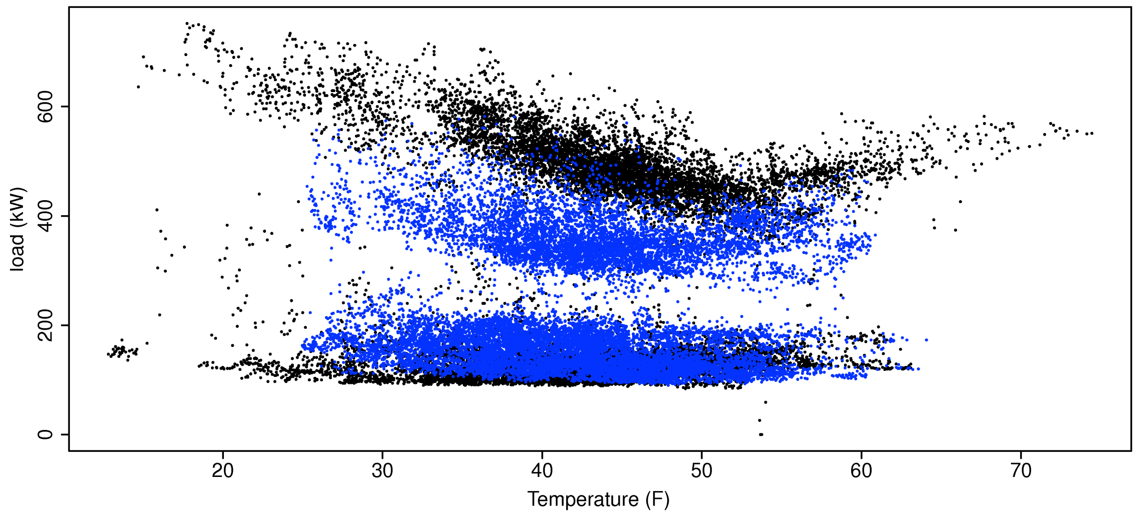


Figure 5: Load as a function of outdoor air temperature during the pre period (black dots) and post period (blue).

3. Choose an appropriate baseline method, given what is learned from plotting the time series and the load versus temperature.

When comparing a pre- and post- period that are several months long and exactly one year apart, it is often not crucial to account for the dependence on temperature: as long as the statistical distribution of temperatures is about the same in both periods the temperature-dependence will average out.

A. What to do if there are no load anomalies and if the pre and post period have comparable outdoor air temperature distributions:

This is the case for the subject building and the pre and post periods used for the Evaluation. In this case adjustment for temperature is not needed and the average load from the pre and post period can be directly compared, as was done in the Evaluation.

To get a valid uncertainty estimate, autocorrelation of the individual data points must be taken into account. If the autocorrelation is low beyond a certain lag, a bootstrap method (Efron, 1977) can be used to sample chunks of load data – that is, contiguous segments of data – from the pre period with replacement in order to generate a simulated year of data, and the same from the post period, and the difference in energy consumed during these years provides one estimate of the energy savings. Repeating this process many times will yield a statistical distribution of estimated savings. This is in the same spirit as the uncertainty estimate used in the Evaluation, but in the subject building we find that the sampling timescale must be approximately 1 week rather than 15 minutes, which makes a rather large difference in the uncertainty estimate: this procedure leads a 50% confidence interval of 528-581 MWh saved, a span of 53 MWh; this compares to the Evaluation's estimate of 563-577, which is a span of only 14 MWh and is clearly too narrow.

B. What to do if there are no load anomalies but the pre and post period are not comparable in terms of outdoor air temperature distributions:

The more temperature-sensitive the load in the building, the smaller is the difference in weather between pre and post period that can be tolerated without some sort of weather adjustment. It is usually advisable to adjust for temperature dependence: if the pre period is substantially harsher or milder than the post period, the estimated energy savings will be under- or over-estimated.

The standard recommended approach to adjusting for temperature is to create a "baseline" statistical model that adjusts for temperature and is fit to data from the pre period, and to use this model to predict the energy consumption during the post period. This approach is described in IPMVP Option C, sections 3.4.3.3-3.4.3.7 (IPMVP, 2002). Until recent years it was standard to create such models using monthly electricity data (usually the only available data) and integrated weather metrics such as Heating Degree-Days and Cooling Degree-Days; nowadays it is possible to create the models using interval load and temperature data, which allows much better estimation of the effect of temperature from even relatively small amounts of data. (For instance, even one hot week can be enough to quantify the effect of high temperatures, which was not true with monthly aggregated data only).

This approach is illustrated in Figure 6 - Figure 8. Figure 6 shows the load and temperature data from the Pre period, and predictions from the resulting model fit to these data. The model whose results are shown here is described in Mathieu et al. (2011); it is a regression model in which the prediction at a given time depends on the time during the week – analogous to time of day, but counting sequentially through a 7-day period rather than a single day – and on the outdoor air temperature, where the dependence of load on outdoor air temperature is assumed to be piecewise-linear. This particular model is used here only to illustrate the point that a statistical model can capture at least some of the time- and temperature-dependence; we don't claim this is the best model that could be created for this building.

The model predictions are, obviously, not perfect, although (since an unbiased model was used) the average of the predictions is equal to the average load during this period. Higher load during some especially cold periods is evident in both the actual data and in the model predictions.

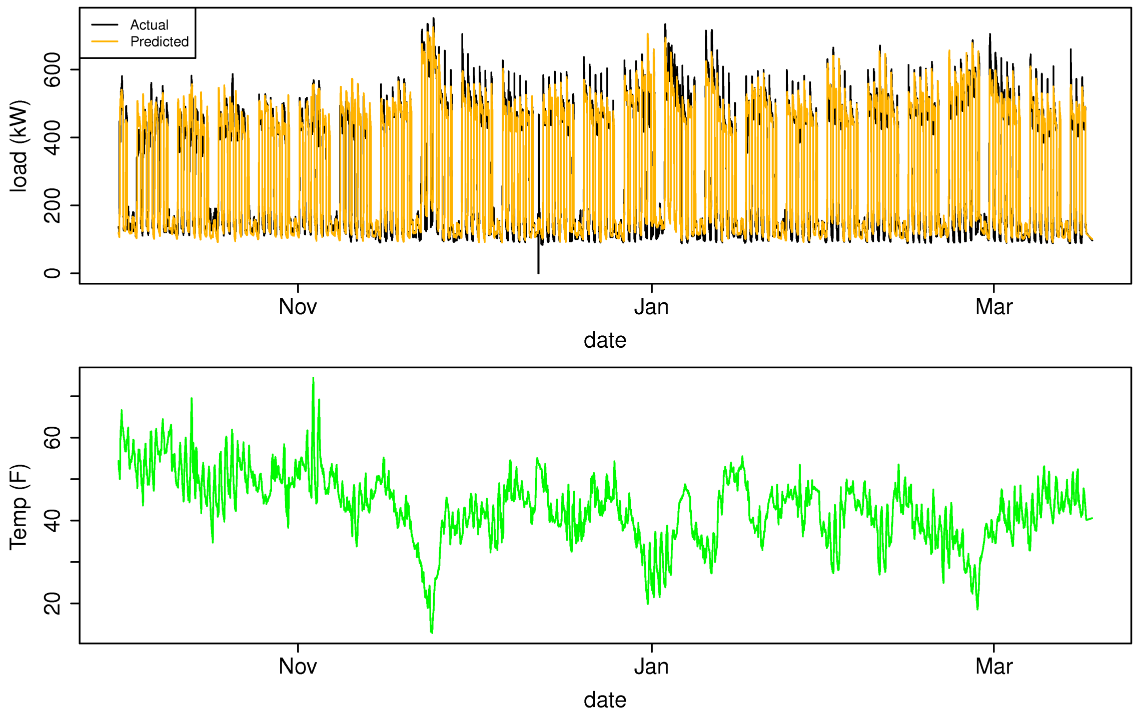


Figure 6. Top: Actual load and predicted load of the subject building during the Pre period, from a statistical model that accounts for the time of week and the outdoor air temperature. Bottom: Outdoor air temperature during the Pre period.

Figure 7 shows (in orange) the result of using the same statistical model that was fit to the pre-ECM data, but using the outdoor air temperatures that occurred in the post period. The relationship between predictions and actual data is hard to see in this plot – as previously noted, plotting choices are important when looking at load data because important features can easily be obscured.

Figure 8 shows only the first six weeks of the post period. On this plot it is easy to see that the load is substantially below the baseline prediction during the day on weekdays, and about the same as the baseline prediction during nights and weekends.

The load in the building varies somewhat with temperature. However, the predicted mean load during the post period is essentially identical to the predicted (and actual) mean load during the pre period, because the statistical distribution of temperatures was very similar in the pre and post periods. So for these pre and post periods and this particular building, there is no difference between estimating the baseline using the regression model as (which adjusts for temperature) and using the average load from the pre period (which does not). In this case the temperature-dependence can be ignored, as was done in the Evaluation. However, this would not be true of other buildings, or of this building if the pre and post periods had substantially different weather, such as a harsh winter compared to a mild one.

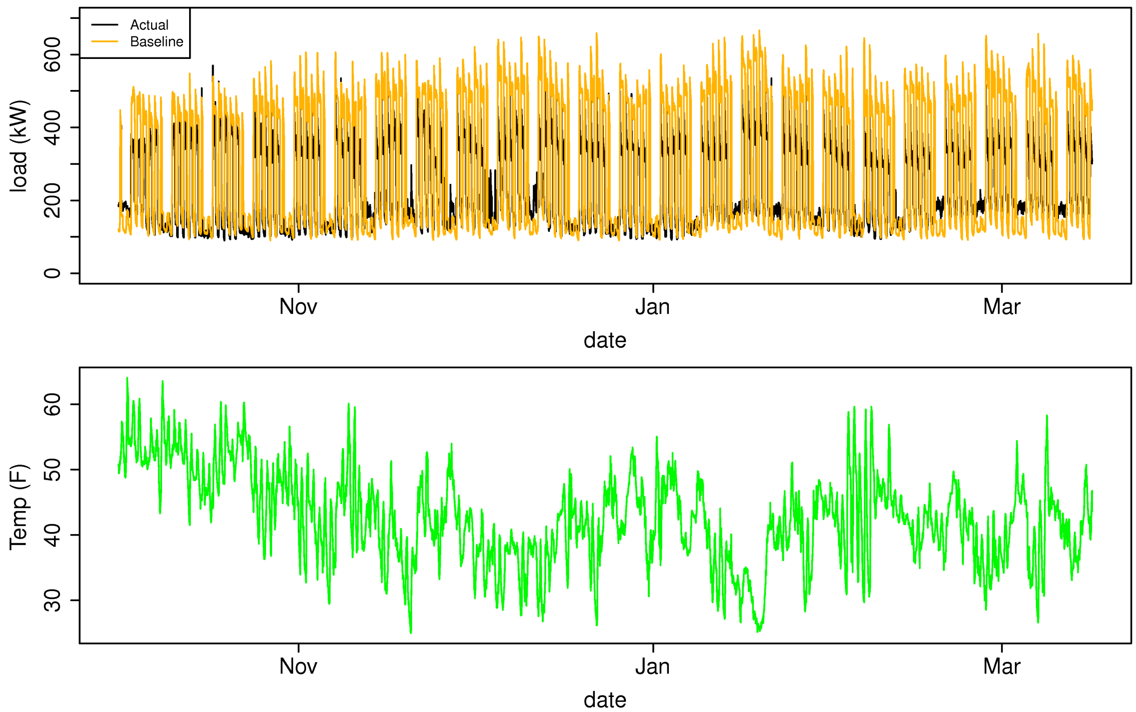


Figure 7: Actual load and predicted baseline load of the subject load during the Post period. Bottom: Temperatures during the Post period.

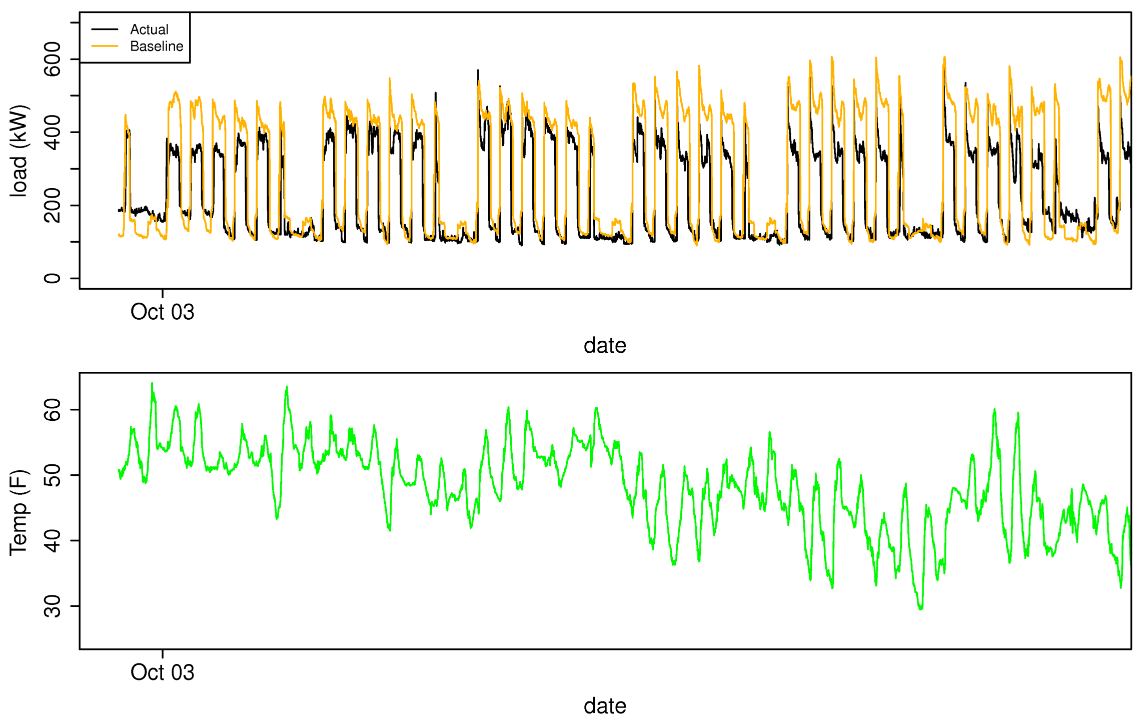


Figure 8: Same as Figure 7, but showing only the first six weeks of the Post period.

Even if the pre and post period experience identical weather, adjusting for outdoor air temperature can be important if either the pre or post period (or both) is substantially shorter than a year. For example, if an effect of the ECM is to reduce the *slope* of the dependence of load on temperature (as would be the case with some HVAC improvements) then the portion of the year that is represented by the pre and post period can matter: if the periods sample mostly mild weather, the effect of the ECM will show up less strongly than if the periods sample more extreme weather. The subject building may be an example: Figure 5 suggests that the ECM may save more energy when the outdoor air temperature is below 40F than when outdoor air temperature is near 55F, so if the pre and post periods oversample lower temperatures (compared to the temperature distribution during a full year) the savings may be over-estimated.

On the other hand, the changes in the subject building might also save substantial energy for outdoor air temperatures above 60F, a temperature range that is nearly absent in the post period but presumably would be experienced at times during the rest of the year, so it's also possible that the savings are underestimated or are correctly estimated. Given the fact that the post period does not include the summer, there's simply no way to know the building's energy performance in the case of higher outdoor temperatures.

One somewhat unconventional approach to handling the problem of pre and/or post periods for which the statistical distribution of temperature differs from the year as a whole is to (a) fit a statistical model for the pre period and another model for the post period, (b) use both models to predict the energy consumption during a typical year, and (c) estimate the average annual energy savings as the difference between these predictions. If this approach is used, it is important to use statistical baseline models are unbiased or nearly unbiased (in the formal statistical sense). We think ASHRAE Guideline 14 sets an overly strict requirement for bias, but we agree with the general principle that the bias should be low. This may prevent the use of otherwise attractive approaches such as using linear regression on the logarithm of the load rather than on the load itself, unless a bias adjustment is made.

To recap: there is the conventional approach of subtracting the actual load in the Post period from the baseline prediction to estimate the savings; or the unconventional approach of fitting baseline models to both the pre and post periods, predicting the energy consumption during a typical weather year, and then subtracting to estimate the savings.

Either way, one obtains only a point estimate that does not automatically provide a confidence interval. To estimate a confidence interval, one approach is to again rely on a bootstrap method, but this time to sample from the distribution of model residuals. The following is the recipe for doing this.

1. Fit the statistical model to the Pre period.
2. Apply the statistical model to generate a baseline prediction for the Post period.
3. Calculate the residuals from the Pre period as a function of time (that is, actual load minus predicted load).
4. Draw random chunks of residuals to generate a simulated year of residuals. The chunks – contiguous intervals of residuals – should be about the length of the autocorrelation scale of the residuals, which is roughly 1 week in the subject building. Sampling should be done with replacement.
5. Add the simulated year of residuals to the baseline prediction for the Post period, to get a simulated baseline that includes the effect of stochastic variability.

6. Subtract the actual load used in the Post period to get the estimated savings at each time interval, and add these up for the entire Post period to get the estimated energy savings during the Post period.
7. Repeat steps 4 through 6 many times to get a statistical distribution of savings estimates.

The procedure described above will account for only the uncertainty in the baseline estimate for the Post period, and not for stochastic variability of the actual data from the Post period. As such, it will provide a statistical distribution that estimates the uncertainty in the energy that was actually saved during the Post period, but not the uncertainty in the energy that might be saved in a similar period in a different year.

V. Summary

1. The Evaluation provided a reasonable estimate of the energy saved during the post-install period in the subject building.
2. Adjustment for outdoor air temperature is generally recommended. In the case of the subject building and the specific seasons that were analyzed, such an adjustment has almost no affect on the estimated energy savings. In other cases it can be important.
3. Both the Pre- and Post-install periods included only the five coolest months of the year. Extrapolating the savings in these months to an entire year without adjusting for outdoor air temperature is not recommended in general but is probably acceptable in the subject building because the energy savings appears be roughly the same for a wide range of temperatures.
4. The Evaluation provides a confidence interval for the energy saved during the year. The method is based on sampling from a distribution of possible energy observations. The basic approach is quite reasonable but the specific implementation assumed that each data point is statistically independent, an assumption that is strongly violated by the data; the result is a confidence interval is much too narrow.
5. A modified sampling approach will provide a much more reasonable confidence interval. The modification is to sample contiguous segments of data that are long enough in duration that the data points at the beginning of the segment are approximately uncorrelated with those at the end. In the case of the subject building, this duration is about 1 week.
6. An improvement over the methods in the Evaluation would be to use a baseline method based on linear regression to adjust for outdoor air temperature when estimating savings (as described in IPMVP Chapter 3), and using a method based on sampling of the residuals in order to estimate the uncertainty. There may be other good methods as well.

VI. References

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

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

Efron B, Bootstrap Methods: Another Look at the Jackknife. Annals of Statistics 7:1-26, 1979.

IPMVP: International Performance Measurement & Verification Protocol Volume I: Concepts and Options for Determining Energy Savings, U.S. Department of Energy Report Number DOE/GO-102002-1554, 2002

Mathieu JL, Price PN, Kiliccote S, and Piette MA. Quantifying changes in building energy use, with application to Demand Response. IEEE Transactions on Smart Grid 2:507-518, 2011.

Price PN. Methods for analyzing electric load shape and its variability. Lawrence Berkeley National Laboratory report LBNL-3713E, a project report to California Energy Commission, June 2010.

Price PN, Granderson J, Sohn MD, Addy N, Jump D. Commercial Building Energy Baseline Modeling Software: Performance Metrics and Method Testing with Open Source Models and Implications for Proprietary Software Testing. Lawrence Berkeley National Laboratory report LBNL-6602E and Pacific Gas and Electric report ET12PGE5312. September 2013.