Evaluability Assessment for the BPA Commercial SEM Pilot Program

Prepared for **Bonneville Power Administration**
Carrie Cobb

Prepared by **SBW Consulting, Inc.**
2820 Northup Way, Suite 230
Bellevue, WA 98004

August 11, 2017

ENERGY + WATER + EFFICIENCY

TABLE OF CONTENTS

1. Purpose

The purpose of this memo is to document our findings from an investigation of BPA's Commercial Strategic Energy Management (SEM) Program and to identify factors that are important for BPA to consider related to future evaluations of the program's savings. The memo first describes our understanding of how the program will be operated. This is followed by our thoughts on the program's evaluability. Last, we describe findings from our review of the savings estimation modeling conducted for three pilot buildings and provide recommendations on how this modeling may be improved. Our review of the pilot buildings helped us to identify and refine factors that should be considered in planning for future evaluation of this program.

2. Executive Summary

The Bonneville Power Administration is developing a Strategic Energy Management (SEM) program for commercial buildings served by its customer utilities. BPA conducted a pilot test of the program, including a test of the program's savings estimation procedure. We reviewed the program design, the pilot projects' data and energy savings estimation models, and identified factors relevant to future evaluations of the program.

There will be two program delivery tracks: "SEM Lite" and "SEM Heavy." The former is more flexible and is designed for Option 2 and large Option 1 utilities, who have technical staff that can take the lead in providing SEM services to program participants. SEM Heavy will provide additional resources from BPA via training and individual support to utilities and participants.

Both program tracks require screening of candidate commercial buildings for availability of energy data, potential for an accurate statistical model of energy use, and savings potential. The energy models are based on regression, typically using ASHRAE change-point models, and include a calculation of uncertainty, which is used to judge the accuracy of the saving estimates.

There are reporting requirements, including: a site assessment; the baseline model and a Baseline Model Analysis Report; general M&V Plan; and the Custom Project Proposal and Completion Report describing the site, baseline, implemented measures, and final estimated savings.

An energy model is created for each participating building. The energy modeling is adherent to IPMVP and follows recommendations found in the BPA Verification by Energy Modeling Protocol and the BPA Regression for M&V: Reference Guide.

2.1. Factors Important in Future Evaluation

The following factors should be considered in planning for future evaluations of this program:

- Pre-implementation savings estimates can be helpful to evaluators. They provide corroboration of the evaluated savings, or an indication of a need for review of the evaluation calculations. Providing information about what times or under what conditions the SEM actions are expected to save energy can also be useful to evaluators when reviewing savings models. Further research is needed to determine how to handle buildings where pre-implementation savings estimates are not within the uncertainty bounds of the regression model estimate of savings.
- The regression models should provide accurate and transparent estimates of savings and precision. Models with physical significance facilitate the comparison of results with expectations. Models that follow industry guidelines and common practice will be transparent and readily evaluated.
- Non-routine adjustments may best be estimated using statistical methods. When engineering calculations are used, the calculations should reside in the same workbook as the regression model so that the adjustments can easily be accessed by evaluators. If

records can be kept regarding the timing of SEM actions or other program activities, they will be very helpful in verifying that the energy use changes visible in the data can be attributed to the program.

- Even more so than for other program changes, it is important to know the timing of implementation for capital measures. This will facilitate reconciliation of the capital measure savings with the regression model estimates of savings.
- BPA needs to have a policy and provide clear direction on how negative savings estimates are treated for both the setting of incentives and the estimation of program savings. We recommend that negative savings be treated the same as positive savings in estimating program savings. When program models show negative savings, the program should document any reasons why the savings estimates are negative.
- Since BPA's SEM program may cover a wide geographical area, encompassing many utility and regulatory jurisdictions, it is very important that the program document whether the applicable baseline is pre-conditions or current practice. The need for a current practice, instead of a pre-conditions baseline, needs to be established as part of project screening.

2.2. Pilot Project Review and Recommendations

The pilot projects were well-analyzed using the energy models developed by BPA. Our suggestions for improvement were mostly toward using the simplest models possible that fit the data, but the choices made for the pilot project models were all reasonable. The overall conclusion is that projects can be evaluated using the proposed process. Savings can be estimated with sufficient precision except for sites and meters that have very low savings, and most of these types of sites are not expected make it through screening.

Four important recommendations were made as a result of this review. These recommendations will help ensure that evaluation is efficient by reducing the changes needed to be made to the program's models and analyses.

- We recommend that engineers use models with the fewest number of parameters that still fit the data.
- We recommend that a model of the reporting period data always be created.
- Modelers should ensure that significant modeling choices are traceable.
- Analysts should document analyses and decision-making rationales.

3. Program Design and Expected Operation

We conducted the following activities to learn about the program's design and its expected operation:

- Studied program M&V flow diagram and checklist.
- Examined documentation from three pilot projects.
- Interviewed Erik Boyer and discussed the program's expected operation with Erik and other BPA staff.

The following describes the expected operation of the program.

3.1. Program Tracks

To design the Commercial SEM program, BPA leveraged experience and adopted procedures from its custom project processes and other SEM-like programs in the region such as the programs offered by BPA's Energy Smart Industrial program, Energy Trust of Oregon (ETO) and Puget Sound Energy (PSE). Based on these experiences and the results from three pilot projects, BPA is developing two tracks for delivering the program: SEM Lite and SEM Heavy. The "Lite" and "Heavy" distinctions convey the amount of resources BPA intends to put into each track.

SEM Lite is focused on constructing a set of requirements which participating utilities must meet to claim savings with BPA. BPA's intention with this track is greater flexibility such that the utility can design and run its own program. This track is generally designed for Option 2 and large Option 1 utilities which have in-house technical staff. Engagement with the program will likely be initiated by the utility energy service companies or by interested end users. BPA engineers will be available to support developing whole building regression models and assessing model performance among other support activities.

SEM Heavy is targeted toward supporting a cohort infrastructure in which participants attend regular workshops and webinars with their peers to learn about and share effective energy management strategies. BPA will develop the materials for and host these training sessions. Utilities will be responsible for recruiting and collaborating with individual participants. Additionally, for Option 1 utilities, BPA will provide more hands-on individual support to the utilities and end users as well as help recruit participants.

Regardless of track, BPA anticipates the program will operate as follows.

3.2. Project Initiation

Once a candidate building has been identified, it must be screened to ensure reliable savings can be estimated. Two hurdles must be met to participate: 1) the candidate must be able to commit the resources to participating in the program; and 2) the last 12 months of energy consumption must be representative of typical building operations.

Resources required from program participants include staff (energy champion) knowledgeable about SEM practices and available to implement these practices themselves or in collaboration with contractor, utility or BPA staff. Furthermore, participants must be able to provide sufficient documentation of SEM activities and non-routine adjustments throughout the performance period.

Criteria for screening the building and developing the baseline model(s) are shown in Appendix B.

At least one year of monthly electric use data is required, and three years desired, to support the development of a whole building regression model for the baseline period. During the baseline period, there must be no major changes in building structure, equipment or systems, occupancy, or operations. Similarly, no major non-routine changes, e.g., large change in building floor area or major occupancy change, can be planned for the performance period. A preliminary model will be developed. Current plans are to use the Energy Charting and Metrics (ECAM) tool. BPA can assist with developing the model if the candidate or utility does not have the expertise to do so.

If the candidate passes these two hurdles, then an assessment is made of potential opportunities for savings. This could identify O&M and capital savings opportunities as well as gather information about building operations, but it could also simply mean that the energy use shows potential for reductions. Information from the assessment is used to estimate savings potential as a percentage of consumption and make refinements to the baseline model, such as changes to the schedule of occupancy or operational modes. Statistical results for the model are examined for goodness of fit and to determine whether savings potential is greater than model uncertainty.

Based on the comparison of savings potential and model uncertainty, if the program has confidence that savings can be achieved at the site, the program develops a Custom Project Proposal (CPP) and M&V Plan for presentation to the participant.

At this time, it is also important to make an initial determination of whether the appropriate baseline is simply the existing conditions, or whether current practice baselines will be needed.

Upon agreement between the program and participant, the SEM project is initiated at the site and the performance period begins.

3.3. Savings Analysis

During the performance period, the program interacts with the energy champion at least monthly, monitoring consumption and ensuring SEM actions are being taken and documented. Documentation will include the date measures become operational and the magnitude of savings. Additionally, the baseline model may be further refined if unanticipated non-routine adjustments are made during the performance period.

BPA intends for the contract performance period to last one year and SEM project savings to have a one-year measure life. At the end of each performance year, whole building savings are determined by subtracting the performance year consumption from the adjusted baseline consumption *forecast* for that year using the baseline model. To determine SEM savings, the savings from any energy efficiency projects implemented during the performance period and claimed by another program are prorated based on project completion date then subtracted from whole building savings.

Additionally, savings from non-routine adjustments not associated with the program are removed. Negative savings, regardless if caused by removal of savings claimed in another program or increase in consumption, are set to zero. If the participant renews the contract, future performance periods continue to be compared to the original baseline period unless major changes were made at the site. If significant changes occurred, a new baseline model must be developed after a steady consumption pattern is reestablished.

3.4. Reporting

Several documents are generated in support of each SEM project. At a minimum, these include:

- Site assessment report containing an estimate of potential savings for the site and baseline data that can be used to refine the baseline model.
- ECAM or other baseline model workbooks and supporting data.
- Baseline Model Analysis Report provides the results from the baseline model and recommendation on whether it can confidently predict energy savings from SEM activities.
- M&V Plan outlines a general M&V approach to ensure verifiable site energy savings for the SEM project.
- Custom Project Proposal and Completion Report workbook records information about the site, baseline conditions, implemented SEM measures, and final savings analysis.

4. Evaluation Considerations

The program design contains a number of elements that will contribute to accurate savings claims. However, there may be opportunities for further improving the reliability of savings and for facilitating future impact evaluations. BPA should consider the following while refining the program's design and operation.

4.1. Pre-Implementation Savings Estimates

Key evaluation considerations are:

- Having information about what times or conditions the SEM actions are expected to save energy.
- Providing corroboration of the evaluated savings, or an indication of a need for review of the evaluation calculations.
- Knowing the timing of capital improvements and SEM actions.

Estimates of the expected savings are prepared as part of each project's site assessment. These savings estimates are often based on general rules-of-thumb for the types of SEM actions expected, or just an expected percentage savings from baseline consumption. In some cases, the program may provide more detailed engineering analyses. Similarly, the savings estimates will usually be expected totals for the site, but may include savings for specific opportunities found during the site assessment. The program uses the savings estimates in determining whether the estimated expected savings are larger than the baseline model uncertainty, and then determining whether to proceed with the project.

Detailed engineering estimates are most applicable for capital measures, and estimates may not be easy to make, nor very accurate, for many SEM measures. Some controls measures—such as schedule changes for motors with fixed loads and speeds—may have estimates that are relatively easy to make. However, the energy impact of many types of controls measures, such as changes to pressure setpoints, resets of those setpoints, or changes to VAV box minimum flows, are very difficult to estimate. This challenge is one of the reasons for SEM-type programs for which savings are estimated using regression models rather than engineering calculations.

When more detailed engineering estimates are worthwhile, they could be used as a reasonableness check on the whole building regression estimate of savings.

The routine development of these engineering estimates and gathering information on what measures were implemented and when they became operational will greatly facilitate impact evaluation, especially when they are combined with the program's documentation of when the measures become operational. Knowledge of when program-induced changes occur can help with attribution of total savings to the program, of measure savings to each measure, and identifying non-program related changes in the building. It can help separate savings caused by capital projects from those attributable to SEM.

When engineering and statistical estimates of savings differ by more than the uncertainty of the statistical estimate, the program should have a defined process for trying to reconcile the differences. Such a process might include:

- Review to ensure that the estimates are covering the same measurement boundary.
- Review to ensure that the estimates are covering the same time period.
- Review of the assumptions and calculations in the engineering estimate.
- Review of the statistical model to see whether the data shows other time-related variation in the model residuals, indicating un-modeled changes in the building, and the possible need for a non-routine adjustment.
- If capital measures were implemented, and the savings estimates for a capital measure were from another program, the models and residuals should be analyzed to see whether the data supports the capital measure's savings estimate.
- If the estimates cannot be reconciled to within the uncertainty of the statistical estimate, the statistical estimate may be the best choice. However, additional research is needed to validate this choice.

4.2. Model Type.

Key evaluation considerations are:

- Accurate and transparent estimates of savings and precision.
- Models with physical significance so that results can be better compared with expectations.
- Models that fit industry guidelines or common practice.

The SEM Program will use regression models as described in the BPA *Verification by Energy Modeling Protocol*[1], to estimate savings for each participating building. The Energy Modeling Protocol includes the general procedure, equations, and types of data used, and introduces the subject of uncertainty in savings estimates from these models. The BPA *Regression for M&V: Reference Guide*[2] goes deeper into the equations and uncertainty, and covers some of the principles of ordinary linear regression.

The International Performance Measurement and Verification Protocol (IPMVP)[3] describes two types of savings:

- Avoided Energy Consumption or Demand (also known as Reporting Period Basis)
- Normalized Savings (also known as Fixed Conditions Basis)

The Avoided Energy approach is described in the IPMVP Core Concepts as follows:

---

[1] Verification by Energy Modeling Protocol, Bonneville Power Administration, 2012

[2] Regression for M&V: Reference Guide, Bonneville Power Administration, 2012

[3] Performance Measurement and Verification Protocol Core Concepts, Efficiency Valuation Organization, 2016

> When savings are reported under the conditions of the reporting period, they can also be called savings of the reporting period, or avoided energy consumption. Savings stated as avoided energy consumption quantifies savings in the reporting period relative to what energy would have been without the ECM(s). When reporting savings under reporting period conditions, baseline period energy needs to be adjusted to reporting period conditions. The term forecasting is sometimes used to describe the adjustment of baseline period energy to reporting period conditions.

Normalized savings is described:

> In this method, energy of the reporting period and possibly of the baseline period are adjusted from their actual conditions to the common fixed or normal set of conditions selected. Another term describing the process of stating savings under some different set of conditions than the baseline or reporting period is chaining.

The SEM Program will use the actual time periods after the start of program at a site for savings estimation. Therefore, it will use the Avoided Energy approach.

The "Avoided Energy Use" process is: A model is developed using the baseline period data. An adjusted baseline is created by projecting the baseline model to the post period by using the post conditions. Savings are estimated by subtracting the actual post period energy use from the adjusted baseline. As noted in the IPMVP Core Concepts, this is a "forecast" approach. A forecast approach estimates energy savings for each month (or data interval) in the post period.

The recent evaluation of the BPA Industrial SEM program, documented in *Industrial Energy Management Impact Evaluation*, "used the forecast method as a default, but employed the pre-post method for certain facilities when this method was expected to produce a more accurate savings estimate. The team developed decision logic to determine when to apply the pre-post method…"

According to the BPA Industrial SEM program evaluation report, a pre-post approach "is used extensively in program evaluation…" A pre-post model includes both the baseline and reporting periods in the regression. The energy savings for the reporting period are represented by an indicator variable. So, the regression fits all the data while including an offset that represents the average savings per interval. For an ordinary linear regression, the slope fits both baseline and reporting periods, the intercept fits the baseline data, and the savings indicator adjusts the intercept to fit the reporting period data. This is shown in Figure 4-1 for a synthetic data set.

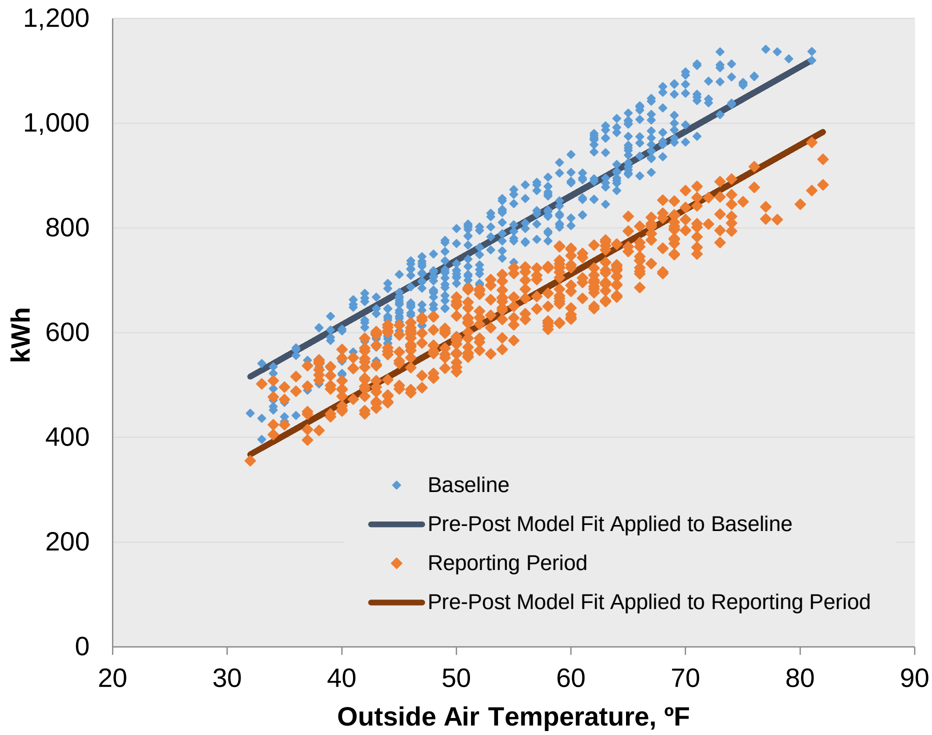


Figure 4-1. Example of Regression Fits from Pre-Post Model

The *Industrial Energy Management Impact Evaluation* used a forecast model for 29 of the 33 sites evaluated.

The forecast method is also expected to be the predominant model type for the Commercial SEM program, for several reasons:

- Most models of commercial buildings can be modeled well using only time and temperature as independent variables. Hence, models will almost always be simpler than the models for industrial applications. In a few situations, such as hotels, occupancy may be an important variable.
- The SEM program will be using a forecast model, and evaluation time will be minimized by using the same type of model.
- Since it estimates savings for each interval, rather than an average, performance can be assessed soon after implementation, whereas the pre-post model requires a full year of data.
- It is arguably more transparent, since it uses actual post period energy use.
- It is consistent with IPMVP and ASHRAE Guideline 14.
- Publicly-available tools for site-specific M&V use a forecast approach.
- The uncertainty in savings is solely the uncertainty in the baseline model's predictions for the adjusted baseline.

It is possible to make a pre-post model perform similarly to a forecast model, but it requires additional interaction variables, making it more complex. The forthcoming *IPMVP Application Guide to Uncertainty Assessment* says, “Including interactions between the ‘post’ indicator variable and explanatory variables is an excellent way to develop normalized impact estimates for more complex model specifications. Unfortunately, the estimation of precision is not straightforward because savings are derived via multiple coefficients.”

The *Superior Energy Performance Measurement and Verification Protocol* for Industry also includes a “backcast” model. This entails modeling the post period, and creating an adjusted post model by projecting the post model back to the baseline period. From *Industrial Energy Management Impact Evaluation*:

> The backcast method can be used when the reporting period is inclusive of baseline period conditions, but not when the reporting period excludes some baseline period conditions. For example, for an industrial facility that only produced low levels of output during the baseline period, but had low and high levels during the reporting period, the backcast model might result in a more accurate estimate of energy savings than the forecast method.

In the scenario described above, the use of a forecast model would require extrapolation that is not necessary if using a backcast model: With commercial buildings, the primary independent variables are almost always temperature (or other weather variables) and time. These do not vary sufficiently year-to-year to require extrapolation. When minor extrapolation is required, e.g., the post period includes record low or high temperatures, the extrapolation will be very minor, and the time period impacted by extrapolation will also be very short, so the energy estimation impact will be tiny. Therefore, backcasting is usually of little utility for commercial buildings.

One possible exception is if the building has inconsistent energy use during the baseline period, but is expected to have more consistent operation during the reporting period, such as might be expected with controls enhancements or retrofits. In that case, the backcast method might be preferred.

Pooled models are often used in residential evaluations. Since commercial buildings are less similar to each other than are homes, pooled models are less valuable. There was no investigation of pooled models in the aforementioned industrial evaluation. Perhaps most importantly, a pooled model would not provide sufficient information about the performance of each building. In a program that pays for measured performance of individual buildings, this is not satisfactory.

Pacific Gas and Electric is relying on forecast models for its pilot Commercial Whole Building performance program. Similarly, California will be relying on forecast models for its coming whole building programs based on AB802. Appendix B to the “Assigned Commissioner and Administrative Law Judge’s Ruling Regarding High Opportunity Energy Efficiency Programs or Projects” listed two forecast models, including ECAM, as “Examples of Implementation.”

4.3. Non-Routine Adjustments.

Key evaluation considerations are:

- Again, knowing the timing of SEM actions.
- Use of statistical methods to estimate non-routine adjustments that can be reviewed by the evaluators.
- Access to engineering calculations of adjustments in same workbooks as regression models.

The site screening procedures are designed to minimize the number of projects that involve non-routine adjustments in the baseline or performance periods. However, the design anticipates that some will still occur. The program could apply the BPA M&V protocols in estimating the impact of these non-routine adjustments, treating them as if they were energy efficiency measures. In particular, the ECwV protocol could provide guidance on how to estimate these impacts without incurring large costs. These calculations could be implemented and documented on additional worksheets in each building's modeling workbook. The adjustments to savings could then be linked to existing savings reporting sheets. This would make the work more transparent and facilitate both BPA QC and impact evaluation.

Perhaps better, where possible, is the use of data analysis to quantify non-routine adjustments. "Where possible" means that the timing of the non-routine change is not concurrent with other changes, and the signal-to-noise ratio in the data is sufficient for the quantification. There are multiple possible ways these changes can be quantified from the data, and they can include analysis of uncertainty. If the program uses these approaches, rather than engineering calculations, it can facilitate evaluation by avoiding the need for reconciling statistical and engineering estimates.

- Look at the time series of residuals for a model that includes the time period of change, and estimate the magnitude of the change from the change in the residuals.
- Use a pre-post model of the period surrounding the adjustment (but confined to the baseline period), treating the time prior to the adjustment as a baseline period and the time after as a post period. The model for this period will require an indicator variable for the post period, and the coefficient on the indicator variable is the required non-routine adjustment. This is simply a more robust method of looking at the time series of residuals.
- Treat the time periods immediately preceding and immediately following the non-routine change as a "mini" baseline and a "mini" post period. These mini periods may cover as much time as possible before and after the non-routine change, but do not go beyond the full baseline or post period within which the non-routine change occurs. Model the change by subtracting the mini post period energy use from an adjusted baseline developed from the mini baseline period.

The first two approaches are likely best if the non-routine adjustment is not expected to have a relationship to weather. The third approach may be best if the non-routine adjustment is expected to be related to weather, but it may suffer from inadequate data in one or both "mini" periods surrounding the non-routine adjustment. All of the methods would benefit from daily or hourly energy use data. Further, they all will need to be appropriately extrapolated to get the effect over a full year.

4.4. Treatment of Capital Measures.

Key evaluation considerations are:

- Knowing the timing of implementation of the capital measure(s).
- Reconciling the capital measure savings estimates with the regression model, regardless of the source of the capital estimates.

The program may integrate capital improvements within the SEM offering or divert them to other program offerings. If integrated, the capital measure savings will be an unknown component of the whole building regression estimate. If handled by other programs, the SEM program will assume that the other program's estimate of savings is correct and deduct it from the whole building regression estimate for the relevant performance period.

The operational date for each capital measure is of concern. More reliable statistical and engineering modeling can be achieved if the program determines when the capital measure becomes operational. If the capital measure savings is being deducted, similar to other non-routine adjustments, the impact should be allocated over time starting from its operational date.

If the program integrates capital measures it may be asked to demonstrate that the embedded savings for these measures is comparable to what would have been reported if the measures had come through other programs. Are they the same as would be estimated using engineering methods that conform to BPA M&V protocols, e.g., using the BPA Lighting Calculator?

The program could test the two approaches on a sample of SEM participants. It would involve developing BPA M&V protocol-compliant estimates of savings for all measures, including important capital measures, then adding up these savings and comparing to the whole building regression estimate. Good test cases would be buildings that have hourly recording meters and where the capital measures were implemented some months before or after the O&M measures. Another possible approach would be to develop whole building regression estimates for buildings in other programs that just implement capital measures. If the two savings estimates are the same or close, this will bolster confidence in both. If they are different, lessons learned in the comparison may lead to improvements in both.

4.5. One Year Measure Life.

A key evaluation consideration is:

- The implications of one-year measure life on cost-effectiveness and incentive levels.

The program plans to associate a one-year measure life with SEM savings. This might be appropriate for some SEM measures, but some SEM measures will persist much longer. If capital measures are integrated with the SEM offering, the reporting period would have to extend throughout the life of the capital measures. Even for O&M measures, if the reporting period was cut short, e.g., only a single year, savings might still accrue after the reporting period has ended, and these savings would not be counted, and hence not contribute to cost-effectiveness. The implications of one-year measure life on cost-effectiveness and incentive levels will be of concern to the participating end users and to an impact evaluation, and should be further investigated.

4.6. Treatment of Negative Savings.

Key evaluation considerations are:

- Clear direction and agreement from BPA on how negative savings estimates are treated.
- Information in program documentation that may explain why savings estimates are negative.

Negative savings are mentioned in multiple places in the *Industrial Energy Management Impact Evaluation*, and are covered in detail in Appendix K. Negative Savings Details in that document. "Negative SEM savings may occur for three reasons... First, there may be an error in the estimated savings. The error can arise in two ways. First, the facility savings estimate is accurate, but the capital project savings are overestimated, causing the SEM savings estimate to become negative. Second, the true facility savings may be positive, but the savings estimate may be negative because of modeling error. Finally, estimated savings may be negative because the implementation of SEM caused the facility to increase consumption."

For some buildings, the whole building regression model may reveal higher use in the performance period than in the forecasted baseline for the performance period. The program plans to set the savings to zero for these buildings. This makes sense in the context of determining the incentive paid to end users, i.e., the incentive cannot be less than 0. However, it is possible that SEM measures suggested by the program will have unintended effects and thus the buildings with negative savings should be investigated to confirm whether certain measures should be changed or removed. It is also possible that there were undetected or incorrectly estimated non-routine adjustments, including savings from capital projects that led to the negative result.

Although it is appropriate to set negative savings to zero in determining end user incentives, it is not appropriate to do this when estimating the total savings from the program. Modeling errors may be present, although perhaps undetected, in the engineering or statistical estimates of savings. These errors can lead to either over or under-estimates of savings. If there are reasons why the savings estimates tend to be biased toward over or under-estimation, these reasons must be fully understood. Once they are understood, empirically grounded techniques should be developed for correcting this bias. In the absence of such empirically grounded adjustments, arbitrarily setting negative savings to zero introduces a bias and results in an overestimate of the savings for the entire program, as reported for the industrial SEM program.

This overestimate is because the bias goes only one way: Sites with negative savings estimates have their savings estimates raised to zero, but sites with high savings estimates do not have their savings estimates decreased.

4.7. Current Practice Baselines

Key evaluation considerations are:

- Program documentation of applicable baseline type
- Additional data may needed by the program and evaluators to model a current practice baseline.

SEM measures could require a current practice baseline, as defined by the RTF. This is more likely to happen if capital measures are integrated in the SEM offering, but could also happen with some O&M measures, such as those that may be required by the Seattle Tune-Up Ordinance. The program's screening process should identify the expected baseline type for the site. The same should be done for measures identified during the performance period whose baselines may be different than the baseline for the site. BPA's ECwV M&V protocol could be followed in estimating non-routine adjustments that would account for the impact of current practice baselines. Some additional data may need to be collected for these measures to support modeling current practice baseline. These data would also be needed for any impact evaluation and might be difficult or impossible to obtain at the time of the evaluation.

California may be including code or standard practice baselines even with AB 802. The CPUC *Staff White Paper on Energy Efficiency Baselines For Implementation of Assembly Bill 802* states that "Key staff recommendations in this paper include:

> The appropriate baseline (i.e., existing conditions or code) can be identified and applied broadly for certain programs, while for other programs the appropriate baseline depends on the measure or other situation-specific conditions."

It further states that "The statute provided the CPUC with discretion to determine the exceptions to the use of an existing conditions baseline. The legislation did not address how the CPUC should treat equipment standards (e.g. Title 20 and Federal standards) in baseline determination or industrial processes where equipment is often installed based on industry standard practice (ISP)."

The industrial program SEM program and evaluation did not address current practice baselines.

These types of baseline adjustment needs cannot be quantified by looking at the data. They would need to be quantified using engineering calculations. The SEM program needs to clearly consider this, and explain if and when current practice baselines apply. Closely related to this consideration is how the savings from capital measures are estimated and treated.

For background, we note that IPMVP only lightly covers the possibility that baselines are not just existing conditions, but can also be standard practice or code.

The older IPMVP 2012 Volume I says,

> "When a certain level of efficiency is required either by law or the facility owner's standard practice, savings may be based on the difference between reporting-period energy and that minimum standard."

The current IPMVP 2016 Final EVO Core Concepts v2 says, in Section 7.1 IPMVP Adherent Plan, Baseline: Period, Usage and Conditions,

> "Significant equipment problems or outages during the baseline period: In some cases, existing systems or facilities may not function properly, meet code, or otherwise may not be reflective of the true baseline conditions. In these cases, the baseline may be adjusted so that it reflects the operation while meeting code or operation after needed repairs."

In Section 7.1 IPMVP Adherent Plan, Basis for Adjustment, it says,

> "Another basis of adjustment is to account for baseline equipment problems or code compliance issues that must be addressed prior to ECM implementation. In these cases, the baseline may be adjusted so that it reflects the operation while meeting code or operation after needed repairs. If the baseline is to be adjusted, include a description of the exact adjustments to the algorithms, variables or terms that affect baseline energy use."

5. Evaluation of Pilot Site Modeling Choices

This section describes our review of the ECAM model creation and analyses developed for the commercial SEM pilot program. We reviewed the models and savings analyses for the seven meters at the three pilot sites.[4]

5.1. Summary

We reviewed the models and savings analyses to examine modeling choices made by BPA engineers. Once familiar with each site's model, we re-modeled each meter from each pilot site in ECAM version 5 using the program's chosen model type. If we thought a different model type might be better, we re-modeled using that type, and compared the results and the uncertainty.

We concluded that the BPA engineers made reasonable choices for all the models. We recommended simpler models for two meters, although one of those recommendations was for a model type that was not available in the earlier versions of ECAM used for the pilot.

Note that the program determines the suitability of the baseline models based on the expected savings: The uncertainty has to be low relative to the expected savings. Evaluation looks at the uncertainty of the baseline model relative to the savings estimated by subtracting reporting period energy use from the adjusted baseline. If the evaluated savings are lower than expected, it is possible that the baseline models are not adequate to state that the savings are statistically different than zero, even though the baseline models were thought to be adequate at time of project initiation.

Furthermore, for some sites, the program models showed bias. This should not be the case. This could have been an ECAM issue, but models should always be evaluated for bias.

We found that the savings could be satisfactorily estimated for all but one meter. The meter with a model inadequate for savings estimation covered only a small part of the site. (More details are provided in the Appendix A. See Site B, Meter 2.)

The program models' uncertainty approached the pre-implementation savings estimates (before efficiency changes are implemented) for Site C, with both meters uncertainty included. With the higher uncertainty in ECAM v5, the uncertainty exceeded the pre-implementation savings estimates. For this site, the pre-implementation savings estimate was about 121,000 kWh. The program models' aggregated uncertainty was about 86,000. The evaluation models' aggregated uncertainty was about 122,000 kWh. So, the program model indicated that the model could estimate savings only with high uncertainty. If the updated uncertainty had been available, the site might have been excluded from the program.

[4] ECAM v5 was used to review these pilot models, and it shows whether the savings were statistically significant relative to the uncertainty. ECAM v5 was used because it has an improved estimate of savings uncertainty. The program had used ECAM v4. It and prior ECAM versions underestimated uncertainty by roughly 41% (√2 -1), although the actual underestimation depends on the length of the baseline and post periods, and the difference in outside air temperature between the two periods.

All other models were very good; any considerations for evaluation would be due to relatively low expected savings. There were two sites that had evaluated savings estimates of only about 0.5%, and the uncertainty was greater than this, so that the range of potential estimates included zero, but we still considered the models useful because the pre-implementation savings estimates were much higher. For example, the pre-implementation savings estimate at Site A was about 280,000 kWh for building 1 at Site A. The evaluated savings estimate was less than 30,000, with an uncertainty of about 84,000. So, the maximum estimated evaluated savings was 114,000. Although we could not conclude that the actual savings were statistically different than zero, we can say that they were far below the pre-implementation estimate.

For one meter, we recommend a different model, because the recommended model was simpler, but resulted in little change in the point estimate of savings and the uncertainty of the savings.

One meter at one site had evidence of the need for a non-routine adjustment. However, with just monthly data, the increase in energy use during the post period was not statistically significant. The range for the estimate of the increase is wide, -40,000 to 580,000 kWh, because there are few data points available for estimating the change, and the exact timing of any non-routine changes was unknown. This range could probably be reduced with information about when improvements to the building actually occurred and when the non-routine event actually occurred, and possibly what the non-routine event actually included.

The estimated impact also includes a major assumption that the energy impact of the non-routine event is not weather-dependent. This assumption probably cannot be verified with monthly data. For sites with shorter interval data, it may be possible to determine whether the impact of the non-routine event is weather-dependent.

This analysis, however, indicates that non-routine building changes can sometimes be detected even with just monthly data. With sufficient information, or more granular data, it may be possible to better estimate the energy impact of those changes.

Further site-specific information on the models is available in Appendix A.

5.2. Recommendations

These recommendations will help ensure that evaluation is efficient by reducing the changes needed to be made to the program’s models and analyses.

- We recommend that engineers use models with the fewest number of parameters that still fit the data.
- We recommend that a model of the reporting period data always be created.
- Modelers should ensure that significant modeling choices are traceable.
- Analysts should document analyses and decision-making rationales.

An explanation of these recommendations follows.

- We recommend that engineers use models with the fewest number of parameters that still fit the data.

  While all of the model choices were reasonable, for two meters it was possible to use models with fewer parameters while still maintaining accuracy. Simpler models, where appropriate, may provide better normalization to typical conditions, and better extrapolation where needed due to different conditions (e.g. weather) between baseline and post periods.[5] Models with fewer parameters may also have lower uncertainty, if there is not a significant increase in Root-Mean-Square-Error (RMSE).

  In support of this recommendation, we believe that users could benefit from clearer information in ECAM about the significance of parameters.

- We also recommend that a model of the reporting period data always be created.

  This is the same type of forecast model as used for the baseline, but it uses reporting period temperatures and energy use. We are not aware of this recommendation having been made elsewhere for IPMVP "avoided energy use" type of savings estimates. There are multiple, related reasons to create a reporting period model. Such a model is valuable to:

  - Observe trends in energy use in the post period.
  - Visually align the timing of changes visible in the model with known changes in the building attributable to the program.
  - Look for non-routine events or gradual changes in energy use.
  - Identify whether any change points are different between baseline and post. Does any observed movement in the change point(s) make sense?

  A reporting period model is especially valuable for an SEM program if the SEM-related changes are documented with the dates they were made. This allows the benefit of changes with significant impacts to be observed in the data as per the second bullet above.

- Modelers should ensure that significant modeling choices are traceable.

  Keep raw (unprocessed) data files (interval, billing, weather) and provide to reviewers along with the baseline and adjusted baseline models

  Site B used weather from more than one site, and we could not determine the sources. It did not have a significant effect on the savings, since the energy use did not have a strong dependency on weather, but to facilitate evaluation, this type of choice should be documented with the data sources and data available, along with the reasons a choice was made. Without the raw weather data, it was difficult to assess how much

[5] Whenever possible, baseline periods should cover the full range of operating conditions present in the reporting period. Since baseline periods are chosen before the reporting period, occasionally some extrapolation will be needed.

missing data was replaced from a different station and what the impact was on the model.

- Analysts should document analyses and decision-making rationales

  Documentation should include why monthly data was chosen instead of interval data. Also, why a particular weather station was chosen and how much of the missing data was replaced with data from a different station, and for what dates.

  Evaluation may need additional information to understand pre-implementation savings calculations, or they may need explanation from the engineer that developed those calculations. Such clarifications may be needed for evaluators to understand, and subsequently explain to others, the difference in pre-implementation and ex post savings estimates. The need for additional information is dependent on:

  - The quality of the regression model,
  - The uncertainty of the estimates relative to the magnitude of the pre-implementation estimates, and
  - Any trends in energy use or the timing of changes in energy use.

  Standardized residuals can help engineers and analysts recognize changes in a building's energy use. Standardized residuals—residuals divided by their standard errors—are similar to standard deviations, and provide the *significance* of a difference between modeled and actual values. Therefore, these can be used to identify significant changes in energy use, and when those changes occur. A heat map[6] based on the standardized residuals can assist by visually showing persistent changes in energy use.

[6] A heat map is a graph that uses a range of colors to represent the different ranges of values in the data. An example is the common display of temperatures in televised weather reports.

Appendix A. Details of Site-Specific Model Evaluations

For all sites, the uncertainty estimated by the evaluation is higher than the uncertainty estimated by the program, because of the changes in uncertainty calculations from ECAM version 4 to version 5. In the tables below, all savings and uncertainty values are in kWh or percentages.

The tables show the results for both the program models and the evaluation models.

The charts show only the evaluation models unless the program used a different model specification or there was a visual difference between the program and evaluation models.

A.1. Site A

Meter 1, Main Site

The program used a 4-parameter model and we considered both 4-parameter and 3-parameter models. Either one would be a reasonable choice, with the 3-parameter having just slightly higher uncertainty. For both models, the uncertainty of the savings was close to 1.5% of baseline, and the evaluated savings were statistically insignificant relative to that uncertainty: The range of savings included zero.

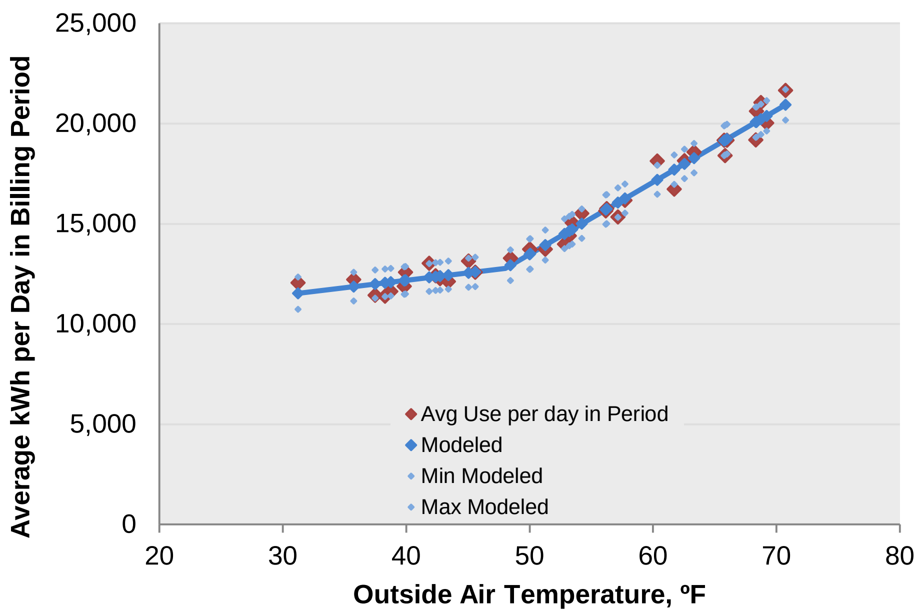
	281,156	Pre-implementation Savings Estimate (for total Site)

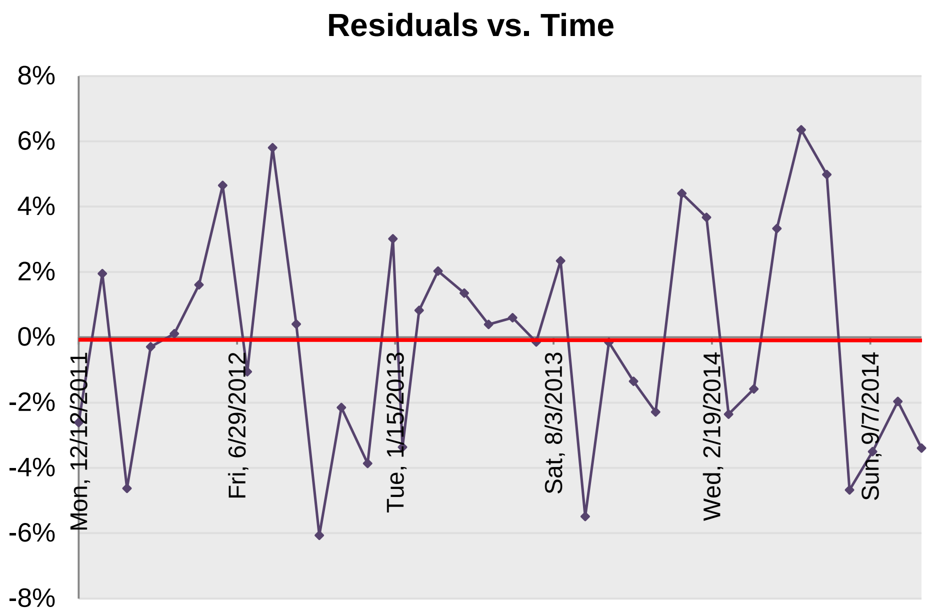
4p program	5,778,670	Projected Baseline Energy
CP=48.0	5,748,876	Measured Energy
	29,794	Avoided Energy Use
	42,592	Projected Baseline ±Uncertainty @ 80% Confidence Level
	29,794 ±42,592	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.7%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0.5% ±0.7%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

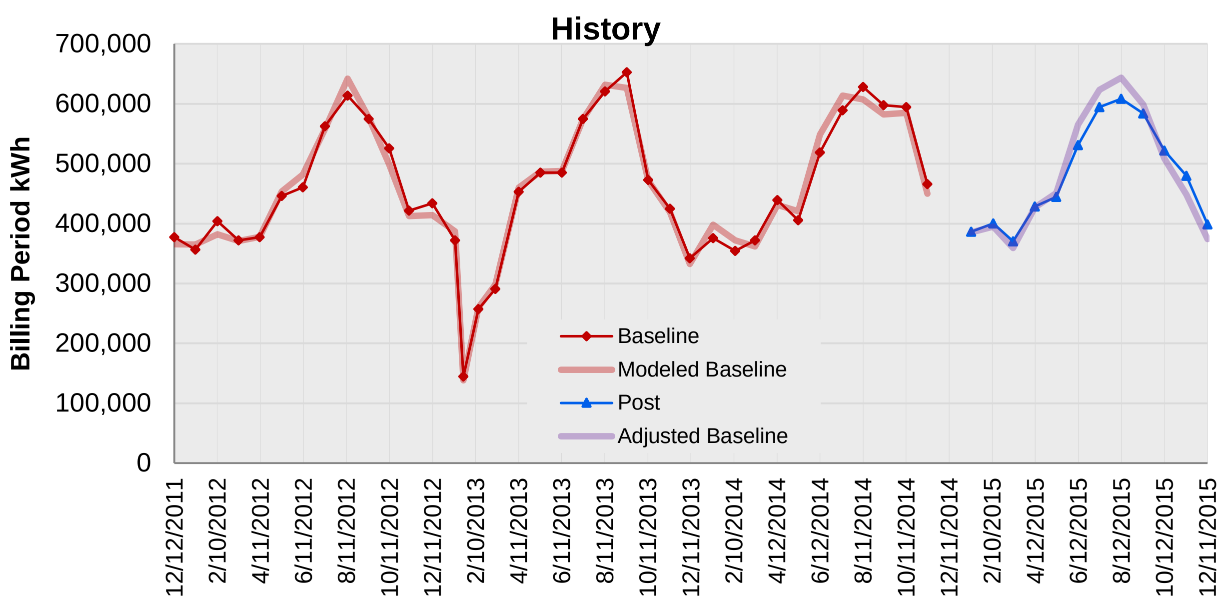
4p evaluation CP=48.0	5,778,525	Adjusted Baseline Energy
	5,748,876	Measured Energy
	29,649	Avoided Energy Use
	83,558	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	29,649 ±83,558	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.4%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.5% ±1.4%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

3p evaluation CP=46.4	5,761,345	Adjusted Baseline Energy
	5,748,876	Measured Energy
	12,469	Avoided Energy Use
	86,905	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	12,469 ±86,905	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.5%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.2% ±1.5%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

4p Evaluation Model




Residuals vs. Time
8%
6%
4%
2%
0%
-2%
-4%
-6%
-8%
Mon, 12/12/2011
Fri, 6/29/2012
Tue, 1/15/2013
Sat, 8/3/2013
Wed, 2/19/2014
Sun, 9/7/2014


History
Billing Period kWh
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0
Baseline
Modeled Baseline
Post
Adjusted Baseline
12/12/2011
2/10/2012
4/11/2012
6/11/2012
8/11/2012
10/11/2012
12/11/2012
2/10/2013
4/11/2013
6/11/2013
8/11/2013
10/11/2013
12/11/2013
2/10/2014
4/12/2014
6/12/2014
8/11/2014
10/11/2014
12/11/2014
2/10/2015
4/12/2015
6/12/2015
8/12/2015
10/12/2015
12/11/2015

Meter 2, Building 2

The program used a 2-parameter (linear) model and we considered both 2-parameter and 1-parameter (mean) models, since the linear model had a slope approaching zero. For both models, the uncertainty of the savings was close to 0.8% of baseline, and the evaluated savings were statistically insignificant relative to that uncertainty. The program did not have a mean model to choose, since it was not available in ECAM at the time. The evaluation used the mean model since the slope parameter was not significant.

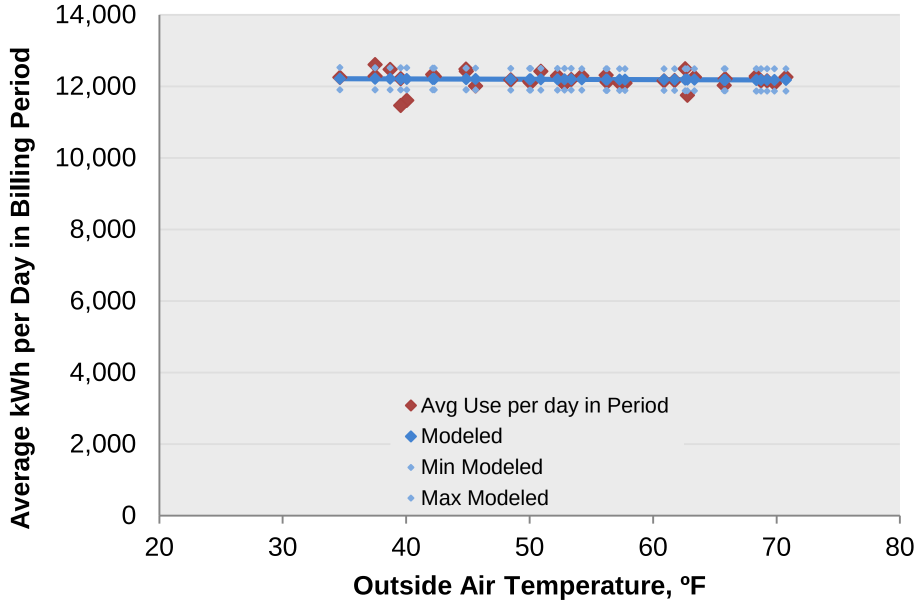
		no pre-implementation estimate of savings separate from total Site

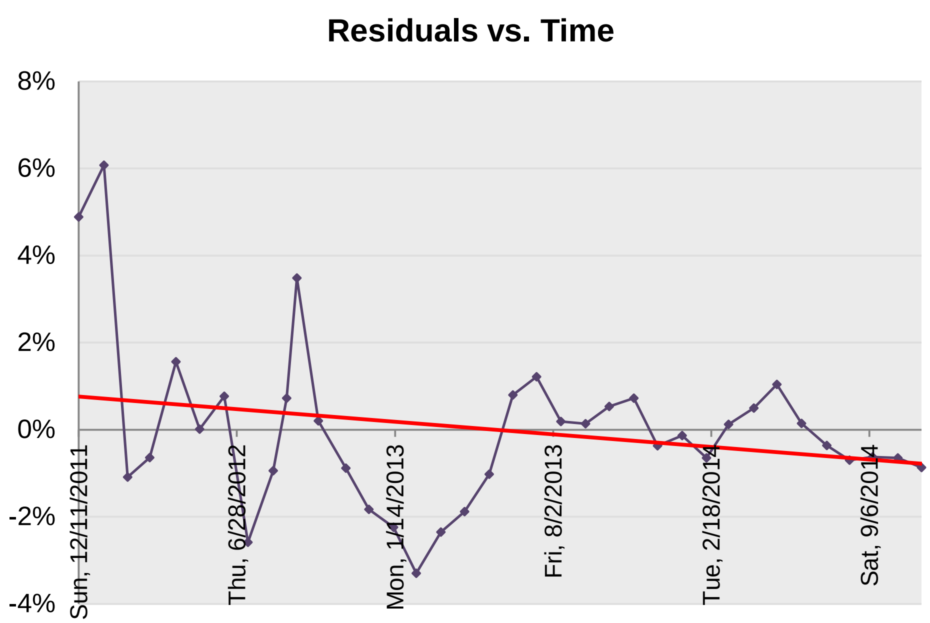
2p Program	4,447,723	Projected Baseline Energy
	4,428,400	Measured Energy
	19,323	Avoided Energy Use
	15,649	Projected Baseline ±Uncertainty @ 80% Confidence Level
	19,323 ±15,649	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.4%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0.4% ±0.4%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.05%	Model Bias

2p Evaluation	4,450,047	Adjusted Baseline Energy
	4,428,400	Measured Energy
	21,647	Avoided Energy Use
	36,376	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	21,647 ±36,376	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.8%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.5% ±0.8%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

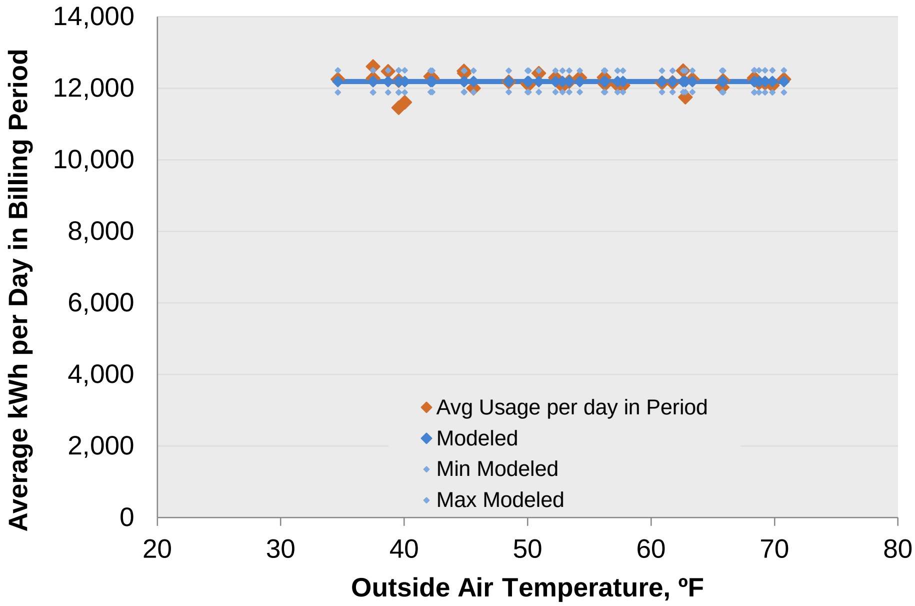
1p Evaluation	4,451,068	Adjusted Baseline Energy
	4,428,400	Measured Energy
	22,668	Avoided Energy Use
	35,899	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	22,668 ±35,899	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.8%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.5% ±0.8%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

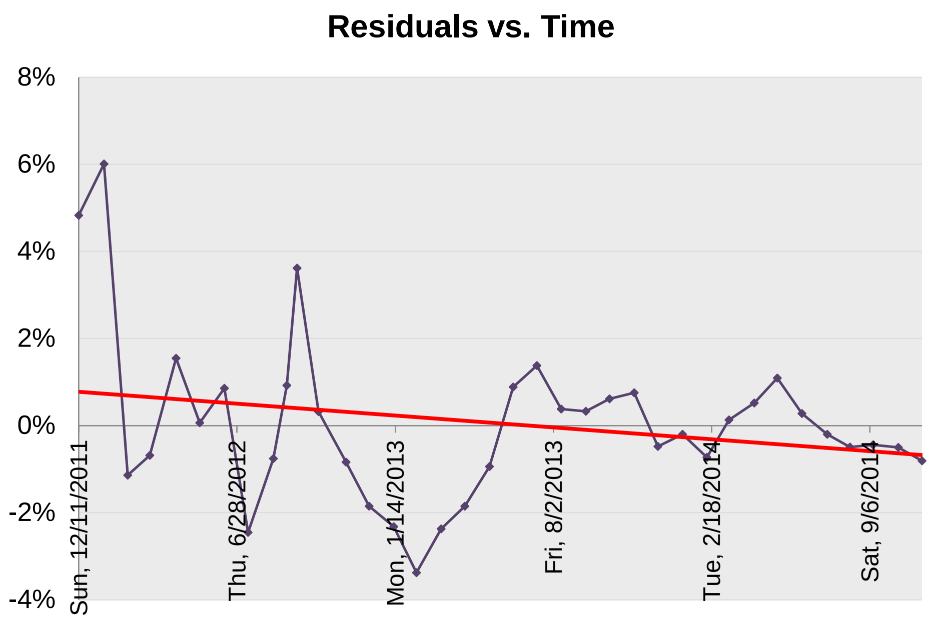
2p Program Model

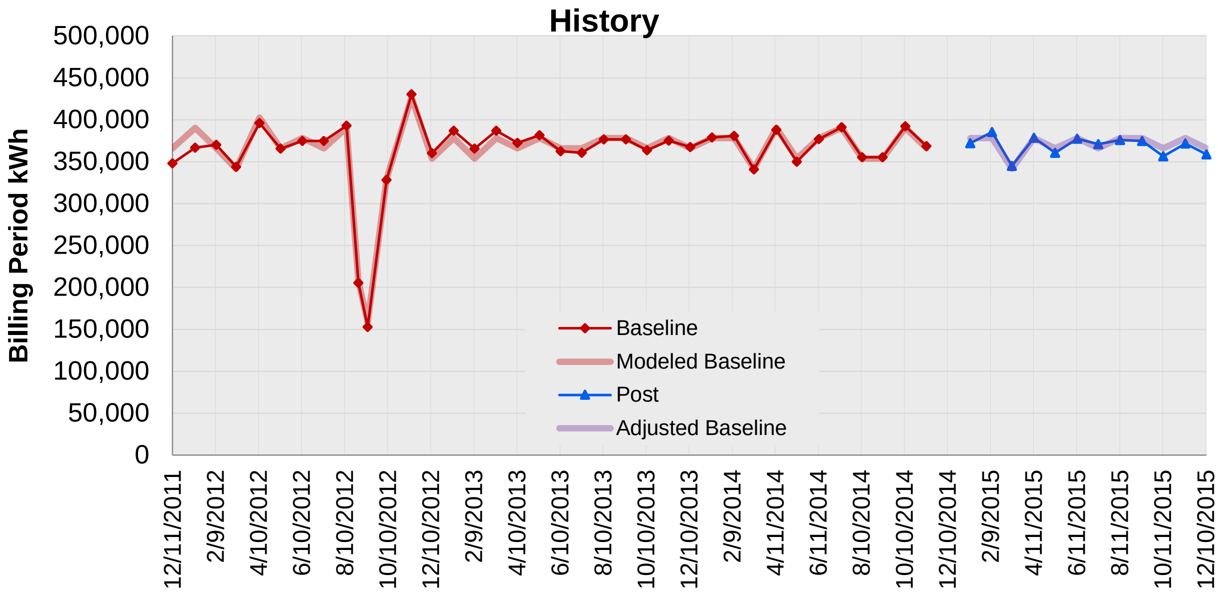




1p Evaluation Model






History
Billing Period kWh
500,000
450,000
400,000
350,000
300,000
250,000
200,000
150,000
100,000
50,000
0
Baseline
Modeled Baseline
Post
Adjusted Baseline
12/11/2011
2/9/2012
4/10/2012
6/10/2012
8/10/2012
10/10/2012
12/10/2012
2/9/2013
4/10/2013
6/10/2013
8/10/2013
10/10/2013
12/10/2013
2/9/2014
4/11/2014
6/11/2014
8/10/2014
10/10/2014
12/10/2014
2/9/2015
4/11/2015
6/11/2015
8/11/2015
10/11/2015
12/10/2015

Meter 3, Building 3

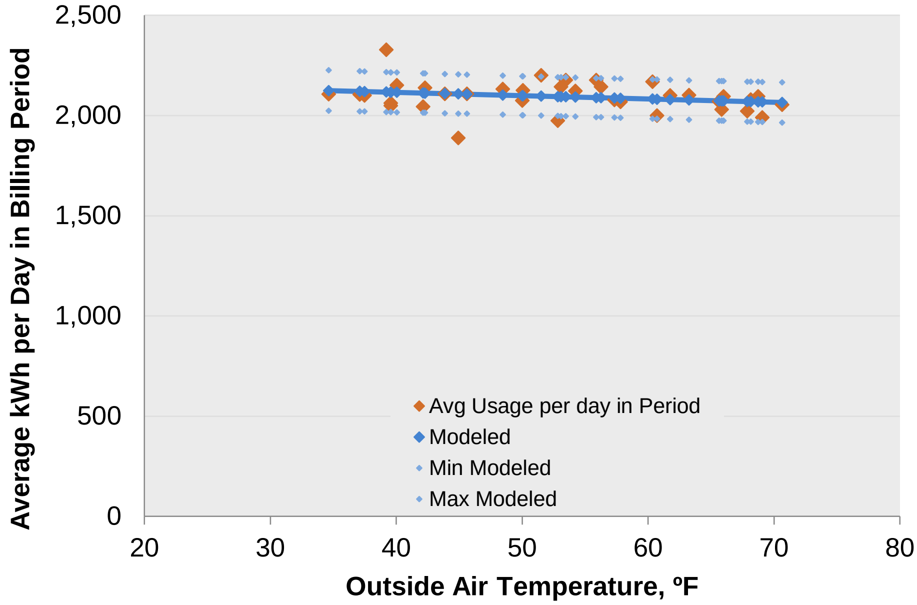
The program and our review both chose 2-parameter models. A 1-parameter (mean) model could also have been a reasonable choice, since the slope parameter was barely significant, but that choice was not available to the program at the time. There was a statistically significant increase in energy use. The M&V 'savings' were -4.7% ±1.5%.

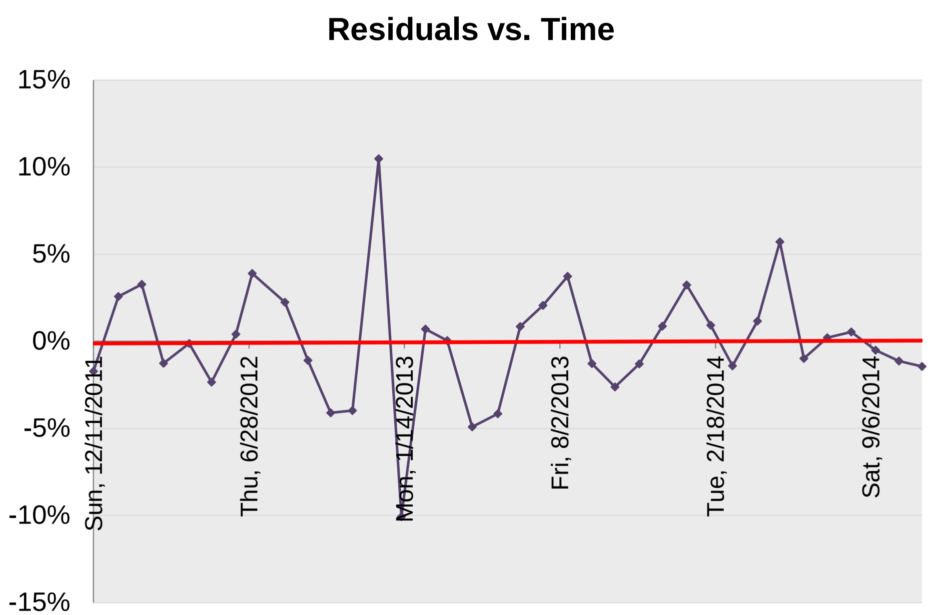
		no pre-implementation estimate of savings separate from total Site

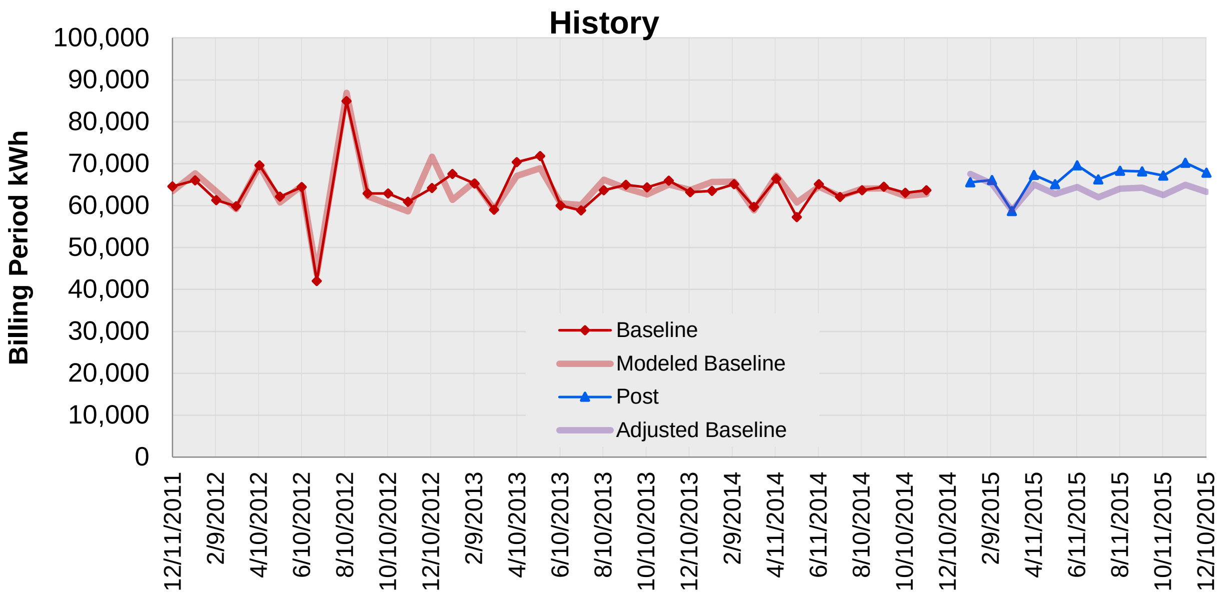
2p Program	765,179	Projected Baseline Energy
	800,480	Measured Energy
	-35,301	Avoided Energy Use
	5,152	Projected Baseline ±Uncertainty @ 80% Confidence Level
	-35,301 ±5,152	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.7%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	-4.6% ±0.7%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.045%	Model Bias

2p Evaluation	764,832	Adjusted Baseline Energy
	800,480	Measured Energy
	-35,648	Avoided Energy Use
	11,755	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	-35,648 ±11,755	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.5%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	-4.7% ±1.5%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

2p Evaluation Model






History
Billing Period kWh
100,000
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
12/11/2011
2/9/2012
4/10/2012
6/10/2012
8/10/2012
10/10/2012
12/10/2012
2/9/2013
4/10/2013
6/10/2013
8/10/2013
10/10/2013
12/10/2013
2/9/2014
4/11/2014
6/11/2014
8/10/2014
10/10/2014
12/10/2014
2/9/2015
4/11/2015
6/11/2015
8/11/2015
10/11/2015
12/10/2015
Baseline
Modeled Baseline
Post
Adjusted Baseline

A.2. Site B

There are numerous things to comment about for this site.

- weather data was sourced from more than one station
- version of ECAM being used changed between initial baseline analysis and savings analysis, with different results
- analyst chose model based on monthly data when interval data was available

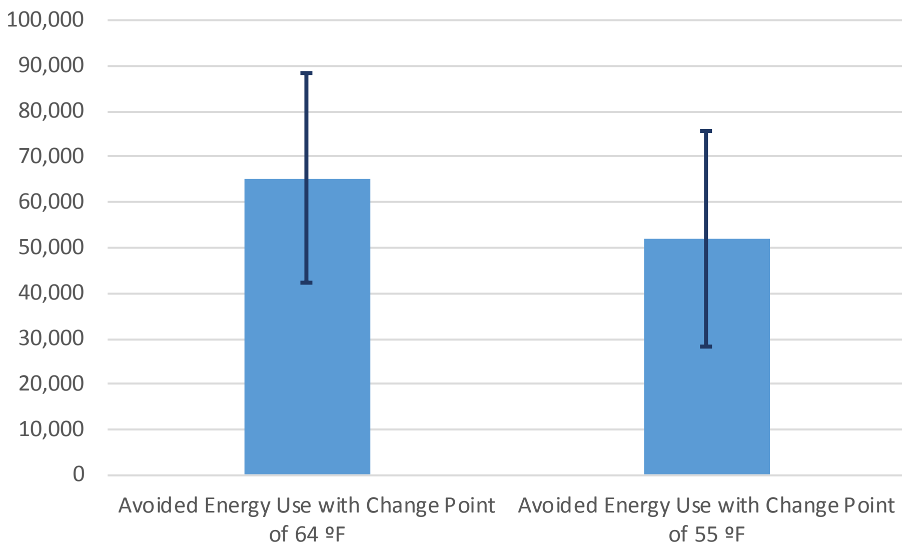
Note, the need for documenting reasons for choosing specific weather data is recommended in the body of this report.

The change in ECAM results was a concern. There was not a statistically-significant impact on the savings and uncertainty, but the change point was quite different. However, this issue is unlikely to repeat.

The building energy use does not appear to have strong weather dependency. The primary impact on modeling by the change in ECAM versions was a substantial shift in the change point temperature. ECAM v3 initially divided the range of x-values into 12 bins to find the x-value range containing the change point. ECAM v4 and v5 used a finer initial grid, dividing the range of x-values into 20 bins before searching for the exact change point. This had an almost negligible impact on the uncertainty (from 23,815 to 23,040). The big shift in balance point with minimal impact on uncertainty indicates low weather dependence. The point estimate of savings changed by 26%, from 51,981 kWh with the 55 ºF balance point to 65,324 kWh with the 64 ºF balance point. However, this is not a significant change because it is still within the error band. It is still potentially problematic because a single value is reported for savings and the models show essentially the same amount of uncertainty for both of these savings results.

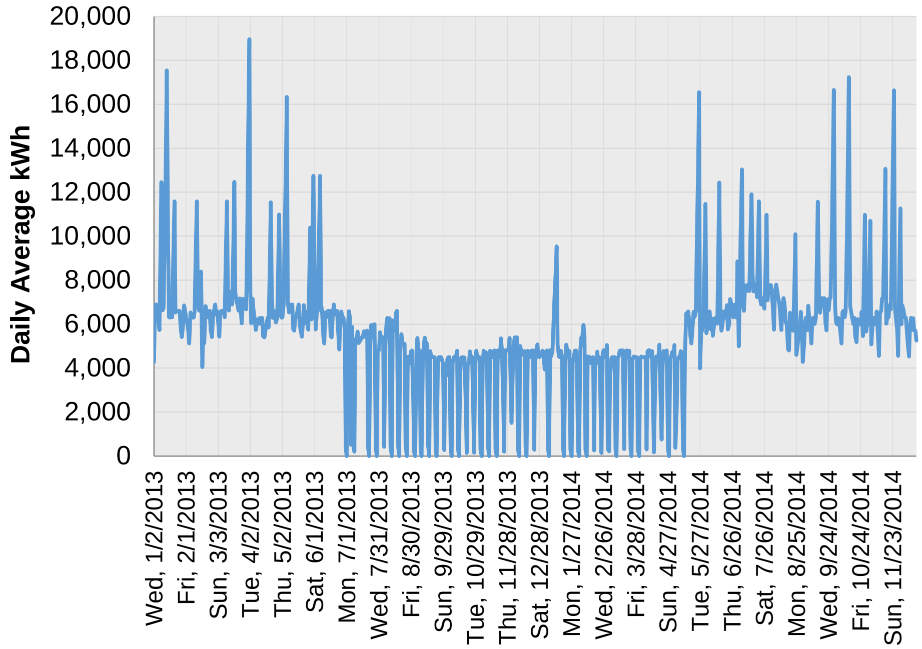
So, which is the correct answer? Visualizing the model and data in a scatter chart, the earlier estimate of a 55 ºF balance point appears superior. It appears in the middle of the range of possible balance points for which the RMSE varies relatively little. However, the later estimate of 64 ºF provides the best fit, with the lowest RMSE and lowest estimate of savings uncertainty. Subsequent analyses of uncertainty and maximum likelihood change point using a bootstrap analysis indicated that the 55 ºF balance point is the best choice.

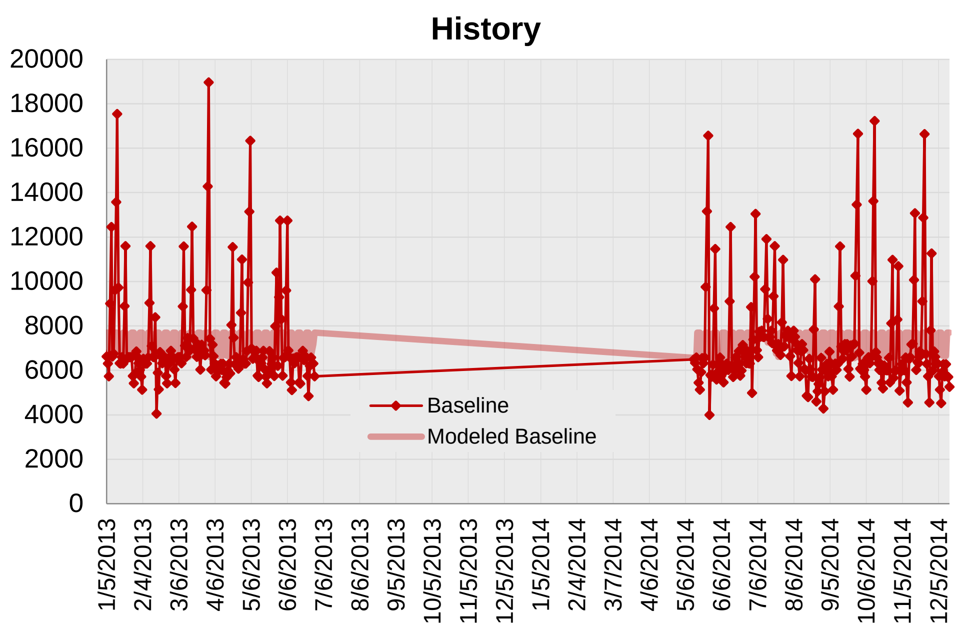
	Savings	Uncertainty
Avoided Energy Use with Change Point of 64 ºF	65,324	23,040
Avoided Energy Use with Change Point of 55 ºF	51,981	23,815



After investigating the choice to use monthly billing data instead of interval data, we reached the following conclusions:

- The interval data provided did not extend into the post period and at first appeared highly erratic, with monthly totals not matching the data used in the monthly model
  - Found that the interval data included both meters.
  - The interval (daily) data did not correlate as well with weather data as the monthly data did, and other building operations information (occupancy, schedules, etc.) was not available to provide explanatory variables to improve the baseline model.
  - The interval data had lots of bad data in the middle of the baseline period.
- Agree with program's finding that, given available data and information, monthly data was better to use than interval. All statistics, including regression statistics and final estimates of uncertainty, supported this.


Daily Average kWh
20,000
18,000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
Wed, 1/2/2013
Fri, 2/1/2013
Sun, 3/3/2013
Tue, 4/2/2013
Thu, 5/2/2013
Sat, 6/1/2013
Mon, 7/1/2013
Wed, 7/31/2013
Fri, 8/30/2013
Sun, 9/29/2013
Tue, 10/29/2013
Thu, 11/28/2013
Sat, 12/28/2013
Mon, 1/27/2014
Wed, 2/26/2014
Fri, 3/28/2014
Sun, 4/27/2014
Tue, 5/27/2014
Thu, 6/26/2014
Sat, 7/26/2014
Mon, 8/25/2014
Wed, 9/24/2014
Fri, 10/24/2014
Sun, 11/23/2014


History
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Baseline
Modeled Baseline
1/5/2013
2/4/2013
3/6/2013
4/6/2013
5/6/2013
6/6/2013
7/6/2013
8/6/2013
9/5/2013
10/5/2013
11/5/2013
12/5/2013
1/5/2014
2/4/2014
3/7/2014
4/6/2014
5/6/2014
6/6/2014
7/6/2014
8/6/2014
9/5/2014
10/6/2014
11/5/2014
12/5/2014

Meter 1, Main Meter

The program and our review both used 4-parameter models. Evaluated savings were statistically significant at 3.1% ±1.6%.

This site had interval meter data as well as the billing data. However, the interval data had issues, so the program chose to use the monthly data. We reviewed the interval data and concurred with the decision. Further information on this is available in the appendix.

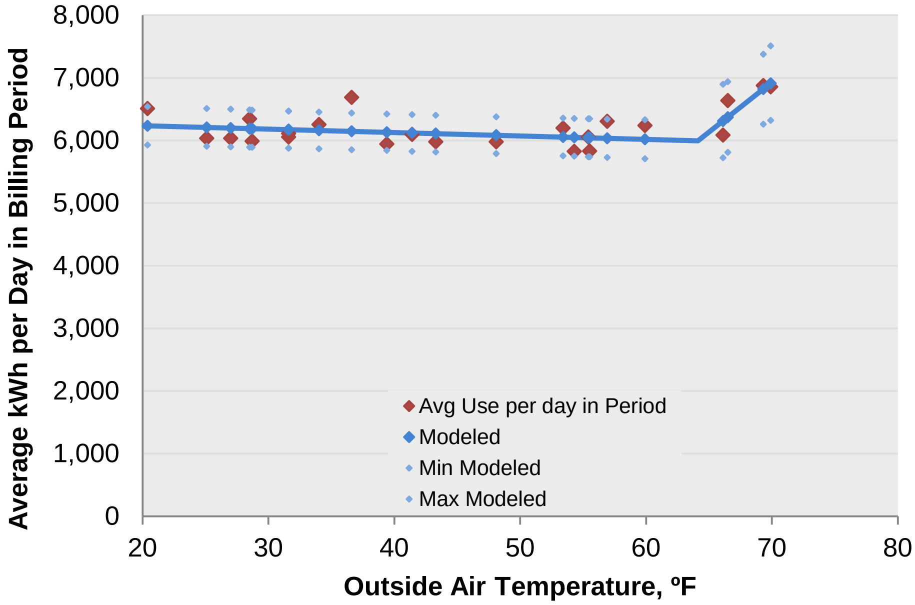
	45,000	Pre-implementation Savings Estimate

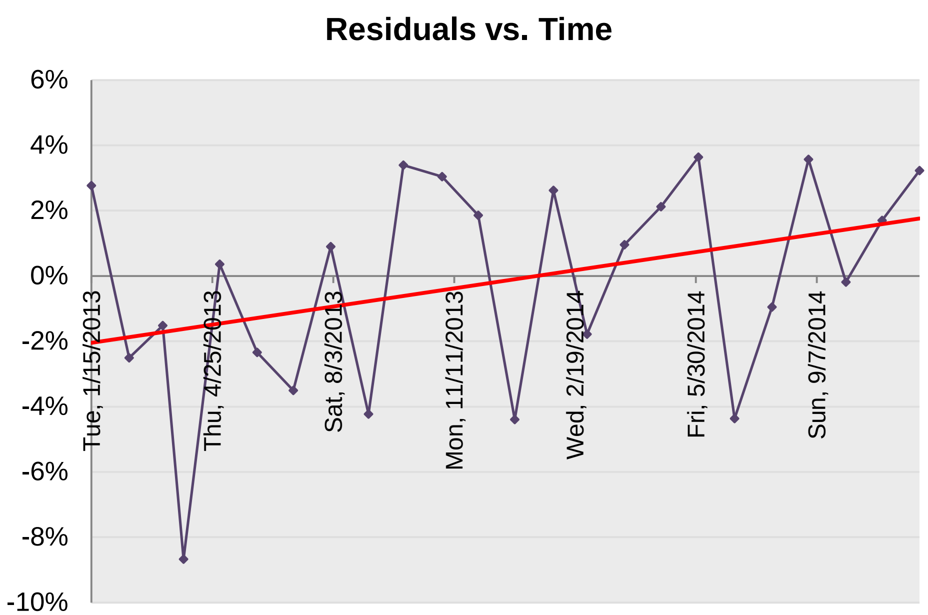
4p Program	2,278,759	Projected Baseline Energy
CP=64.14	2,209,000	Measured Energy
	69,759	Avoided Energy Use
	23,040	Projected Baseline ±Uncertainty @ 80% Confidence Level
	69,759 ±23,040	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.0%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	3.1% ±1.0%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

4p Evaluation	2,278,751	Adjusted Baseline Energy
CP=64.14	2,209,000	Measured Energy
	69,751	Avoided Energy Use
	37,542	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	69,751 ±37,542	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.6%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	3.06% ±1.65%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

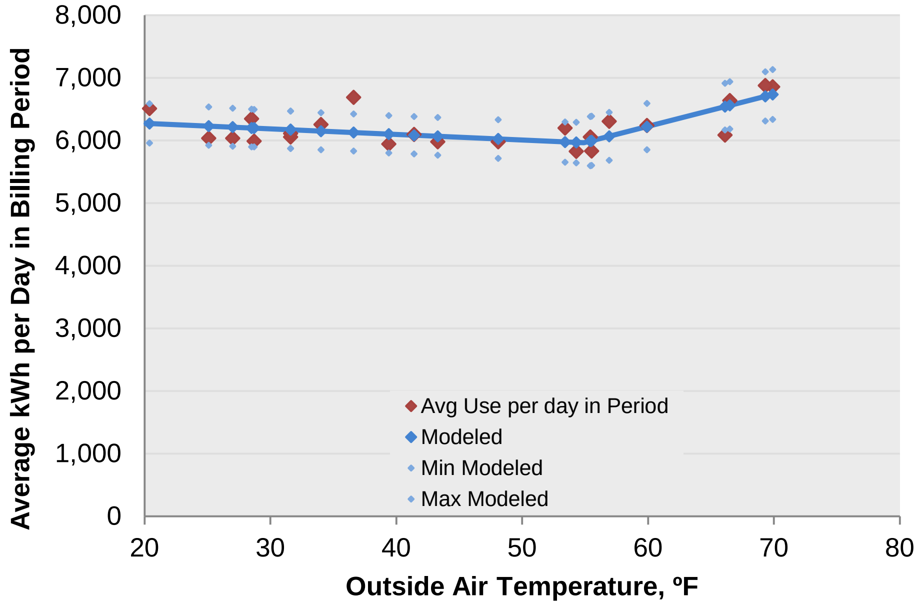
4p Evaluation	2,278,316	Adjusted Baseline Energy
CP=54.89	2,209,000	Measured Energy
	69,316	Avoided Energy Use
	38,186	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	69,316 ±38,186	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.7%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	3.04% ±1.68%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

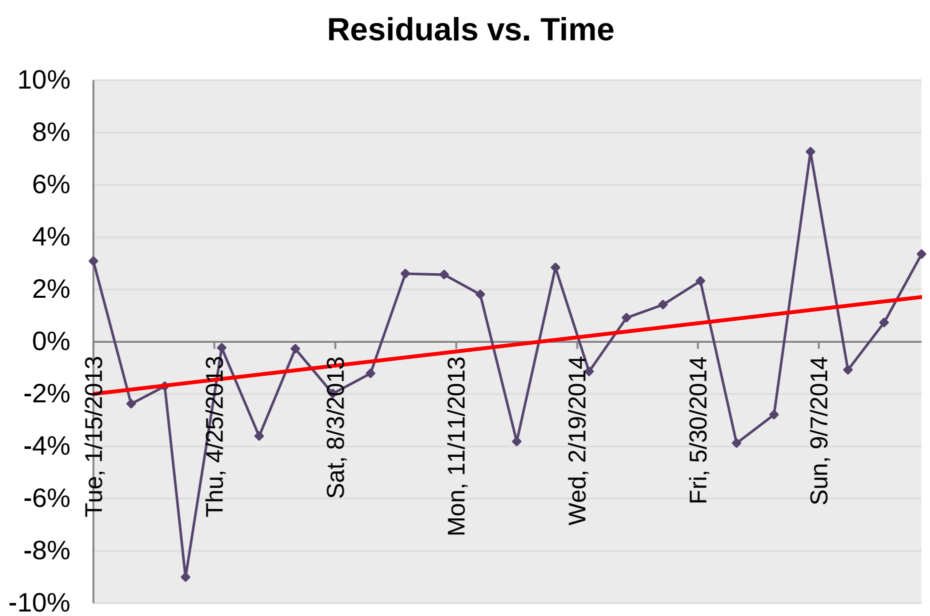
4p Program Model, 64.1 ºF Change Point Temperature

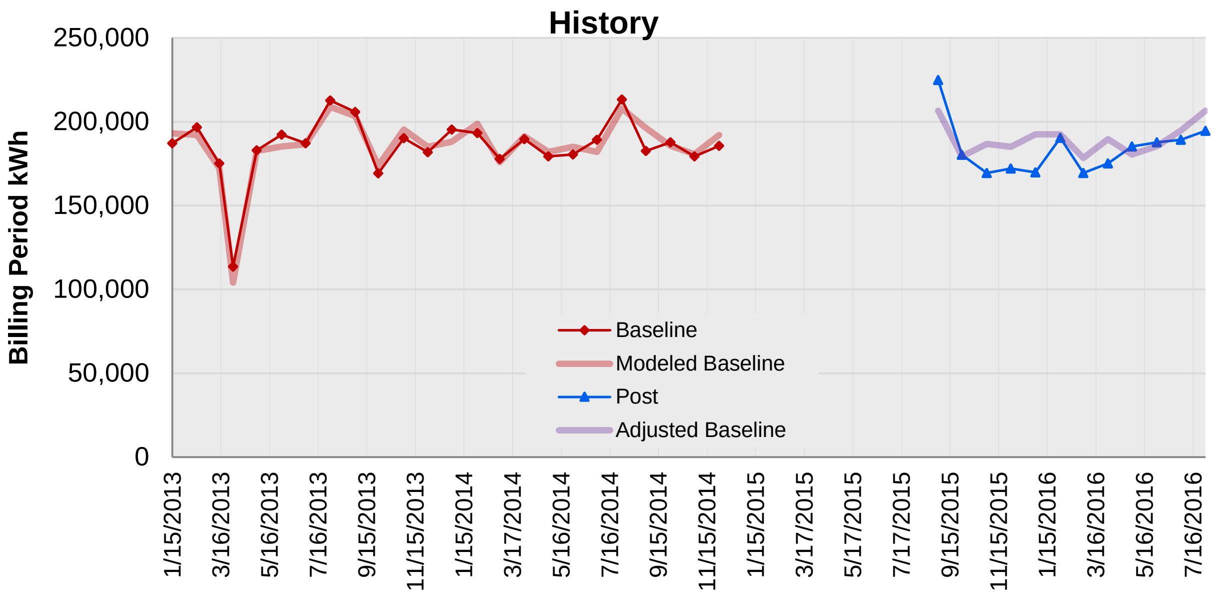




4p Evaluation Model, 54.9 ºF Change Point Temperature






History
Billing Period kWh
250,000
200,000
150,000
100,000
50,000
0
Baseline
Modeled Baseline
Post
Adjusted Baseline
1/15/2013
3/16/2013
5/16/2013
7/16/2013
9/15/2013
11/15/2013
1/15/2014
3/17/2014
5/16/2014
7/16/2014
9/15/2014
11/15/2014
1/15/2015
3/17/2015
5/17/2015
7/17/2015
9/15/2015
11/15/2015
1/15/2016
3/16/2016
5/16/2016
7/16/2016

Meter 2, Expansion Meter

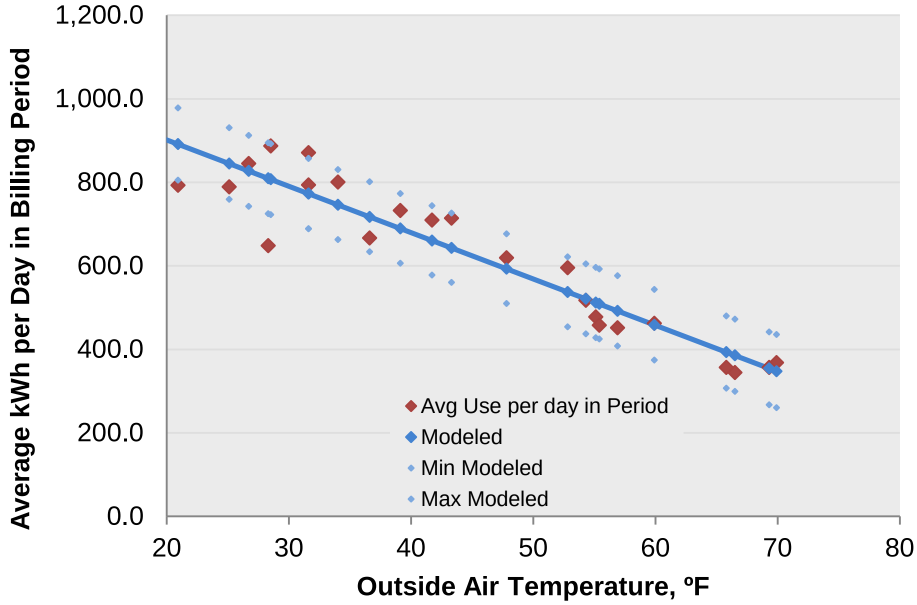
The program and our review both used 2-parameter (linear) models. Evaluated savings were not statistically significant at 4.2% ±4.7%. We note that there was a significant–nearly 10%–decrease in temperature-adjusted energy use over the baseline period, which means that the real savings due to the program were even less.

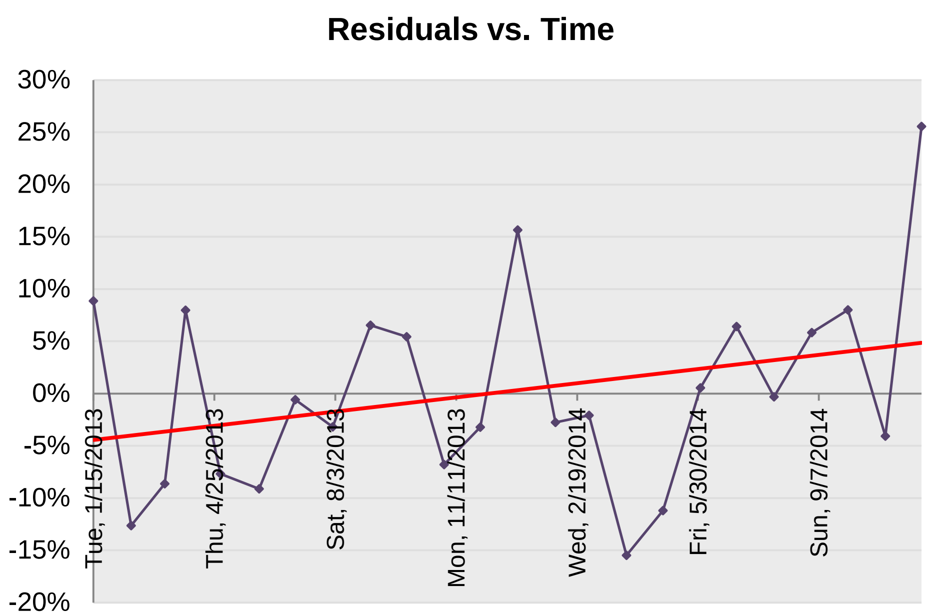
		no pre-implementation estimate of savings separate from total Site B

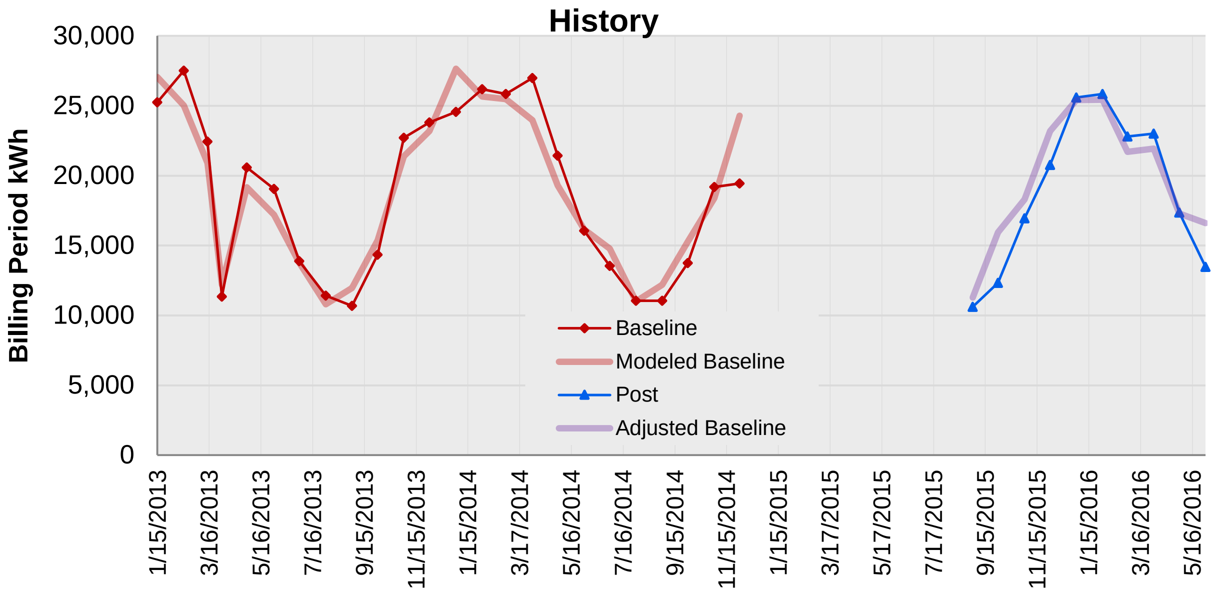
2p Program	196,869	Projected Baseline Energy
	188,791	Measured Energy
	8,078	Avoided Energy Use
	4,238	Projected Baseline ±Uncertainty @ 80% Confidence Level
	8,078 ±4,238	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	2.2%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	4.1% ±2.2%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	-0.13%	Model Bias

2p Evaluation	197,115	Adjusted Baseline Energy
	188,791	Measured Energy
	8,324	Avoided Energy Use
	9,334	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	8,324 ±9,334	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	4.7%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	4.2% ±4.7%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

2p Evaluation Model






History
Billing Period kWh
30,000
25,000
20,000
15,000
10,000
5,000
0
1/15/2013
3/16/2013
5/16/2013
7/16/2013
9/15/2013
11/15/2013
1/15/2014
3/17/2014
5/16/2014
7/16/2014
9/15/2014
11/15/2014
1/15/2015
3/17/2015
5/17/2015
7/17/2015
9/15/2015
11/15/2015
1/15/2016
3/16/2016
5/16/2016
Baseline
Modeled Baseline
Post
Adjusted Baseline

A.3. Site C

Meter 1, Main Meter

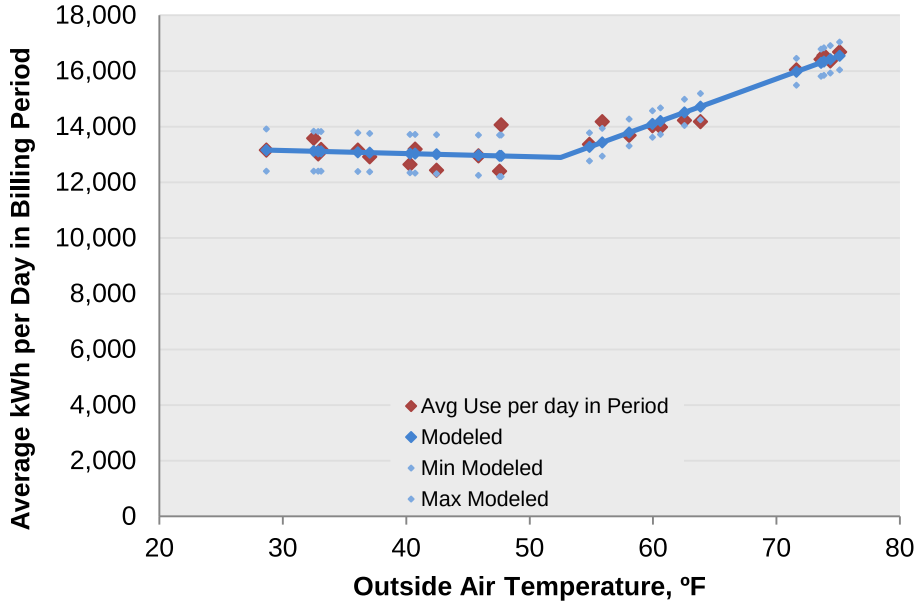
The program and our review both chose 4-parameter models. The uncertainty was 1.1% of baseline consumption. There was no estimate of savings because there was no data for the reporting period. We assumed that the project was not completed.

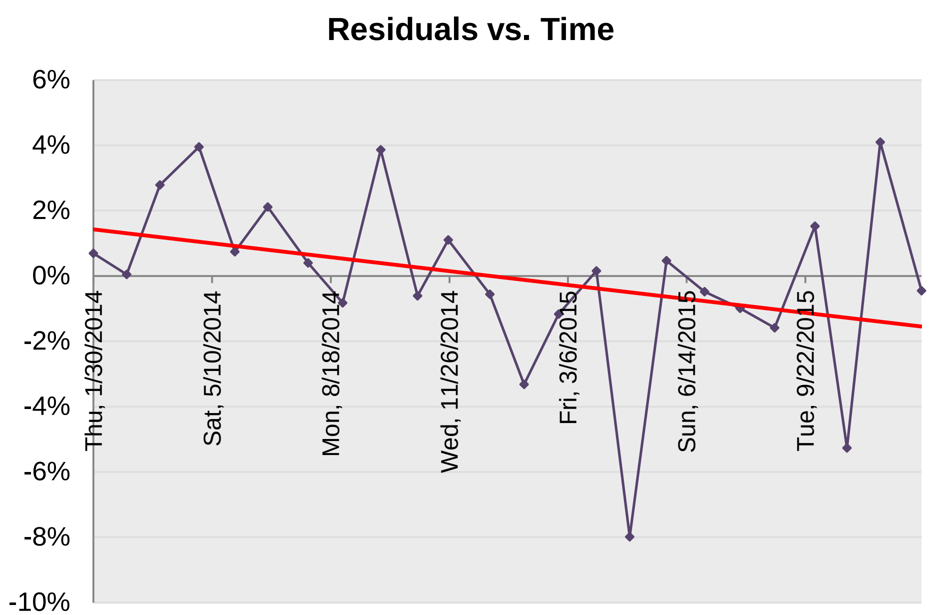
	120,973	Pre-implementation Savings Estimate

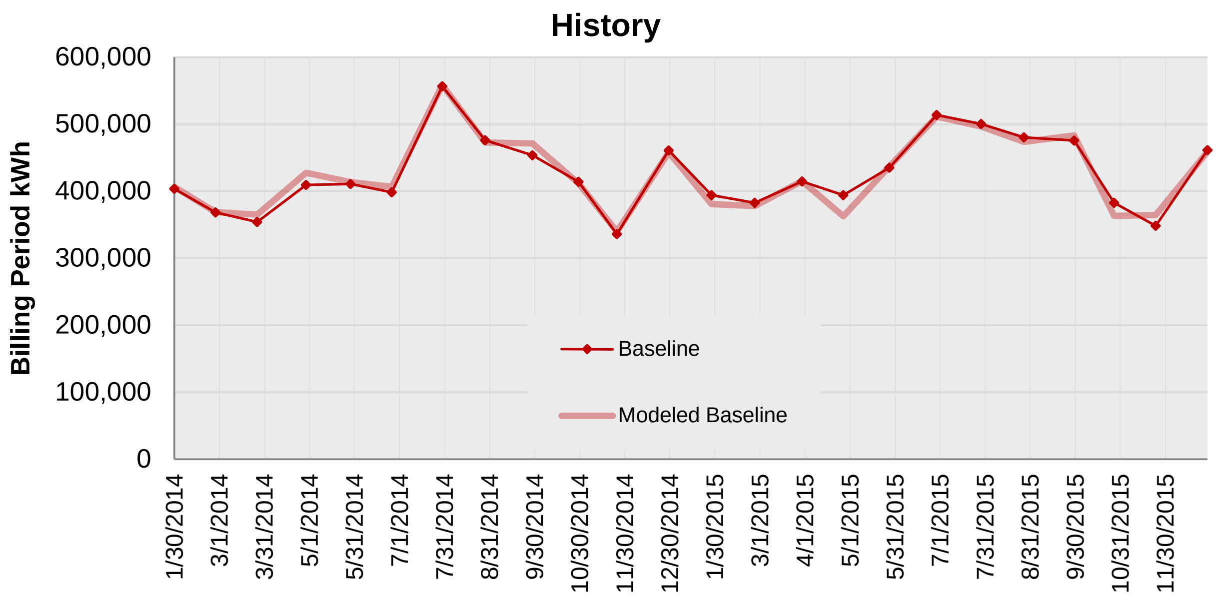
4p Program	10,222,600	Projected Baseline Energy
Baseline only	10,222,600	Measured Energy
	0	Avoided Energy Use
	81,555	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0 ±81,555	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.8%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±0.8%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

4p Evaluation	10,222,600	Adjusted Baseline Energy
Baseline only	10,222,600	Measured Energy
	0	Avoided Energy Use
	115,359	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0 ±115,359	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.1%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±1.1%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

4p Evaluation Model






History
Billing Period kWh
600,000
500,000
400,000
300,000
200,000
100,000
0
1/30/2014
3/1/2014
3/31/2014
5/1/2014
5/31/2014
7/1/2014
7/31/2014
8/31/2014
9/30/2014
10/30/2014
11/30/2014
12/30/2014
1/30/2015
3/1/2015
4/1/2015
5/1/2015
5/31/2015
7/1/2015
7/31/2015
8/31/2015
9/30/2015
10/31/2015
11/30/2015
Baseline
Modeled Baseline

Meter 2, Sub Meter

The program and our review both chose 4-parameter models. The uncertainty was 1.1% of baseline consumption. There was no estimate of savings because there was no data for the reporting period. We assumed that the project was not completed.

		no pre-implementation estimate of savings separate from total Hospital

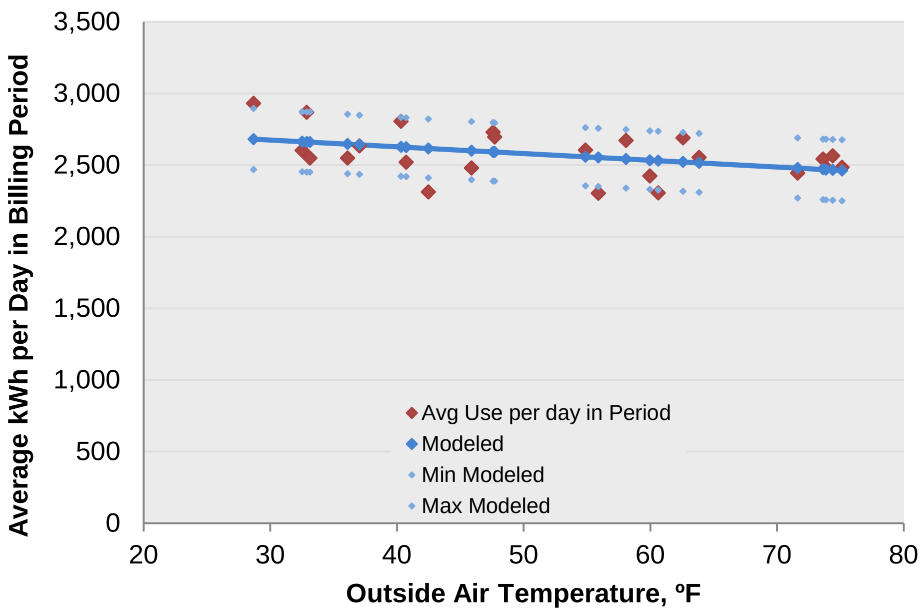
4p Program	1,874,720	Projected Baseline Energy
CP=34.15	1,874,720	Measured Energy
Baseline only	0	Avoided Energy Use
	28,928	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0 ±28,928	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	1.5%	Projected Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±1.5%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

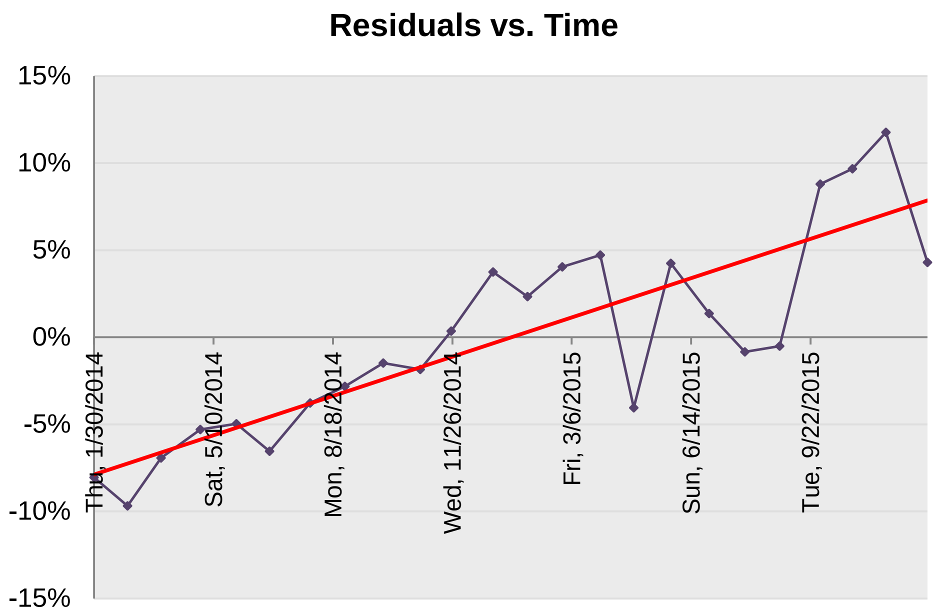
4p Evaluation	1,874,720	Adjusted Baseline Energy
CP=34.15	1,874,720	Measured Energy
Baseline only	0	Avoided Energy Use
	40,317	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0 ±40,317	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	2.2%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±2.2%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	Model Bias

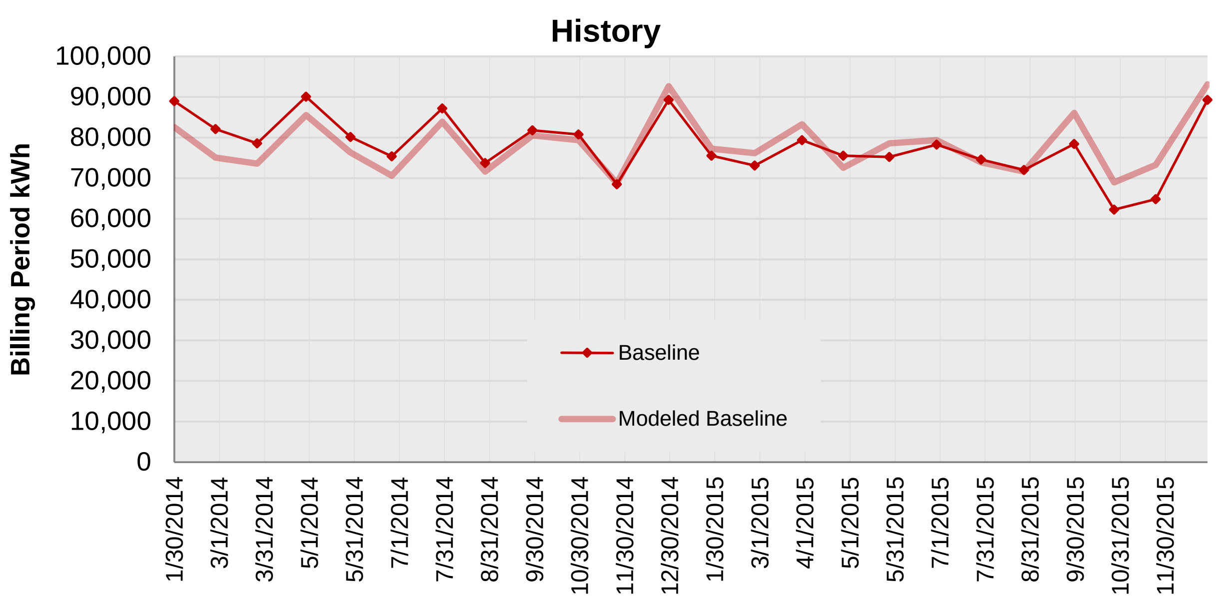
3p Evaluation	1,874,720	Adjusted Baseline Energy
CP=35.0	1,874,720	Measured Energy
Baseline only	0	Avoided Energy Use
	40,243	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0 ±40,243	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	2.1%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±2.1%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	

2p Evaluation	1,874,720	Adjusted Baseline Energy
Baseline only	1,874,720	Measured Energy
	0	Avoided Energy Use
	41,940	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0 ±41,940	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	2.2%	Adjusted Baseline ±Uncertainty @ 80% Confidence Level
	0.0% ±2.2%	Avoided Energy Use and Uncertainty @ 80% Confidence Level
	0.00%	

2p Evaluation Model




Residuals vs. Time
15%
10%
5%
0%
-5%
-10%
-15%
Thu, 1/30/2014
Sat, 5/10/2014
Mon, 8/18/2014
Wed, 11/26/2014
Fri, 3/6/2015
Sun, 6/14/2015
Tue, 9/22/2015


History
Billing Period kWh
100,000
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
Baseline
Modeled Baseline
1/30/2014
3/1/2014
3/31/2014
5/1/2014
5/31/2014
7/1/2014
7/31/2014
8/31/2014
9/30/2014
10/30/2014
11/30/2014
12/30/2014
1/30/2015
3/1/2015
4/1/2015
5/1/2015
5/31/2015
7/1/2015
7/31/2015
8/31/2015
9/30/2015
10/31/2015
11/30/2015

APPENDIX B. CRITERIA FOR SEM SCREENING

Site Screening	Y	N	Notes
Measurements: Is there a single meter per building or multiple meters per building?			If there are multiple buildings on a single meter then it's difficult to detect smaller changes in performance
Measurements: Are the meters utility-grade?			If not then the meter data uncertainty might be a problem
Measurements: At a minimum do utility meters read out monthly?			If utility read out is less than monthly then the ability to model and predict energy usage is more difficult and would drive up uncertainty.
Measurements: There have been no major changes in usage/occupancy in the past 12 months?			This is to ensure that at a minimum the last 12 months will represent the building operations and energy consumption.
Measurements: There are no major usage changes anticipated within the next 12 months?			This is to ensure that the performance period is not affected by large non-routine adjustments
ECMs: Are O&M/RCx opportunities part of the project scope?			If not then maybe the project should be routed under BPA's other program offerings (i.e. deemed or CPP).
Baseline Analysis-To Dos			
Baseline Model: Is ECAM being used for Analysis?			If not, then BPA can support developing this.
Baseline Model: Has 3 years of historical billing consumption with read dates been obtained?			If not, how much is available? A minimum of 12 months is needed.
Baseline Model: Has local weather data been obtained for associated billing consumption?			If not, look at Weather Underground, NOAA or Agrimet.
Baseline Model: Has other probable variables been obtain for associated billing consumption (i.e. occupancy)?			If not, then just investigate whether or not weather is the sole independent variable.
Baseline Model: Have competing baseline models been developed? (i.e. monthly vs daily, 12 month vs 24 month vs 36 month)?			If not, then BPA can support developing this.

Baseline Model: Is CV-RSME <5%.			If not, then further model investigation and discussion with BPA is needed. Current Standards indicate a reasonable model if CV<20% , but BPA's experience has indicated that CV's of <5% is needed for projects with small savings.
Baseline Model: Take a look at the Residuals Scatter Plot. Make sure there is no pattern. Has this been done?			This analysis is somewhat subjective, but in general the plot should be evenly distributed with no discernable pattern. If there is a discernable pattern, then it suggests that a wrong change point model was selected.
Baseline Model: Take a look at the Time Series of Residuals Plot. Make sure the slope is near 0. Has this been done?			If there is a fairly high positive or negative slope (possibly use total slope change of < 50% the savings) then it suggests that during the baseline period there has been a large increase or decrease in energy consumption which might indicate a change in use, previous energy efficiency project or something else that should be taken into account.
Baseline Model: Is Fractional Savings Uncertainty < .50 with a CL=80% (which basically is stating that the savings should at least be twice the uncertainty)? If not, test other models and lengths of pre/post period.			If not then savings may not be able to discerned outside the uncertainty.
Implementation/Reporting			
Model Adjustments: Are there any non-routine adjustments that are needed to the baseline model? (i.e. increase in ventilation rates, increase in occupancy, added load, etc.)			If so, what is the basis for the required adjustments? Are these documented?
Savings Adjustments: Are there any other energy efficiency project that has been implemented and the energy savings claimed thru a different program?			If so, then the energy savings will need to be deducted from the whole building energy savings.
Savings Reporting: Has a final savings calculation and completion report been developed for reporting?			If not, then BPA can support developing this.
Savings Reporting: Was a project log kept during the project and if so was it provided with savings reporting?			If not can one be created to document activities that were performed.

Supplemental			
Policy: Is there an energy management policy in place in the facility?			If so, is this available?
Other Fuels: Are other energy sources (i.e. NG, propane, etc.) used on site?			If so, is that historical data available for review?
Data: Does the site have interval level data either at the building or sub-system level?			This data can possibly be used in lieu of utility data for tracking and reporting.
Executive Support: Is there an energy champion supporting SEM efforts?			If so, who is this person? If not, is this facility in the best position to participate in SEM cohort or is it best suited for P4P or standard programs (CPP or deemed)?