February
2017

Industrial Strategic Energy Management (SEM) Impact Evaluation Report

Submitted by
SBW Consulting, Inc. & The Cadmus Group

Bonneville
POWER ADMINISTRATION

Acknowledgements

We acknowledge the tremendous efforts made by BPA staff and its program contractors, utility staff, and end users to support this study. Many hours were spent reviewing plans and draft work products, as well as gathering documentation, answering questions from the evaluation team, and discussing the results and their implications. This study could not have been completed without these efforts.

Executive Summary

The Bonneville Power Administration (BPA) began offering its Energy Management (EM) Program to industrial facilities in 2010. Through the program, BPA provides long-term energy management consulting services to educate and train industrial energy users for two primary purposes: (1) to develop and execute a long-term strategy for energy planning and (2) to permanently integrate energy management into their business planning. BPA's EM Program was one of the nation's first large-scale deployments of a strategic energy management (SEM) program in the industrial sector, which had engaged 65 facilities by the end of 2014.

BPA offers two components through the EM Program: High Performance Energy Management (HPEM) and Track and Tune (T&T). HPEM provides industrial facilities with training and technical support and engages upper management and process engineers to implement energy management in their core business practices. Through T&T, BPA helps industrial facilities improve operation and maintenance (O&M) efficiencies at little to no cost, while establishing systems that allow the facilities to track energy performance and savings over several years. BPA also offers co-funding for an energy project manager in conjunction with these two components to enable a facility to devote staff time to energy management.

BPA's Energy Performance Tracking (EPT) team developed monitoring, targeting, and reporting (MT&R) guidelines that include the methodology for measurement and verification (M&V) of energy savings for EM Program participants.[1] The methodology aligns with best practices from the International Performance Measurement and Verification Protocol (IPMVP) Option C – Whole Facility.[2] The EPT team analyzed facility meter data, production data, and other relevant data to estimate annual energy savings for each facility, and BPA recorded savings in its reporting system.

The EPT team estimated two types of savings: facility savings and SEM savings. The team estimated facility savings, based on electricity savings at the billing meter level, using the MT&R facility consumption model. Facility savings included SEM savings and savings from capital equipment projects that received rebates through either BPA's Energy Smart Industrial (ESI) Program or other energy efficiency programs. To avoid double counting, the team considered SEM savings equal to the difference between the MT&R facility savings and the savings from prorated capital equipment projects.

---

1 BPA (Energy Smart Industrial EPT Team). "MT&R Guidelines: Monitoring, Targeting, and Reporting (MT&R) Reference Guide." February 20, 2015. Available online: https://www.bpa.gov/EE/Policy/IManual/Documents/MTR-Reference-Guide-Rev5.pdf

2 IPMVP Committee. "International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings." January 2012. Available online: http://www.coned.com/energyefficiency/PDF/EVO%20-%20IPMVP%202012.pdf

BPA also recorded SEM savings in its reporting system. Reported savings equaled the MT&R SEM savings, except when the MT&R SEM savings were negative. When the MT&R SEM savings were negative, BPA recorded zero SEM savings.

Evaluation Objectives

For this assessment, the evaluation team (Cadmus and SBW) focused on the performance between 2010 and 2014 of HPEM and T&T facilities that had the longest history of participation in BPA's EM Program. The evaluation team estimated savings for these facilities and did not extrapolate to the program population.

The evaluation included the following objectives:[3]

- Estimate SEM energy savings and characterize year-to-year SEM savings trends for sampled facilities.
- Verify the EPT Team's estimated SEM savings and BPA's reported SEM savings.
- Survey participants about their adoption of SEM practices and assess whether differences in adoption can explain the energy savings results.
- Develop recommendations, as needed, on how to improve the MT&R guidelines and impact evaluation methods for this program.

The evaluation team independently estimated annual energy savings for each facility using regression analysis. Similar to the MT&R process, we estimated annual facility savings by comparing metered consumption during program engagement to an adjusted baseline. We estimated SEM savings as the difference between total facility energy savings and energy savings from any capital projects incentivized by other energy efficiency programs.[4] BPA provided the data we used to estimate savings, which it collected by working closely with each participating customer.

Evaluation Findings

Finding 1. The EPT team carefully documented the program implementation and collected the data required for evaluation. Overall, the EPT team's EM Program data collection and documentation can serve as an industry standard for SEM programs. The EPT team's ongoing communication with participants through several program years resulted in the collection of high-quality data for the evaluation. The evaluation team was able to estimate savings for most facilities because the EPT team had thoroughly documented the program's implementation. For each facility and year, the EPT team prepared a project completion report, which described the facility operations and energy consumption, documented implemented SEM activities, and provided an estimate of the SEM energy

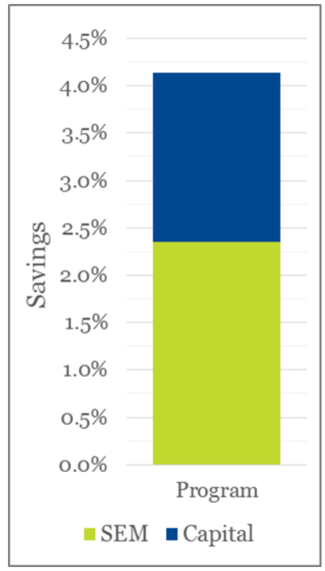
3 The scope of this evaluation did not include an assessment of program cost-effectiveness.

4 EM Program participants were eligible to receive incentives for capital or custom projects from BPA's ESI Program or other utility programs.

savings. In addition, the EPT team collected energy consumption data and production data required for evaluating participating facilities.

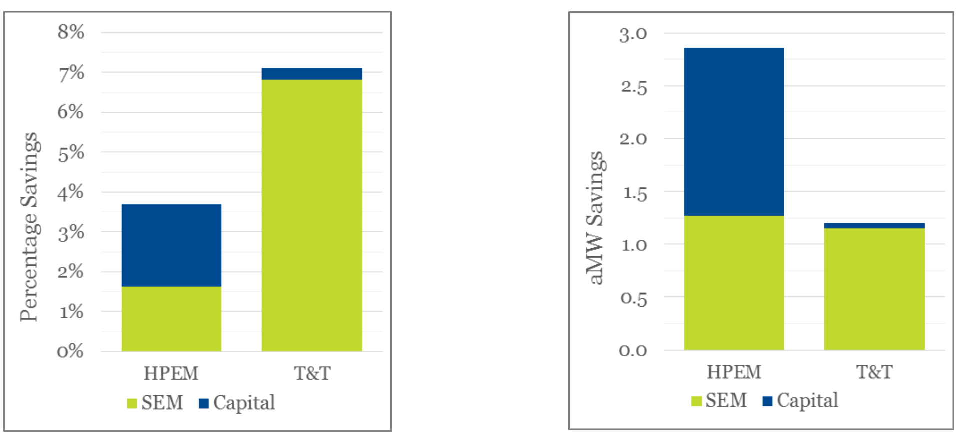
Finding 2. SEM saved 2.3% of facility electricity consumption. The evaluation team estimated that, across all years, sampled EM Program facilities saved 4.1% of electricity consumption from the combination of SEM and capital projects, for an annual average savings of 3.8 average megawatts (aMW).[5] Capital project savings equaled 1.8% of electricity consumption.[6] SEM savings equaled 2.3% of electricity consumption, an average of 2.1 aMW per year. The percentage savings are summarized in Figure 1.

Figure 1. EM Program Savings



Finding 3. SEM savings varied by Energy Management Program component. Sampled T&T facilities saved the most energy as a percentage of consumption, with total facility savings of 7.1% and SEM savings of 6.8% (an average of 1.1 aMW). Sampled HPEM participants achieved facility savings of 3.7% and SEM savings of 1.6% (an average of 1.3 aMW). These results are summarized in Figure 2.

Figure 2. EM Program HPEM and T&T Savings

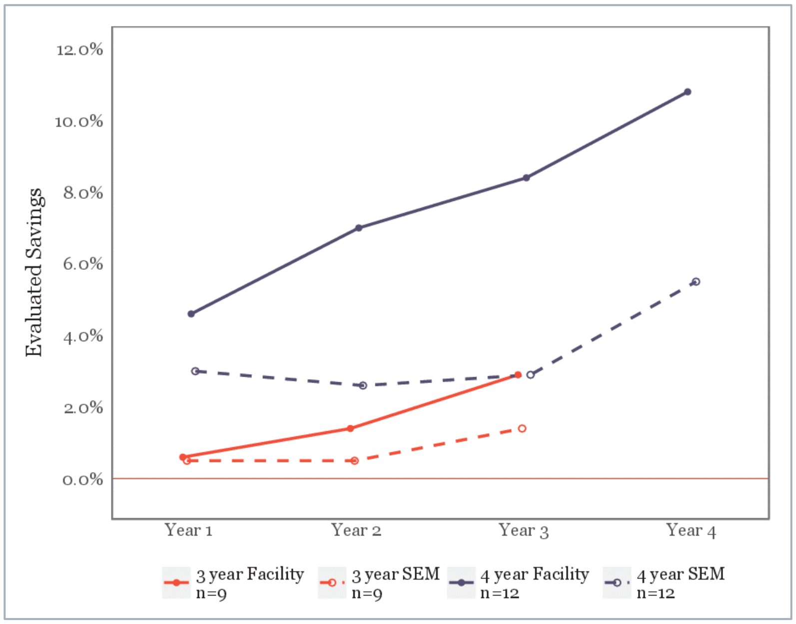


5 Percentage savings were the sum of electricity savings for all facilities and years divided by the sum of adjusted baseline consumption for all facilities and years. The aMW savings were average annual MWh savings per hour and obtained by dividing the annual MWh savings by 8,760.

6 Capital project savings were not evaluated in this study. The evaluation team obtained these savings from original M&V estimates, contained in the MT&R reports.

Finding 4. SEM savings persisted during the participation period. The evaluation team tracked the energy savings of sampled HPEM facilities that participated for three or four years. As Figure 3 shows, facility savings increased throughout the participation period and SEM savings (dashed lines) persisted after the first year and increased slightly in the last year. This persistence of savings suggests that facilities continued to practice energy management activities throughout the engagement.

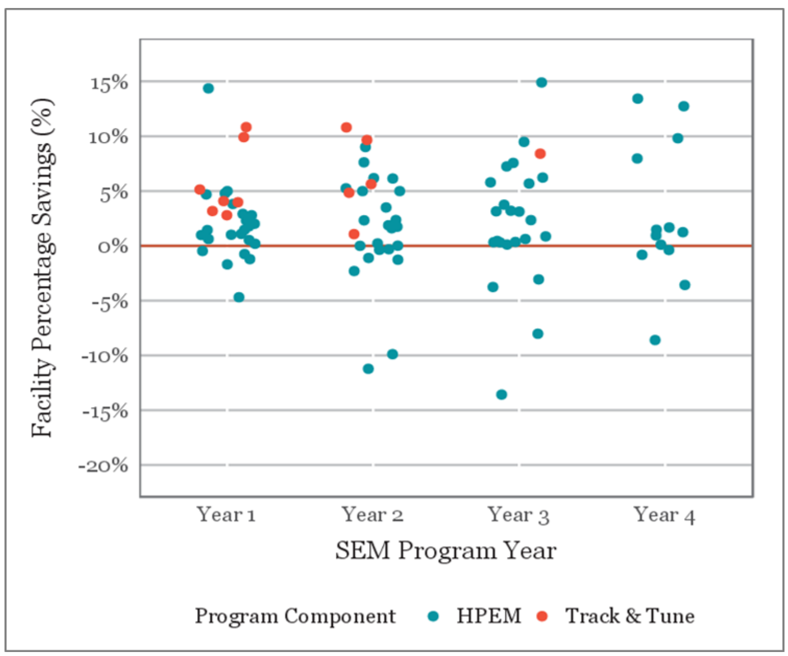
Figure 3. Annual Percentage Savings by Years in Program



Finding 5. Individual facility savings were variable. There was significant variation in savings between facilities and from year-to-year for individual facilities. The percentage savings coefficient of variation (the ratio of the sample standard deviation to the sample mean) was 201%. This variation in annual savings likely reflected differences in SEM implementation, changes in electricity consumption, and uncertainty of the savings estimates.

Figure 4 shows the evaluated annual SEM savings for individual facilities by program year.

Figure 4. Summary of Variability of Annual SEM Percentage Savings Estimates



Note: Each dot represents the annual SEM savings for an individual facility in a program year.

Finding 6. Some facilities had estimated consumption increases. In the majority (78%) of facility program years, evaluated SEM savings estimates were positive. However, in 22% of facility program years, the SEM savings estimate was negative. This includes 10% of cases where both facility and SEM savings were negative, as well as 12% of cases when the facility savings estimate was positive but the SEM savings estimate was negative after subtracting capital project savings.

Estimated increases in consumption likely reflect difficulties in the measurement of savings because of omitted variables, degradation in capital equipment performance, or unaccounted for non-programmatic effects—not that the program caused consumption to increase. However, an increase in facility consumption (e.g., because of a program implementation error) cannot be ruled out. As there is no accepted method for differentiating between omitted variables and a program causal effect, the evaluation results included estimated consumption increases.

Finding 7. The adoption of SEM elements was not correlated with SEM percentage savings. The Consortium for Energy Efficiency identified 13 management practices, called "elements," for facilities to continuously improve their energy performance. The evaluation team surveyed 24 EM Program participants in both program components to assess their adoption of these elements. We analyzed whether facilities that implemented a larger number of SEM elements or that adopted specific elements saved more energy. The results in Appendix N show no pattern of specific SEM elements. This may be due to the small sample size, unexplained variation in percentage savings between facilities, or because savings depended on factors outside this survey (such as how well participants implemented the SEM practices).

Finding 8. The evaluation team verified the MT&R SEM savings estimates. The evaluation team's estimate of SEM savings (2.3% of consumption) was slightly higher than the EPT team's MT&R SEM savings estimate (2.2% of consumption). The MT&R SEM savings realization rate—the ratio of evaluated to MT&R savings—was 1.06.[7] The MT&R realization rates were 1.05 for T&T and 1.08 for HPEM. The MT&R and evaluation savings estimates for individual facilities were also similar: in 73% of facility-years, the evaluated savings and the MT&R savings estimates were not statistically different.[8] The evaluation savings estimate was statistically different and

7 The realization rate was the ratio of evaluation savings to either the MT&R or reported savings for evaluated facilities. Realization rates greater than 1.0 indicate that the evaluation savings exceeded the MT&R or reported savings. These realization rates apply to evaluated facilities between 2010 and 2014 and may not represent the current or future performance of the EM program population.

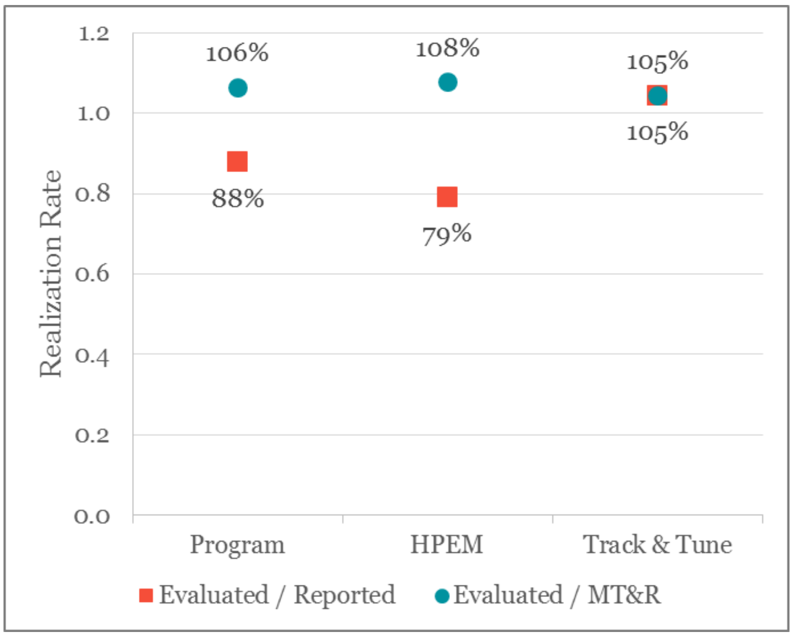
8 The savings estimates were not statistically different when the 80% confidence interval around the evaluated facility savings included the MT&R facility savings.

higher than the MT&R estimate in 12% of facility-years and statistically different and lower than the MT&R in 15% of facility-years.[9]

Finding 9. The evaluation team estimated lower SEM savings than BPA reported due to BPA's reporting practices. BPA reported program SEM energy savings of 2.7% (average of 2.4 aMW per year). The evaluation team estimated savings of 2.3% (average of 2.1 aMW per year), or 12% less. The reported SEM savings realization rate was 0.88. The reported savings realization rates were 1.05 for T&T and 0.79 for HPEM.

Figure 5 shows realization rates for the MT&R and reported SEM savings for the program, as well as the HPEM and T&T components.

Figure 5. Realization Rates of SEM Savings by Program Component



The evaluated savings were less than the reported savings because of BPA's practice of reporting zero savings for facilities with negative savings estimates. BPA reasoned that an increase in facility electrical consumption was not likely to have been caused by SEM implementation. Also, because incentives are based on savings, this convention mitigates a change in payment policies.

However, this reporting convention treats negative and positive savings estimates inconsistently. Positive savings estimates were just as likely to exhibit error as negative savings estimates, and the sign of the savings estimate should not be the reason for accepting or rejecting it. Reporting zero savings for negative facility savings biases the estimates of overall

9 Facility-year savings were savings for a facility during a participation year.

program savings upwards. Appendix K discusses the issue of negative SEM savings estimates.

Finding 10. More research about estimating SEM savings is needed. This evaluation led to new insights about the reliability of different SEM savings estimation methods, estimation of SEM savings uncertainty, causes of negative savings estimates, and ways of controlling for significant, non-programmatic changes in facility operations and energy consumption (non-routine adjustments). Nevertheless, more research is needed in each of these areas.

Key Recommendations for EM Program M&V

The evaluation team makes the following key recommendations for performing measurement and verification of the EM savings.

The EPT team should do the following:

- Continue to use statistical analysis of facility consumption to estimate savings. Specifically, the EPT team should employ the forecast savings estimation approach on a site-specific basis. This approach is widely accepted, familiar to program participants, and expected to produce accurate savings estimates.
- Continue to collect high-frequency consumption data.
- Continue to report estimated increases in consumption in the MT&R model workbooks and to document the application of any non-routine adjustments.
- Use discretion about whether to calculate and report uncertainty of the MT&R facility savings estimates (estimation of savings uncertainty is not essential for M&V).
- Routinely test for the statistical significance of weather variables in the MT&R energy consumption regression model.

BPA should do the following:

- Attempt to improve the accuracy of the reported SEM savings by recording negative SEM savings estimates or making program-level adjustments to savings.

If BPA wants to conduct additional research, we recommend investigating the following topics:

- The relationship between savings and adoption of specific SEM elements.
- How the persistence of capital project savings can impact the accuracy of SEM savings estimates.

- Whether participation in an SEM program increases the number of capital projects implemented and the persistence of capital project savings.
- Program cost-effectiveness by collecting data on participant facilities' costs of implementing SEM and savings from other fuels.
- How the persistence of savings after a facility finishes its engagement can be used to better assess the program's long-term value and cost-effectiveness.

Evaluation Recommendations

Although this evaluation has broken new ground in many areas, we recommend that BPA or other national evaluators of SEM programs further explore further these topics:

- Evaluate the energy savings of the newest EM projects, which were not considered in this evaluation.
- Assess the effect of BPA's new policy of establishing a new baseline for participant facilities every two years on savings realization rates.
- Conduct a process evaluation to understand why HPEM cohorts performed differently and to gain insights about the relationship between savings and implementation of specific SEM activities.
- Study how uncertainty of capital project savings estimates affects SEM savings estimates.

Table of Contents

List of Figures

LIST OF TABLES

This page left blank.

1 Introduction

Bonneville Power Administration (BPA) launched its Energy Smart Industrial (ESI) Program in October 2009, and began delivering the ESI Energy Management (EM) Program in July 2010. Energy management differs from traditional energy efficiency programs, as it consists of a comprehensive energy efficiency strategy that includes both capital projects and the implementation of operations, maintenance, and behavioral changes. Through the program, BPA provides long-term energy-management consulting services that educate and train industrial energy users to (1) develop and execute a long-term energy planning strategy and (2) permanently integrate energy management into their business planning.

The program has two components:

- High Performance Energy Management (HPEM): Through this component, BPA provides industrial facilities with training and technical support, engaging upper management and process engineers to incorporate energy management in their core business practices. HPEM entails applying the principles and practices of strategic energy management (SEM) within an industrial facility.
- Track and Tune (T&T): Through T&T, BPA helps industrial facilities improve operation and maintenance (O&M) efficiencies at little to no cost, while establishing systems that allow the programs and facilities to track energy performance and savings over several years.

BPA also offers co-funding for an energy project manager in conjunction with these two tracks to enable a facility to devote staff time to energy management.

BPA's Energy Performance Tracking (EPT) team developed monitoring, targeting, and reporting (MT&R) guidelines to estimate energy savings from SEM activities for HPEM and T&T participants.[10] This methodology employs regression analysis of facility energy consumption, using pre- and post-participation meter data to establish adjusted baseline electricity consumption and to estimate energy savings associated with program activities. The EPT team estimated the energy savings for each facility and subtracted capital project savings, and BPA engineers made M&V site-specific decisions to record these savings in the BPA energy efficiency (EE) reporting system. When the MT&R model resulted in a negative annual savings estimate, BPA reported zero savings, based on BPA engineers' review and decision that the increase in electricity consumption did not result from the program intervention, but rather from unknown or outside

[10] Energy Smart Industrial (ESI) Energy Performance Tracking (EPT) Team. 2015. MT&R Guidelines: Monitoring, Targeting, and Reporting (MT&R) Reference Guide. https://www.bpa.gov/EE/Policy/IManual/Documents/MTR-Reference-Guide-Rev5.pdf

factors that were not accounted for in the MT&R model. Additionally, this practice mitigates issues from BPA customer utility and end-user payments associated with negative savings on a site-by-site basis.

As of January 2016, the EM Program had five HPEM cohorts and a number of facilities participating in T&T. An HPEM cohort was a group of facilities that began participating in HPEM at approximately the same time. T&T participants began in different years. Table 1 shows the participation levels for each cohort, the number of years each has been in the program, and number of sites included in the evaluation.[11] For this evaluation, the team focused on the HPEM 1 and HPEM 2 cohorts and seven T&T sites.[12] We excluded the HPEM 3 and HPEM 4 cohorts because one evaluation objective was to assess annual savings trends, and these participants had only claimed savings for one year or less at the time of sample selection.[13] The chosen T&T sites had at least one year of claimed savings and did not pose barriers for data collection. Evaluation sample selection is discussed further in Section 2.1: Evaluation Sample Selection.

Table 1. HPEM and T&T Participation Through 2014

Program Component	Participating Sites (n)	Program Start Date	Years in Program	Sites Included in Evaluation (n)
HPEM 1	14	Summer 2010	5	13
HPEM 2	11	Fall 2011	4	11
HPEM 3	6	Spring 2012	3	0
HPEM 4	8	January 2014	2	0
SI-HPEM	8	September 2014	1	0
T&T	18	2010 through 2014	2 to 5	7
Total	**65**	**N/A**	**N/A**	**31**

[11] A site is an industrial location that implemented energy management through the program. A facility is an area over which energy use is measured and modeled. A site may have more than one facility (e.g., multiple buildings at one location).

[12] The evaluation team determined that it was not possible to estimate savings for one HPEM 1 facility because of suspected inaccuracies in the estimate of savings for a large capital lighting project and the poor predictive performance of the facility's baseline consumption model. When estimating aMW savings or percentage savings, the team excluded this facility.

[13] To estimate efficiency savings from SEM improvements over time, BPA tracked energy use to measure energy savings over multiple years. M&V approaches (used by program implementers or evaluators) that measure SEM savings over shorter periods of time (for example, a few months) cannot accurately capture savings from improvements in efficiency over time.

1.1 Evaluation Background

In 2013, Cadmus completed an impact evaluation for the first year (2010–2011) of the EM Program, evaluating first year savings for the HPEM 1 cohort and two T&T sites.[14] This evaluation covers 2010 to 2014 and builds on the findings of that first evaluation, quantifying energy savings in each participation year for the HPEM cohorts and T&T facilities.

1.1.1 Previous Evaluation Findings and Recommendations

In the 2010-2011 evaluation, the team found that the first cohort of EM Program participants achieved facility electricity savings of 4.4% and SEM savings of 2.7% of electricity consumption in the first year of engagement. The program achieved a realization rate of 0.88 for electricity savings based on a comparison of the evaluated SEM savings and MT&R SEM savings estimates.[15] The evaluated first-year pilot electricity savings estimates were statistically different from zero, and the 80% confidence interval of [0.62, 1.15] for the electricity savings realization rate included 1.0, indicating that the confidence interval included the program savings estimate.[16]

The 2010-2011 evaluation report noted several challenges in estimating energy savings. These included the following:

- ***Data Frequency.*** The evaluation team was more likely to detect savings at facilities with high frequency interval data for energy consumption and production.
- **Capital Measures Confounding the Analysis.** At some sites, the installation of capital measures just before or after the start of a facility's participation in HPEM or T&T made isolating SEM savings difficult or impossible.
- ***Implementation Timing of Measures.*** SEM savings for activities implemented near the end of a program year may not have been fully estimated, as not enough months of post-implementation data were available.

As a result of these challenges in 2010-2011, the evaluation team offered several recommendations to help improve the accuracy and precision of the

---

14 Cadmus. "Energy Management Pilot Impact Evaluation." Prepared for Bonneville Power Administration. February 1, 2013. Available online: http://www.bpa.gov/EE/Utility/research-archive/Documents/BPA_Energy_Management_Impact_Evaluation_Final_Report_with_Cover.pdf

15 Realization rate is the ratio of evaluation savings to reported savings. Realization rates greater than one mean that we found more savings than were reported.

16 In statistical terms, the evaluation team could not reject that the pilot savings equaled the MT&R savings estimate.

energy savings estimates. Table 2 shows these recommendations and how they were addressed.

Table 2. 2010-2011 Evaluation Recommendations and Status of Implementation

Recommendation	Status
When beginning an engagement with a site, perform a statistical power analysis to estimate the probability of detecting the expected savings at the site.	BPA added fractional savings uncertainty (a type of statistical power analysis) guidance to its MT&R guidelines.
Collect data for additional months in the pilot's second year and evaluate the second-year pilot savings.	This report presents savings from multiple years of participation, including the second year pilot savings.
When possible, collect higher frequency billing data and production data to provide more certainty in energy savings and to decrease the confidence interval range.	Implementer collected higher frequency billing and production data when available.
Re-estimate the first-year pilot savings for sites with insignificant savings after obtaining data for additional periods in the second year.	In this evaluation, the team re-estimated first year savings for all pilot sites.
The MT&R models should test and account for autocorrelation, especially if addressing higher frequency data (i.e., daily or weekly data).	BPA added guidance for testing and accounting for autocorrelation to the MT&R guidelines. However, this evaluation is no longer recommending the EPT team account for autocorrelation since it does not impact the energy savings estimate.
Report confidence intervals and relative precision for all savings estimates.	BPA added guidance for calculating uncertainty, but not for calculating confidence intervals. However, this evaluation is no longer recommending the EPT report confidence intervals since it can be complex with the forecast method and it does not impact the energy savings estimate.

1.2 Evaluation Objectives

For this evaluation, the team sought to achieve the following objectives:

- Use regression analysis of facility consumption to estimate SEM energy savings and characterize year-to-year SEM savings trends.
- Verify the EPT Team's estimates of SEM savings and BPA's reported SEM savings.
- For selected sites, conduct an exploratory statistical analysis, comparing the MT&R and evaluation approaches for estimating savings.
- Survey participants about their adoption of SEM practices and assess whether adoption can explain the estimated energy savings.

- Assess program data collection, determine evaluability, and identify potential improvements to documentation and data collection.
- Develop recommendations, as needed, on how to improve the MT&R guidelines and impact evaluation methods for this program.

1.3 Definitions of Savings Terms

This report refers to several categories of electric savings, defined here. The methodologies for calculating these savings are described in Section 2.3: Energy Savings Calculation Methods for SEM.

- MT&R facility savings: the savings calculated by the EPT team at the billing meter level using the MT&R model. These savings include both SEM savings and savings from capital equipment projects that received rebates through either the ESI Program or other energy efficiency programs.
- MT&R SEM savings: the savings calculated by the EPT team after taking the difference between the MT&R facility savings and the savings from prorated capital equipment projects funded by other efficiency programs. The differencing avoids double counting of savings from capital equipment projects that received rebates from other programs.
- Reported SEM savings: the SEM savings listed in BPA's reporting system. Typically, reported SEM savings are equivalent to the MT&R SEM savings. They differ when the MT&R SEM savings are less than zero, as BPA reports zero savings rather than negative savings.[17]
- Evaluation facility savings: the savings calculated by the evaluation team using billing meter data. These savings included SEM savings and savings from capital equipment projects that received rebates through either the ESI Program or other energy efficiency programs.
- Evaluation SEM savings: the difference between the evaluation facility savings and the prorated savings from capital equipment projects that received rebates from other programs.

[17] See Section 3: Program MT&R and Reported Savings for BPA's rationale for reporting negative savings as zero savings, and see Section 4: Evaluation Energy Savings Findings for the influence this had on the realization rate.

2 Evaluation Methodology

The evaluation team performed the following activities to evaluate EM Program savings:

- Select the sample of facilities for evaluation
- Collect and review the facility energy consumption and production data and project completion reports
- Independently estimate facility and program savings and conduct exploratory statistical analyses for three case studies

Each of these activities is discussed below.

2.1 Evaluation Sample Selection

Using SEM savings estimates from the first evaluation, the team simulated different facility sampling strategies to test whether the strategies would yield accurate and precise estimates of the program population savings. We concluded that because of the small program population and significant variability in facility savings realization rates, there was a high probability that analyzing a sample of facilities would result in a biased estimate of the program savings. These simulation results are shown in Appendix A.

Based on this review and in consideration of the study objectives, the evaluation team selected all HPEM 1 and HPEM 2 facilities and a sample of T&T facilities for analysis. The evaluation team subsequently determined that it was not possible to estimate savings for one HPEM 1 facility because of suspected inaccuracies in the estimate of savings for a large capital lighting project and the poor predictive performance of the facility's baseline consumption model. Reported, MT&R, and evaluated savings presented in this report do not include savings for this facility. We excluded the HPEM 3 and HPEM 4 cohorts from the study to focus on facilities that had participated for longer (and had more than one year of data). The evaluation team worked with BPA to identify T&T facilities where savings were most likely to be evaluable, or those that had claimed savings for at least one year and did not pose barriers for data collection or risk BPA's relationship with the facility or the facility's utility. We reviewed all facilities and chose seven of 18 for evaluation.

The team focused our evaluation on facilities that had participated for longer because of the EM Program's emphasis on continuous efficiency improvements. Participating facilities are expected to build a workplace culture that emphasizes SEM and the continuous identification and implementation of new efficiency opportunities. This focus on continuous change contrasts with implementing a capital project, which involves a one-time intervention and the measurement of savings over a short time period.

To estimate efficiency savings from SEM improvements over time, it may be necessary to track energy consumption and to measure energy savings over multiple years. The measurement and verification (M&V) approaches used by program implementers or evaluators that measure SEM savings over shorter periods of time (such as a few months) cannot accurately capture improvements in efficiency over time.

A summary of the sample design is shown in Table 3.

Table 3. Summary of Energy Management Evaluation Sample

Program Component	Participating Sites	Number of Sites Included in Evaluation*	Number of Facilities Included in Evaluation*	Years Evaluated
HPEM 1	14	13	14	2010-2014
HPEM 2	11	11	11	2011-2014
HPEM 3	6	0	0	N/A
HPEM 4	8	0	0	N/A
SI-HPEM	8	0	0	N/A
T&T	18	7	7	2010-2014**
Total	**65**	**31**	**32**	**2010-2014**

* Some sites have more than one participating facility, necessitating the development of more than one model per site. The evaluation team determined that it was not possible to estimate savings for one facility because of suspected inaccuracies in the estimate of savings for a large capital lighting project and the poor predictive performance of the facility's baseline consumption model. When estimating aMW savings, MWh savings, or percentage savings, the team excluded this facility.

** T&T began in 2010, though not all enrolled facilities during 2010 – 2014 were evaluated. Of those that were, the facilities included in the evaluation participated between 1 and 3 years.

2.2 Evaluation Data Collection and Review

The team began our impact evaluation with a detailed review of the program documentation and data specific to each facility.

2.2.1 Facility Documentation and MT&R Models

BPA provided annual completion reports and annual MT&R model workbooks for each sampled facility and program year. The EPT team submitted completion reports annually, which documented the facility characteristics and any facility changes, SEM activities completed each year, capital project savings, the regression model and diagnostics, and the resulting savings. The annual MT&R model workbooks contained data, the regression model and cumulative sum calculations, supporting the savings values shown in the completion reports.

The evaluation team reviewed the following information and data for each of the sampled facilities:

- Background information about the industry, facility, and program implementation
- Project implementation data, history, and savings estimates for capital projects
- Project implementation data, history, and savings estimates for SEM projects
- MT&R process reports and documentation
- Raw data from the facility (e.g., billing, weather, production, and other data used in the MT&R model)

We conducted an in-depth review of the data and MT&R models for each sampled facility and participation year, focusing on the following:

- The data series' completeness and quality
- The capital projects' timing and effects
- The baseline period definitions
- Potentially omitted variables correlated with both energy consumption and program participation

The evaluation team reviewed data for each facility and discussed questions about the data with the EPT team. After obtaining answers and determining that we had all the needed data, the evaluation team reviewed the facility documentation, MT&R models, and individual capital measure savings calculations.

Upon completing our review of MT&R documentation and data, the team attempted to replicate the model results and savings estimates in the MT&R reports for each facility. When there were discrepancies between the MT&R analysis and our results, we noted the difference for additional investigation and discussion with the EPT team.

2.2.2 Other Data Sources

The evaluation team conducted phone surveys with facility energy managers and analyzed the survey results. The team considered but did not conduct site visits.

Phone Surveys

The evaluation team conducted phone surveys with facility energy managers. BPA requested that we keep these phone surveys short, since some facilities had recently been contacted as part of other regional research efforts. Therefore, we used the surveys to assess the adoption of SEM elements, but not to verify that particular SEM activities had been completed or verify whether measures had been rebated through other efficiency programs.[18] The team did not use survey responses to verify facility energy savings.

Site Visits

The evaluation team considered but did not conduct site visits, as we were uncertain whether the benefits would justify the cost. The team was uncertain how and the extent to which site visits would improve the accuracy of the SEM savings estimates. However, the evaluation revealed that capital project savings were 40% of the estimated facility savings. In light of the significant contribution of capital projects to the facility savings, BPA should consider whether site visits would improve the accuracy of the capital project savings estimates.

2.3 Energy Savings Calculation Methods for SEM

The evaluation team reviewed different methods for calculating facility savings, including the forecast, backcast, and pre-post methods. These methods are described in the forthcoming U.S. Department of Energy (DOE) Uniform Methods Project Strategic Energy Management Evaluation Protocol,[19] DOE Superior Energy Performance (SEP) Measurement and Verification protocol,[20] and IPMVP Option C – Whole Facility. We also reviewed the pre-post model savings estimation method.[21] Appendix B provides an overview of the various methods.

From these protocols, the evaluation team selected the forecast method as the default for estimating savings. The evaluation team selected the forecast method for the following reasons:

If the energy consumption model is correctly specified, the forecast method is expected to yield an accurate savings estimate.

---

[18] Verifying energy efficiency activities through phone surveys has limitations. The respondent may not understand which activity you are referring to, may not remember the activity, or may not be familiar with the details of the activity or measure.

[19] U.S. Department of Energy. "Strategic Energy Management Evaluation Protocol: The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures." Forthcoming.

[20] U.S. Department of Energy. "Superior Energy Performance Measurement and Verification Protocol for Industry." November 19, 2012. Available online: http://energy.gov/eere/amo/downloads/superior-energy-performance-measurement-and-verification-protocol-industry

[21] Luneski, Robert. "A Generalized Method for Estimation of Industrial Energy Savings from Capital and Behavioral Programs." Energy Systems Laboratory, Texas A&M University. (2011). Available online: http://hdl.handle.net/1969.1/94789

The forecast method is well accepted by program implementers and administrators and is the convention for estimating savings for SEM program facilities.

This aligned the evaluation team's default estimation method with that of the EPT team. This reduced the potential for differences between the two sets of savings estimates.

For a small number of facility-years, the evaluation team employed the pre-post method because the team expected that it would produce a more accurate savings estimate. The evaluation team's use of the pre-post method is described in Appendix D.

2.4 Savings Estimation

The evaluation team estimated energy savings for each of the 31 sites (35 facility energy models) in the analysis sample using the forecast method and following BPA's ESI MT&R Guidelines.[22] Using regression analysis, the team estimated facility savings by comparing a facility's electricity consumption in the reporting period during SEM implementation to its adjusted baseline consumption, which reflects what consumption would have been during the reporting period if SEM had not been implemented. The evaluation team estimated the adjusted baseline consumption using a regression analysis of the facility's baseline period energy consumption. We chose each facility's regression specification to accurately represent the relationship between the facility's energy consumption and its production output(s), weather, and other drivers of energy consumption.

The evaluation team followed five main steps to develop forecast regression savings estimates:

Define the baseline and reporting period and the facility boundaries. For 30 of 35 evaluated facility models, the evaluation team used the same baseline period as that selected by the EPT team.

Build the baseline regression model. We selected model variables by analyzing baseline period data to identify the facility's energy consumption covariates. The purpose of using baseline period data was to build a model that would accurately predict facility energy consumption under baseline conditions during the reporting period.

Calculate adjusted baseline energy consumption for the reporting period using the forecast regression model. The adjusted baseline represents what energy consumption would have been during the reporting period without SEM.

[22] Two sites each had two facilities for which separate consumption models were estimated. One facility at one of the sites had two consumption models estimated for different program years.

Estimate facility savings for each interval of the reporting period and for whole the reporting period as the difference between the adjusted baseline and metered energy consumption.

Estimate SEM savings as the difference between facility savings and savings from any capital projects receiving incentives from other energy efficiency programs.

The team also calculated 80% confidence intervals around the facility savings estimate.[23]

The evaluation team did not independently verify the capital measure savings, which was outside of the scope of this evaluation.

Further details about the team's process to develop and select an appropriate regression model and to estimate facility and SEM energy savings are provided in Appendices C and D.

2.4.1 Non-Routine Adjustments

A non-routine adjustment is an adjustment to metered energy consumption that accounts for a non-programmatic change in facility operations. For example, a facility may have installed a new piece of equipment during the reporting period, causing energy consumption to increase, but also making it difficult to estimate the SEM savings. IPMVP defines a non-routine adjustment as an "individually engineered calculation… to account for changes in static factors within the measurement boundary since the baseline period."[24]

Analysts can make non-routine adjustments during the baseline or reporting period energy consumption by using an engineering estimate to adjust the baseline. In cases when an engineering estimate is unavailable, it may also be possible to account for the non-programmatic change in the facility's energy consumption using a regression model. For example, it might be possible to account for the non-programmatic change by indicating the change in a pre-post regression model. However, a pre-post regression model would only be applicable if high-frequency data were available and the non-routine adjustment and program year indicator variables did not coincide too closely.

The evaluation team developed a logic flow, shown in Appendix D, for determining whether and how to make non-routine adjustments. The

23 The team chose to use 80% confidence intervals based on the Regional Technical Forum's *Guidelines for the Estimation of Energy Savings* (December 8, 2015) for sampling custom measures, page 35, which states, "In general, sampling should not be used unless it is practical to achieve relative error in the estimate of mean unit energy use equal to or less than ±20% at a confidence level of 80%, without introducing substantial bias."

24 International Performance Measurement and Verification Committee. "International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings." January 2012. p. 55. Available online: http://www.coned.com/energyefficiency/PDF/EVO%20-%20IPMVP%202012.pdf

evaluation team used this logic to make non-routine adjustments to the energy consumption for several facilities. The logic flow shows that evaluators may consider a site unevaluable in one or more program years if non-programmatic changes cannot be reliably modeled using any available savings estimation approaches.

2.4.2 Evaluation Treatment of Consumption Increases

The EPT team and evaluation team estimated consumption increases (i.e., negative savings) for some facilities, which arose in two ways. First, in some cases the regression-based estimate of facility savings was negative. Second, in some cases the regression-based estimate of facility savings was positive, but the capital project savings was larger than the facility savings. In these cases, the estimated SEM savings became negative after subtracting the estimate of the capital project savings.

Following best practices, the evaluation team did not differentiate between estimates of positive and negative facility or SEM savings, reporting each without regard to sign. Though the EM Program was not expected to lead to increased facility energy consumption, the evaluation team could not rule out that the program had increased energy consumption. There is not an accepted method for determining for individual facilities whether an estimated consumption increase was a program effect or the result of a variable omitted from the baseline regression model.

Based on BPA engineering M&V site-specific decisions, BPA reported estimated consumption increases as zero SEM savings; however, the evaluation team could not justify treating negative savings results differently than positive savings results. Positive savings estimates were just as likely to exhibit error as negative savings estimates, and the sign of the savings estimate should not determine whether to accept or reject it. Reporting zero savings for negative facility savings would bias the estimates of program savings upwards. The team provided further discussion of this issue in a memo to BPA, which is included as Appendix K.

3 Program MT&R and Reported Savings

The EPT team used the forecast approach to calculate savings and documented the MT&R savings calculations in an Excel workbook and MT&R report. The EPT team calculated annual MT&R facility savings, capital savings, and SEM savings for each facility. As noted above, the capital savings were from capital projects incentivized through other energy efficiency programs, and SEM savings represented the difference between the MT&R facility energy usage and incentivized capital project savings. BPA uploaded the SEM savings into its BPA EE reporting system, then calculated incremental annual SEM savings by subtracting the savings from the previous year.[25] Lastly, BPA applied the busbar adjustment to account for line losses. Facility and capital savings, as documented in the MT&R workbooks, were not reported in BPA's EE reporting system.[26]

BPA's EE reporting system for the years included in this evaluation tracked incremental annual SEM savings. However, because BPA tracks annual savings, the evaluation team focused this report on the average annual savings calculated from the annual savings (which do not subtract the previous year's savings). The average annual savings are the weighted average of annual savings per facility, with weights equal to the number of facilities evaluated in each program year.

3.1 Average Annual MT&R and Reported Savings

Overall, the EPT team estimated average annual MT&R savings of 31,807 MWh or 4.0% of consumption for facilities in the EM evaluation sample.[27] Savings from capital projects were 1.8% of consumption and savings from SEM were 17,599 MWh (19,149 MWh when adjusted for busbar) or 2.2% of consumption.[28]

Overall, the reported savings were higher than the MT&R modeled savings. BPA reported average annual savings of 21,276 MWh (23,203 MWh when adjusted for busbar), which were 21% higher than MT&R modeled savings. This was due to BPA's reporting of zero SEM savings in cases where the

25 This step of subtracting the previous year's savings to calculate incremental savings was not included in the MT&R savings or in the evaluated savings.

26 BPA reports capital savings into its reporting system for savings estimated using traditional M&V methods, consistent with BPA M&V protocols. These are tracked separately from the EM Program savings reporting process.

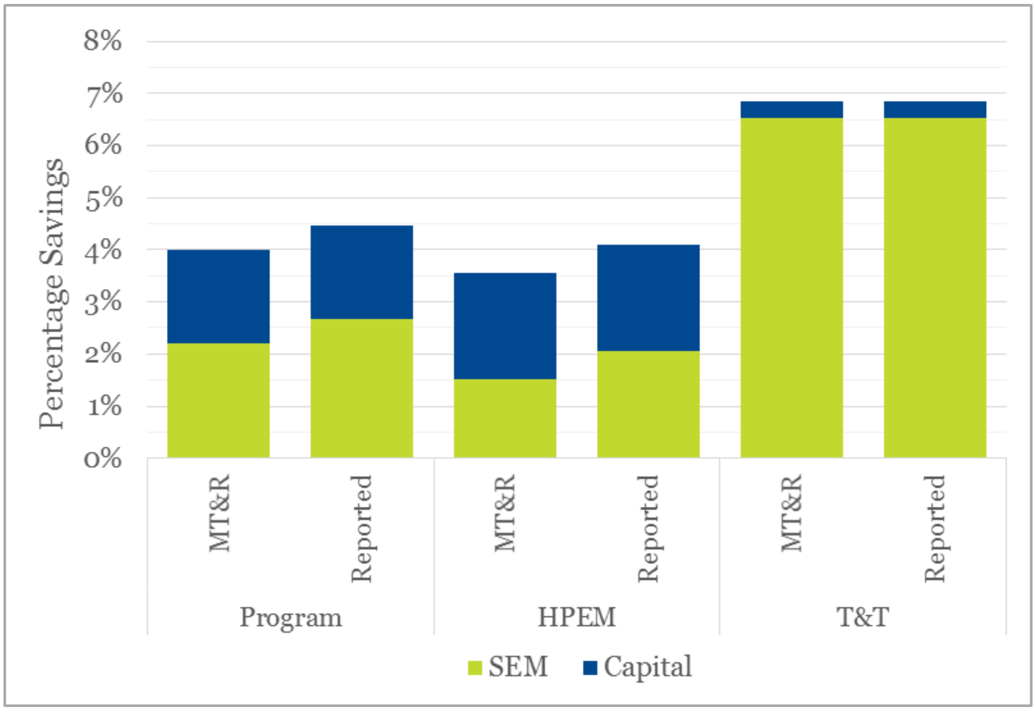
27 The evaluation team determined that it was not possible to estimate savings for one HPEM 1 facility because of suspected inaccuracies in the estimate of savings for a large capital lighting project and the poor predictive performance of the facility's baseline consumption model. When estimating aMW savings or percentage savings, the team excluded this facility.

28 The annual consumption was determined by adding the savings estimate to the metered consumption to estimate the baseline consumption in the absence of the program.

MT&R model estimated an increase in energy consumption (i.e., negative SEM savings).

Figure 6 shows the BPA reported and MT&R percentage savings for the program and the T&T facilities and HPEM cohorts, depicting the SEM savings (yellowish green) and capital project savings (blue). The percentage savings represents the sum of annual savings divided by the sum of annual adjusted baseline consumption for all facilities and program years.

Figure 6. MT&R and Reported Average Annual Savings by Program Component



*Note: BPA's EM reporting system does not track capital savings, so the reported capital savings in this plot are from the MT&R workbooks.

Table 4 shows the percentage savings and average annual MWh savings for the program and for HPEM and T&T estimated by the EPT team and reported by BPA. The team calculated the average annual MWh savings for the program and each component as the average annual savings per facility multiplied by the average annual number of evaluated facilities.

Table 4. MT&R and Reported Average Annual Savings by Program Component

Component	Quantity	Facility	Capital	SEM	SEM with Busbar
All	MT&R MWh	31,807	14,247	17,559	19,149
	MT&R %	4.0%	1.8%	2.2%	2.2%
	Reported MWh	N/A	N/A	21,276	23,203
	Reported %	N/A	N/A	2.7%	2.7%
HPEM	MT&R MWh	24,252	13,916	10,336	11,272
	MT&R %	3.5%	2.1%	1.5%	1.5%
	Reported MWh	N/A	N/A	14,053	15,325
	Reported %	N/A	N/A	2.0%	2.0%
T&T	MT&R MWh	10,073	442	9,631	10,504
	MT&R %	6.8%	0.3%	6.5%	6.5%
	Reported MWh	N/A	N/A	9,631	10,504
	Reported %	N/A	N/A	6.5%	6.5%

Source: MT&R model workbooks, annual completion reports, and BPA's EM reporting system.

The first set of rows in Table 4 show the MT&R and reported savings for all program components (HPEM and T&T). The EPT team's MT&R models' estimated average annual savings of 31,807 MWh. Annual average savings from capital projects were 14,247 MWh and savings from SEM were 17,559 MWh.

The second set of rows shows savings for the HPEM component. This component included 26 facilities from 24 sites that started participating in the EM Program in 2010 or 2011. HPEM facilities saved about 3.5% of consumption, or an annual average of 24,252 MWh. Capital project savings were 2.1% of total consumption and SEM savings were 1.5% of total consumption. Twenty HPEM facilities had implemented capital projects during the EM Program participation, which explains the large share of savings from capital projects.

The third set of rows in Table 4 shows savings for the seven T&T facilities, one of which had participated for three years, four of which had participated for two years, and two which had participated for one year. T&T facilities began EM Program participation between 2010 and 2013, and saved approximately 6.8% of consumption, or an annual average of 10,073 MWh. SEM savings were 6.5% of consumption. Because capital project implementation was not a primary objective of T&T, capital projects only contributed savings of 0.3%. The T&T facilities achieved percentage savings approximately equal to those of the HPEM cohort, although T&T savings derived principally from SEM activities and not capital projects.

BPA reported negative SEM savings estimates as zero in its reporting system. This reporting convention caused the MT&R and reported savings to differ. The reported savings corresponded closely to the MT&R savings for the T&T facilities. However, there were significant differences between the MT&R and reported savings for the HPEM cohort. BPA reported zero savings instead of negative for six sites in year one, ten sites in year two, three sites in year three, and four sites in year four. Across all program components, the difference between the MT&R and reported average annual savings was 3,717 MWh or 0.5% of consumption.

3.2 Reported Incremental Annual SEM Savings

Between 2010 - 2014, BPA tracked in its reporting system incremental savings by program year. BPA calculated these incremental savings by subtracting the previous year's annual SEM savings from the current year's annual SEM savings (e.g., 2012 annual savings were subtracted from the 2013 annual savings to determine incremental 2013 SEM savings). Table 5 shows the reported incremental SEM savings.

Table 5. Reported Incremental and Annual SEM Savings by Program Component and by Year

	Incremental SEM Savings (MWh)					Annual SEM Savings (MWh)			
Component	2011*	2012*	2013	2014	Total	2011*	2012*	2013	2014
HPEM 1	4,836	4,125	1,402	5,302	15,665	4,836	8,961	10,363	15,665
HPEM 2	0	4,647	2,699	4,226	11,572	0	4,647	7,346	11,572
T&T	0	922	9,881	1,855	12,658	0	922	10,803	12,658
Total	**4,836**	**9,694**	**13,982**	**11,383**	**39,895**	**4,836**	**14,530**	**28,512**	**39,895**

Source: BPA EM reporting system. Savings include the busbar adjustment, accounting for line losses.

* BPA claimed 75% of the 2011 incremental SEM Savings in 2011 and claimed the remaining 25% of the 2011 SEM savings in 2012.

Note that the sum of the incremental SEM savings across years match the 2014 annual SEM savings. However, the incremental savings and the *average* annual savings in Table 4 differ because the savings in Table 4 are an average of the annual savings.[29] BPA now tracks and reports annual savings, so the evaluation team focused this report on the annual average savings values. We calculated realization rates based on both average annual savings and incremental savings, which are discussed in Section 4: Evaluation Energy Savings Findings.

29 The average annual savings in Table 5 cannot be calculated from the annual savings in Table 6 because the average annual savings were calculated from annual savings values that do not include BPA's adjustment to the 2011 savings where 75% were claimed in 2011 and the remaining 25% were claimed in 2012.

4 Evaluation Energy Savings Findings

Using the savings estimation methodology described in Section 2: Evaluation Methodology, the evaluation team estimated the facility electricity savings, capital project savings, and SEM savings for each facility and year of program participation.

This section also reports evaluation savings estimates for each program component and for all components across each program year. Estimates for each program year are reported in Section 4.3 Year-Over-Year Trends and in Appendix F. We also calculated realization rates by comparing the evaluation savings estimates with the MT&R and reported savings estimates described in Section 3: Program MT&R and Reported Savings. The reported savings are the final record of program achievement.

We calculated confidence intervals for facility savings but not for SEM savings, because standard errors for capital project savings estimates were not available to determine SEM savings uncertainty.

4.1 Program-Level Evaluation Results

Across all evaluated facilities and participation years, the EM Program saved 4.1% of electricity consumption, which equates to average annual savings of 32,924 MWh. After subtracting capital project savings funded through other energy efficiency programs of 1.8% from facility savings, the evaluation team estimated that the BPA EM Program saved 2.3% of energy consumption, or average annual savings of 18,687 MWh (20,379 MWh when adjusted for busbar).[30]

Figure 7 and Table 6 show the evaluation estimates of the average annual MWh savings, percentage savings, and realization rates for evaluated EM Program facilities.[31] The busbar adjusted savings account for line losses.

30 The evaluation team did not independently estimate the capital project savings. The team obtained capital project savings estimates for evaluated facilities from the MT&R reports.

31 As noted above, average annual savings are the weighted average of annual savings per facility, with weights equal to the number of facilities evaluated in each program year. Percentage savings are the sum of annual savings for all program years divided by the sum of annual consumption for all program years.

Figure 7. Average Annual SEM and Capital Savings Across All Program Components

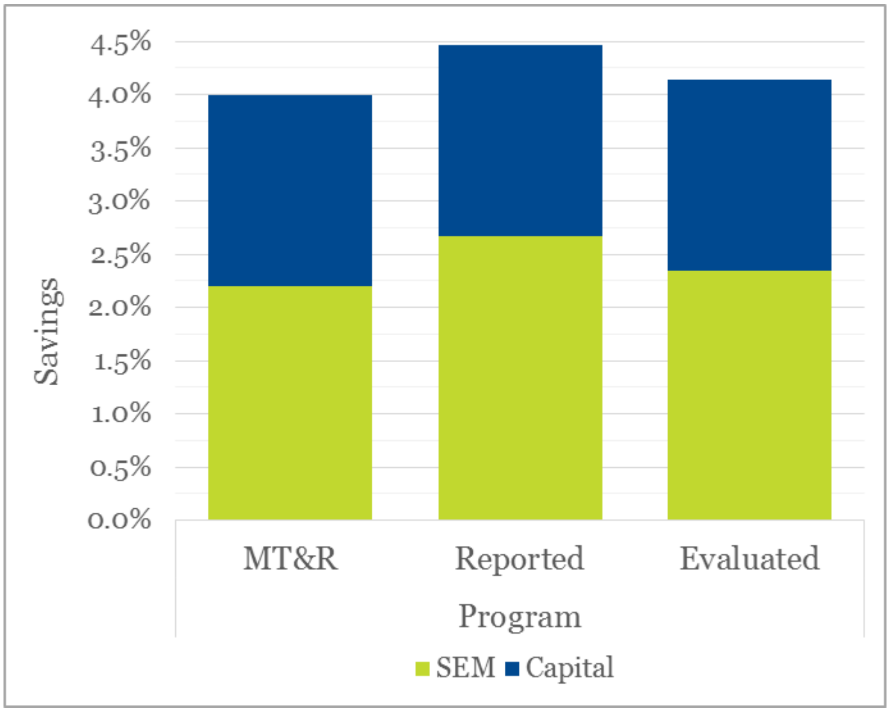


Table 6 provides estimates of uncertainty for the facility savings estimates. The 80% confidence interval for the evaluated facility savings was ±2,829 MWh and included the MT&R facility savings estimate of 31,807 MWh. The evaluation team did not calculate savings uncertainty for the SEM savings estimates, since uncertainty estimates for capital project savings were unavailable.

Table 6. Energy Management Program MT&R, Reported, and Evaluated Savings

	Average Annual Savings			
All Program Components	Facility	Capital*	SEM	SEM with Busbar
MT&R MWh Savings	31,807	14,247	17,559	19,149
MT&R % Savings	4.0%	1.8%	2.2%	2.2%
Reported MWh Savings	N/A	N/A	21,276	23,203
Reported % Savings	N/A	N/A	2.7%	2.7%
Evaluated MWh Savings	32,924	14,237	18,687	20,379
Evaluated % Savings	4.1%	1.8%	2.3%	2.3%
80% Confidence Interval (MWh)**	± 2,829	N/A	N/A	N/A
80% Confidence Interval (%)**	± 0.4%	N/A	N/A	N/A
Realization Rate				
Evaluated / MT&R	1.04	N/A	1.06	1.06
Evaluated / Reported	N/A	N/A	0.88	0.88

* The EPT team pro-rated capital savings for the number of days that the equipment was operational during the SEM period. The evaluation team adjusted this pro-rating in some instances, resulting in lower capital savings than that documented in the MT&R reports.

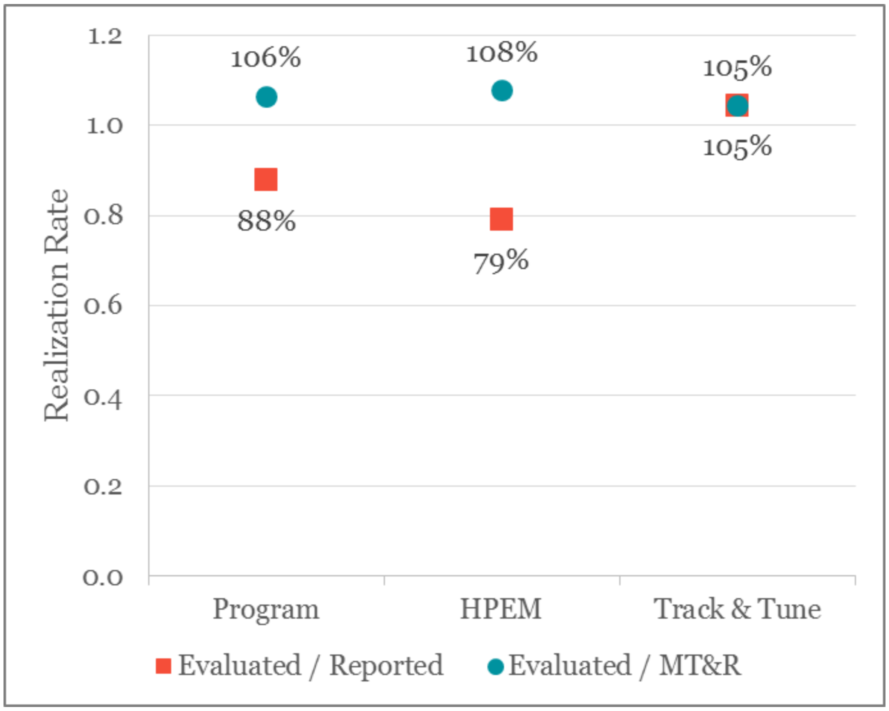
** The team only calculated confidence intervals around facility savings. It was not possible to calculate the confidence intervals around the SEM savings because the uncertainty around the capital measure savings estimates is unknown.

As previously mentioned, BPA reported zero savings instead of negative savings for facilities with estimated consumption increases. Due to this difference in MT&R and reported savings, the evaluation team calculated two sets of realization rates. The first is the ratio of evaluation savings to MT&R savings. The second is the ratio of evaluation savings to reported savings. Realization rates greater than 1.0 indicate that the evaluation savings exceeded the MT&R or reported savings.

The realization rates for the MT&R savings and reported savings are shown in Figure 8 for the overall program and for the HPEM and T&T components. The evaluation team found slightly higher SEM savings than the EPT team, resulting in an overall MT&R savings realization rate of 1.06. However, the evaluation team found fewer SEM savings than reported, with a savings realization rate of 0.88.[32] The HPEM and T&T realization rates are discussed in the following section.

32 These realization rates apply to evaluated facilities between 2010 and 2014 and may not represent the current or future performance of the EM program population. BPA has adopted a policy of estimating new baselines after every two years of participation. It is possible that some estimates of negative savings in this evaluation would not have been obtained under BPA's new policy.

Figure 8. Realization Rates of SEM Savings by Program Component



4.2 Program Component Results

The evaluation team estimated savings for each of the HPEM and T&T components, with the results outlined here.

4.2.1 HPEM

The HPEM component achieved facility savings of approximately 3.7% of consumption, or an average of 25,042 MWh per year. Figure 9 and Table 7 present evaluation savings estimates for the HPEM cohorts. The 80% confidence interval for the evaluation facility savings estimate of ±2,809 MWh contains the MT&R savings estimate. Capital projects incentivized through other energy efficiency programs accounted for approximately 2.1% of consumption, or 13,906 MWh per year. EM Program activity saved approximately 1.6% of consumption, or about 11,136 MWh per year (12,144 MWh adjusted for busbar). As Figure 8 shows, the HPEM SEM savings realization rates were 1.08 for the MT&R savings and 0.79 for the reported savings. The savings realization rate for reported savings was lower because negative savings estimates were recorded as zeros.

Figure 9. Average Annual SEM and Capital Savings for HPEM

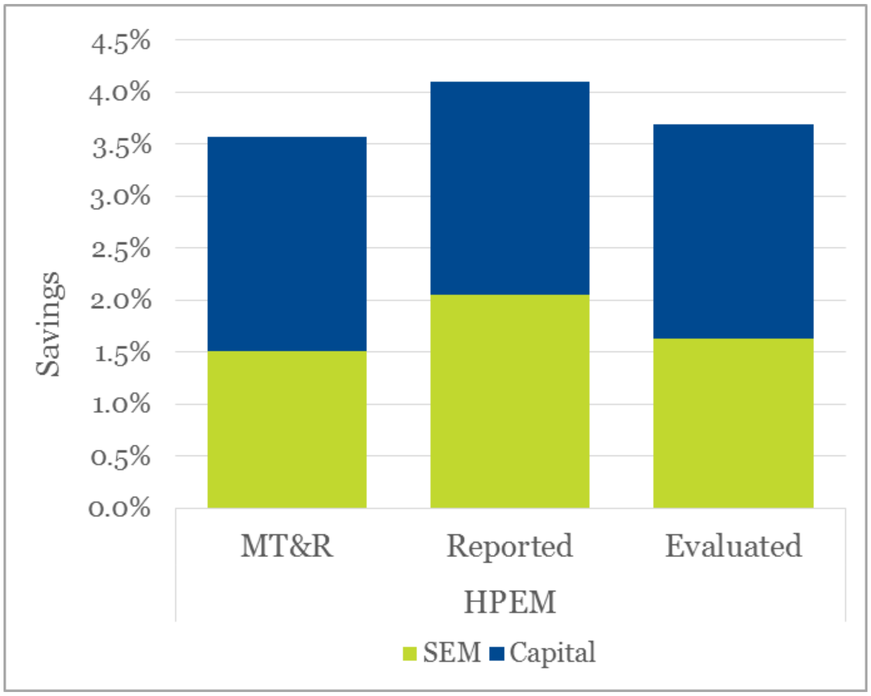


Table 7. HPEM Component MT&R, Reported, and Evaluated Savings

HPEM	Average Annual Savings			
	Facility	Capital*	SEM	SEM with Busbar
MT&R MWh Savings	24,252	13,916	10,336	11,272
MT&R % Savings	3.5%	2.1%	1.5%	1.5%
Reported MWh Savings	N/A	N/A	14,053	15,325
Reported % Savings	N/A	N/A	2.0%	2.0%
Evaluation MWh Savings	25,042	13,906	11,136	12,144
Evaluation % Savings	3.7%	2.1%	1.6%	1.6%
80% Confidence Interval (MWh)**	± 2,809	N/A	N/A	N/A
80% Confidence Interval (%)**	± 0.5%	N/A	N/A	N/A
Realization Rate				
Evaluation / MT&R	1.03	N/A	1.08	1.08
Evaluation / Reported	N/A	N/A	0.79	0.79

* The EPT team prorated capital savings for the number of days the equipment was operational during the SEM period. The evaluation team adjusted this prorating in some instances, resulting in lower capital savings than that documented in the MT&R reports.

** The evaluation team only calculated confidence intervals around facility savings. It was not possible to calculate the confidence intervals around the SEM savings because the uncertainty around the capital measure savings estimates is unknown.

The HPEM 1 cohort, which began the program in 2010, achieved higher savings than the HPEM 2 cohort, which began the program in 2011. The HPEM 1 cohort saved 7.1% of facility consumption. SEM savings were 3.0%. The HPEM 2 cohort saved 1.5% of facility consumption. SEM savings were 0.8%. This difference in savings could have been due to the different types of facilities in each cohort. The HPEM 2 cohort facilities were larger and had more complex production processes and could have required more time to implement SEM activities.

4.2.2 Track and Tune

The evaluation team estimated that energy savings for the T&T facilities were 7.1% of electricity consumption, or 10,510 MWh of average annual savings. The 80% confidence interval for the evaluated facility savings of ±452 MWh contained the MT&R and reported facility savings estimate.

Figure 10 and Table 8 present evaluation savings estimates for the T&T facilities.

For the T&T facilities, EM Program activity was responsible for almost all facility savings. Only 0.3%, or about 442 MWh of average annual savings, was attributable to capital projects incentivized by other efficiency programs. After accounting for capital projects, T&T facilities saved approximately 6.8% of consumption, or 10,068 MWh of average annual savings (10,980 MWh adjusted for busbar).

The MT&R and reported SEM savings realization rates for T&T facilities were larger than those for the HPEM cohorts. As the results in Section 4.4: Facility-Level Savings Estimates show, the savings realization rate was greater than 1.0 because the evaluated savings were significantly higher than the MT&R savings for one T&T facility.

Figure 10. Average Annual SEM and Capital Savings for T&T Facilities

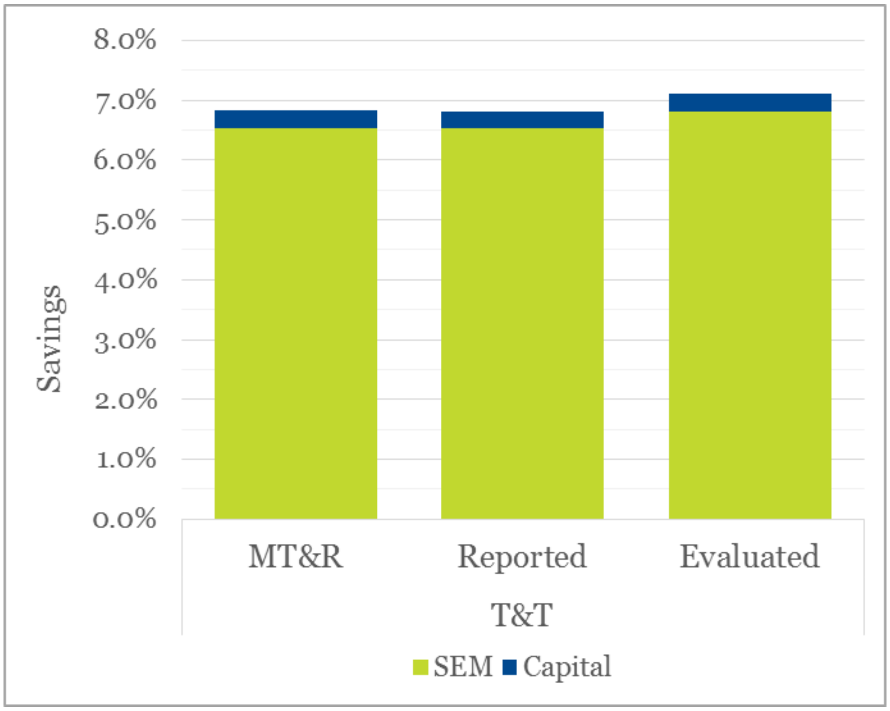


Table 8. T&T Facilities Evaluated Savings

T&T	Average Annual Savings Facility	Capital*	SEM	SEM with busbar
MT&R MWh Savings	10,073	442	9,631	10,504
MT&R % Savings	6.8%	0.3%	6.55%	6.55%
Reported MWh Savings	N/A	N/A	9,631	10,504
Reported % Savings	N/A	N/A	6.55%	6.55%
Evaluation MWh Savings	10,510	442	10,068	10,980
Evaluation % Savings	7.1%	0.3%	6.8%	6.8%
80% Confidence Interval (MWh)**	452	N/A	N/A	N/A
80% Confidence Interval (%)**	0.3%	N/A	N/A	N/A
Realization Rate				
Evaluation / MT&R	1.04	N/A	1.05	1.05
Evaluation / Reported	N/A	N/A	1.05	1.05

* The EPT team prorated capital savings for the number of days the equipment was operational during the SEM period. The evaluation team adjusted this prorating in some instances, resulting in lower capital savings than that documented in the MT&R reports.

** The evaluation team only calculated confidence intervals around facility savings. It was not possible to calculate the confidence intervals around the SEM savings because the uncertainty around the capital measure savings estimates is unknown.

4.3 Year-Over-Year Trends

The savings estimates presented thus far have reflected average annual program or program component performance over several years. However, an integral part of SEM programs is the emphasis on making continuous improvements in facility energy efficiency over time. To gauge whether EM Program facilities made year-over-year improvements in efficiency, the evaluation team estimated savings by year of participation.

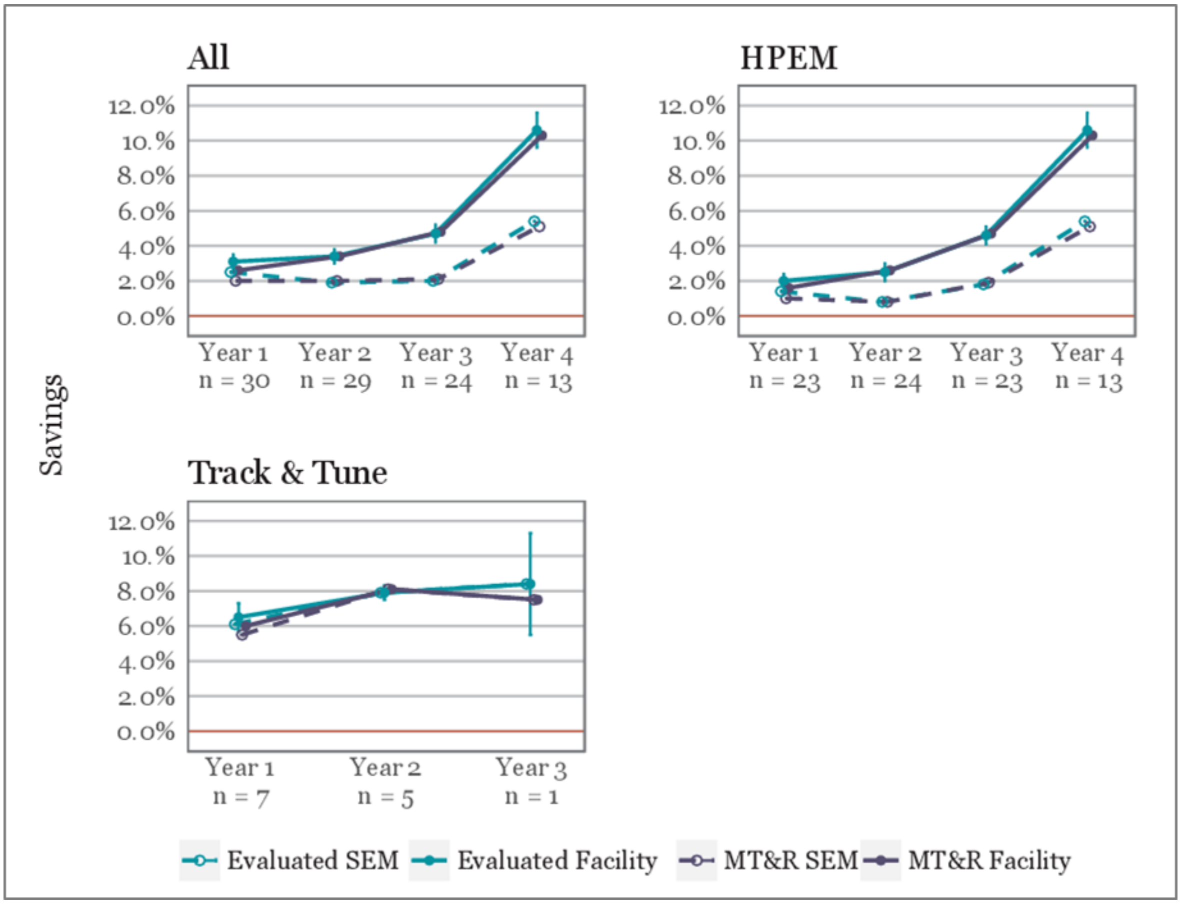
Figure 11 shows evaluated facility and SEM percentage savings, along with MT&R facility and SEM percentage savings, by year of participation.[33] In general, there was close equivalence between the evaluation and MT&R savings estimates for each program component and year.[34]

Evaluated facility savings as a percentage of consumption increased over time. Figure 11 displays an upward trend in average percentage savings for the program and each program component. However, as the number of evaluated facilities changed over time, it cannot be concluded that EM Program facilities made year-over-year incremental efficiency improvements. The upward trends could have reflected the change in sample composition rather than actual increases in annual savings.

33 As described in Section 1: Introduction, participants joined the program at different times. Therefore, the participation-year savings estimates reported here (and in Appendix F) do not correspond to a particular calendar year or to a BPA's program year. For example, the Year 1 HPEM savings would include savings from both the HPEM 1 cohort's first year of participation in 2010 and the HPEM 2 cohort's first year of participation in 2011.

34 The exception was for T&T in year 3, when there was a big difference between the evaluated and MT&R savings estimates for one facility.

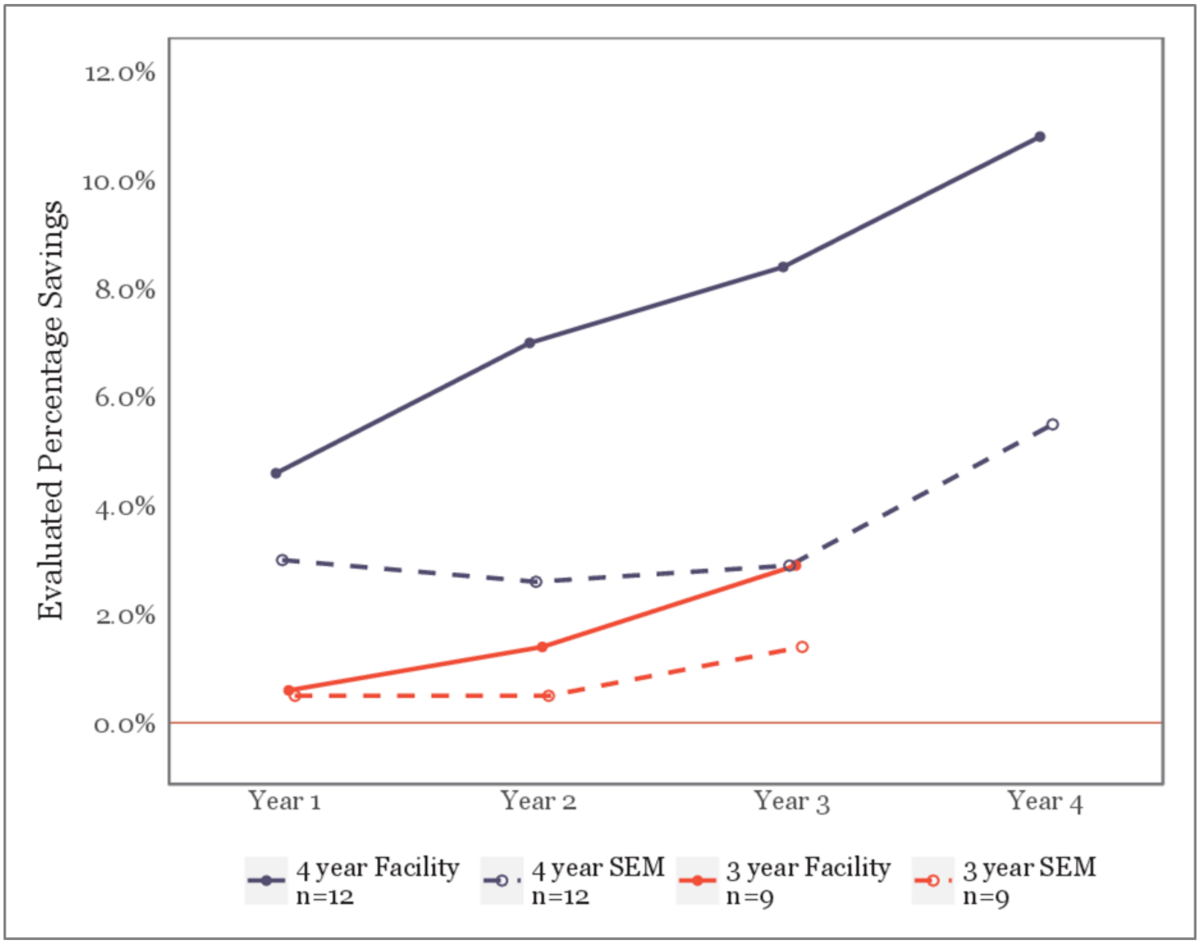
Figure 11. Evaluated and MT&R Savings by Program Component and Year



Note: the error band around the evaluated facility savings indicates whether the MT&R savings were within the evaluation savings 80% confidence interval.

To better assess trends and the persistence of SEM savings during program participation, the evaluation team examined savings of HPEM facilities that had participated in the program for similar duration. There were nine HPEM facilities with evaluated savings for three program years and 13 HPEM facilities with evaluated savings for four program years, allowing us to observe savings trends for the same group of facilities (shown in Figure 12).

Figure 12. Evaluated Savings by Program Year for HPEM Cohorts



Note: The error band around the evaluated facility savings indicates whether the MT&R savings were within the evaluation savings 80% confidence interval.

In HPEM facilities participating for three years, estimated SEM savings increased from 0.5% in year 1 to 1.4% in year 3. In HPEM facilities participating for four years, estimated savings increased from 3.0% in year 1 to about 5.2% in year 4. There was a small decrease in evaluated SEM savings between year 1 and year 2, then an increase in savings during the next two years.

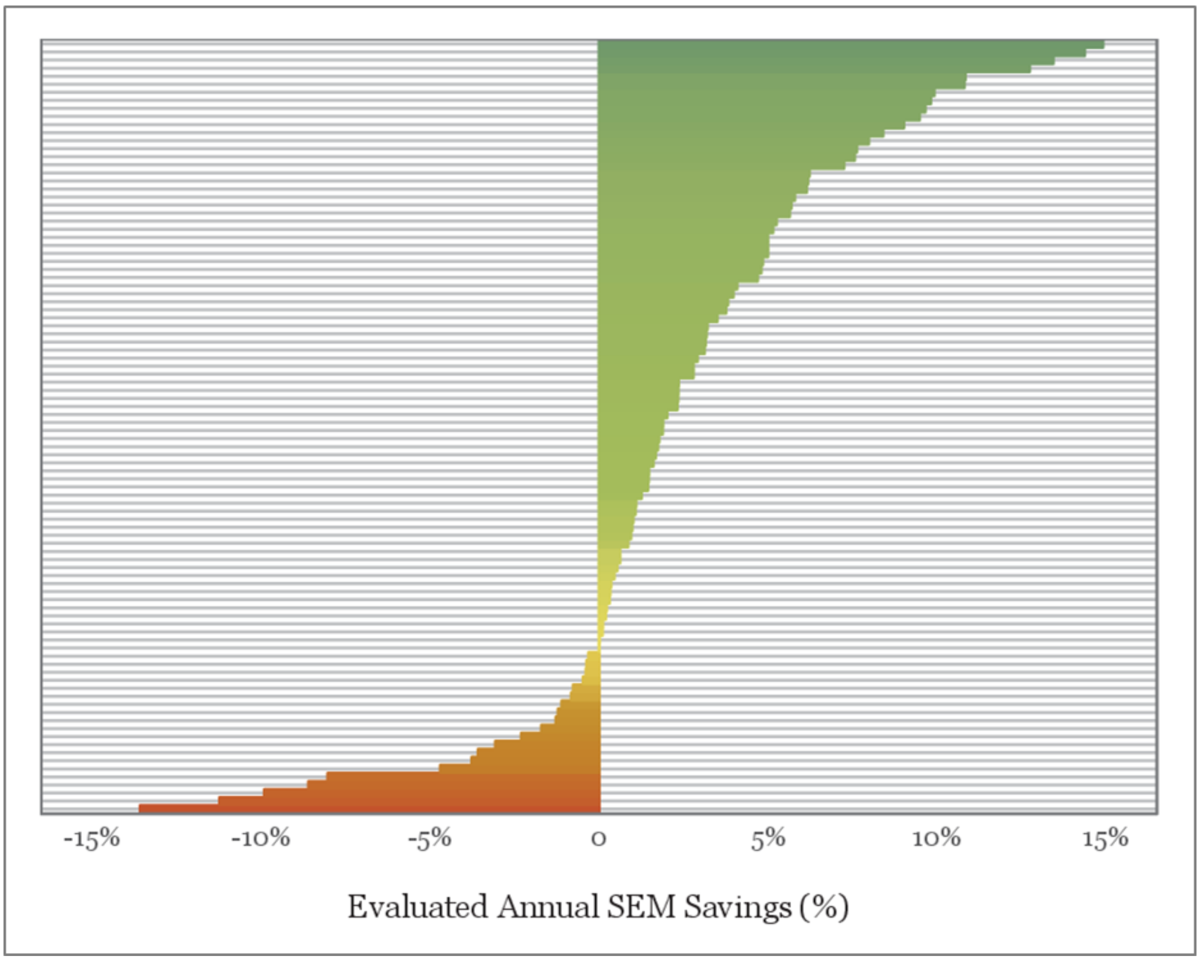
Overall, SEM savings as percentage of consumption in HPEM facilities appears to have persisted over the first three or four program years. We did not find evidence that annual savings decayed over time.

4.4 Facility-Level Savings Estimates

The program-level results showed substantial and statistically significant SEM savings of 2.3%. Yet, averages and totals can mask significant variation in savings between facilities and across years. This section summarizes annual savings estimates for individual facilities.

Figure 13 summarizes the cross-sectional and time series variation in SEM savings. Each bar represents an estimate of SEM percentage savings for a facility in one year. The facility annual SEM percentage savings ranged from approximately negative 14% to positive 15%. However, there were many more facilities and years with positive than negative estimated percentage savings.

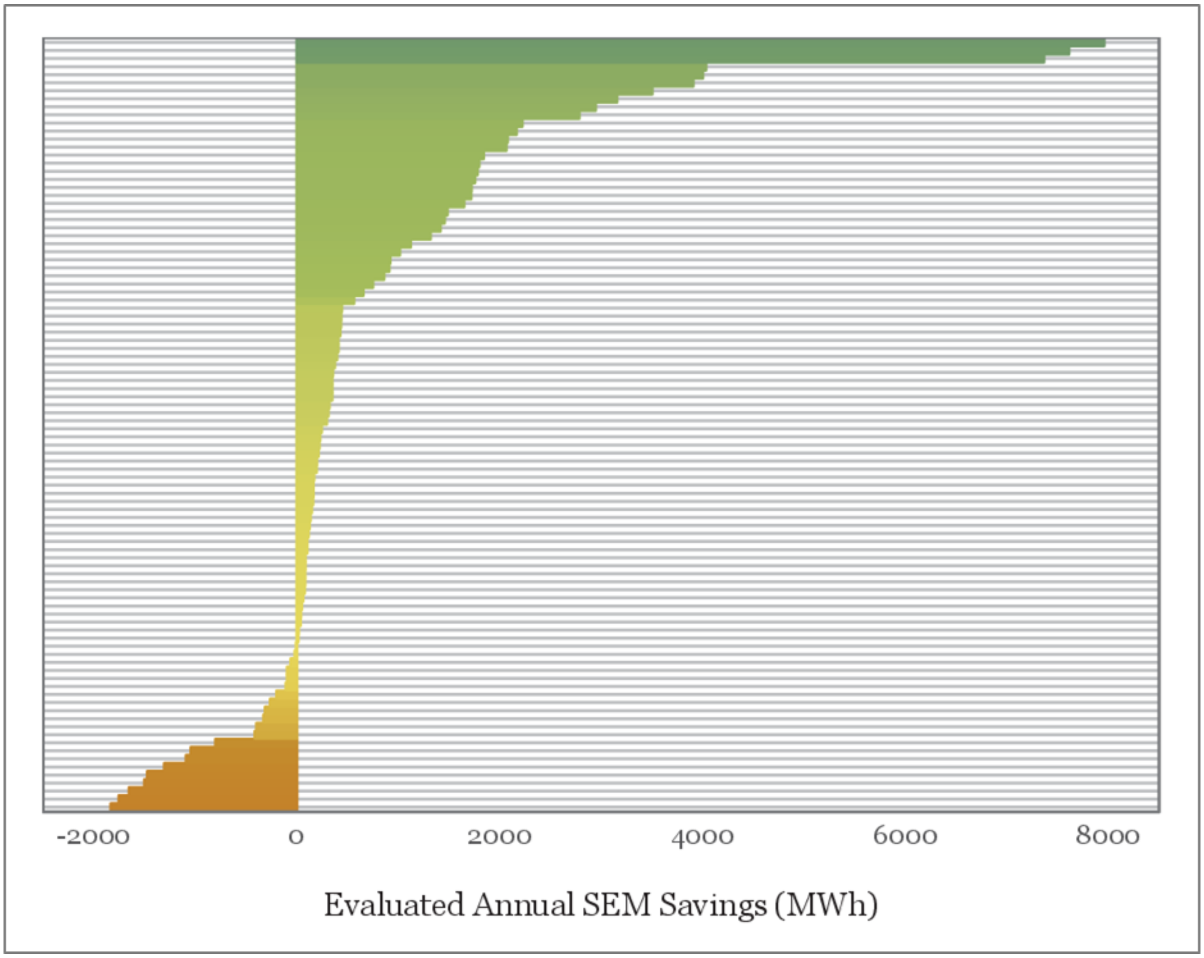
Figure 13. Individual Facility SEM Percentage Savings Estimates



Note: Each line indicates an SEM savings estimate for a facility and year.

Figure 14 shows the variation across facilities and years in SEM MWh savings. Each bar represents an estimate of SEM MWh savings for a facility in a year. The areas to the left and right of zero represent, respectively, the negative MWh savings estimates and the positive MWh savings estimates. It is evident that the area to the right, the positive MWh savings, far exceeds the areas to the left, the negative MWh savings, and that the program saved electricity overall.

Figure 14. Individual Facility SEM MWh Savings Estimates

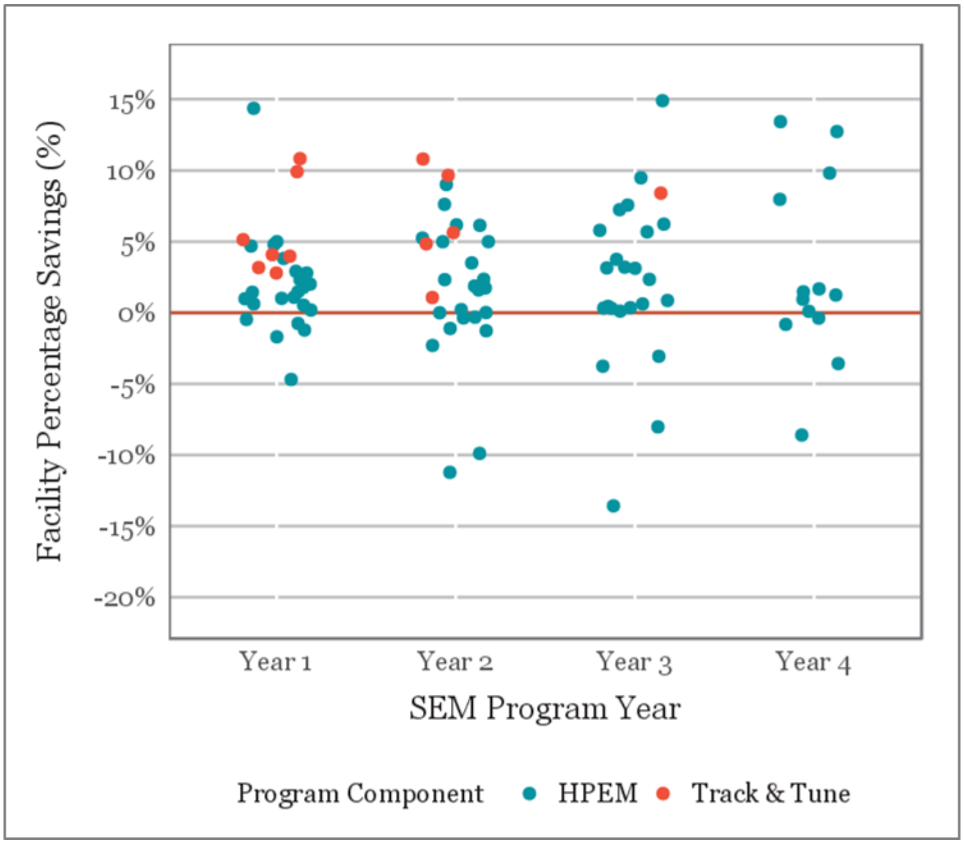


Note: Each line indicates an SEM savings estimate for a facility and year.

Figure 15 summarizes the variation in savings between facilities by program year, showing boxplots of the evaluation estimates of individual facility annual SEM percentage savings. Savings of individual facilities are shown as dots, with the color denoting the facility's program component.

Again, there was significant variation between facilities in estimated savings in each year. This likely reflected annual variation in facility savings performance, electricity consumption, and uncertainty of the savings estimates. This variation appears to have increased over time, shown by the increased dispersion of percentage savings from year to year.[35] Across all evaluated facilities and years, the coefficient of variation of percentage savings (the ratio of the sample standard deviation to sample mean) equaled 2.01.

Figure 15. Individual Facility SEM Percentage Savings Estimates by Program Year

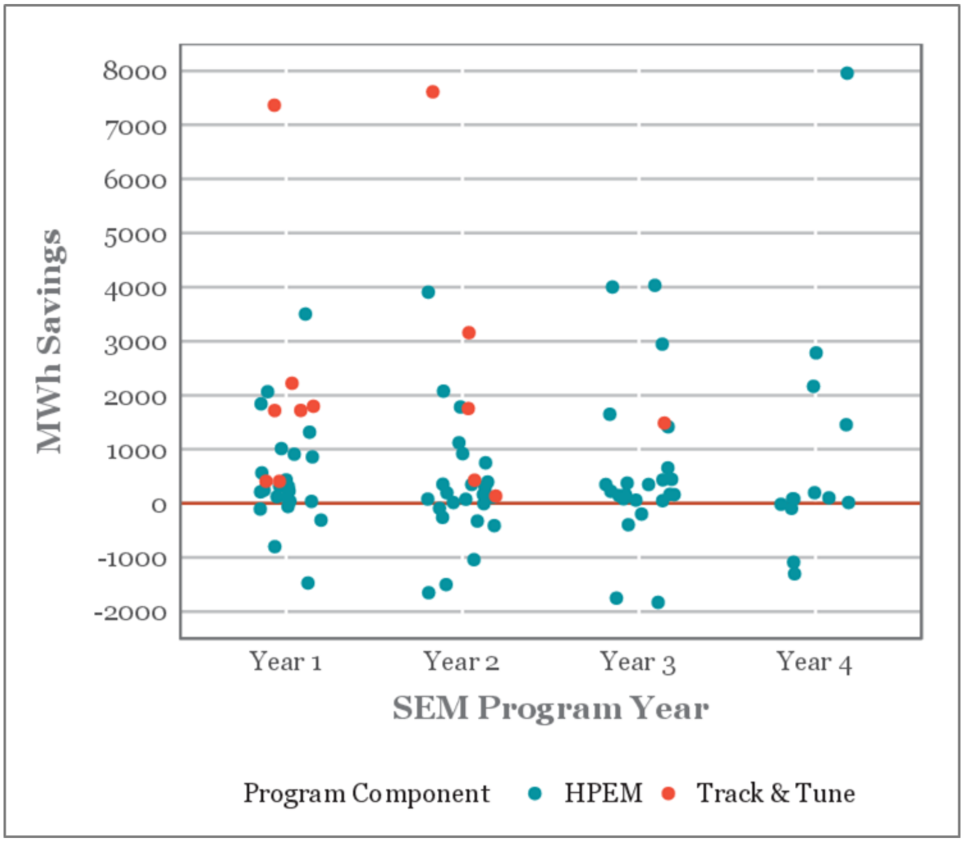


Note: Each dot represents the annual SEM savings for an individual facility in a program year.

Figure 16 presents a boxplot for the SEM MWh savings by program year. The increasing trend in the variability of MWh savings over time is not as evident because MWh savings reflected both the effectiveness of SEM implementation as well as the level of electricity consumption.

35 The number of facilities also decreased over time. Program year 4 only includes HPEM facilities that started the program in 2010.

Figure 16. Individual Facility SEM MWh Savings Estimates by Program Year



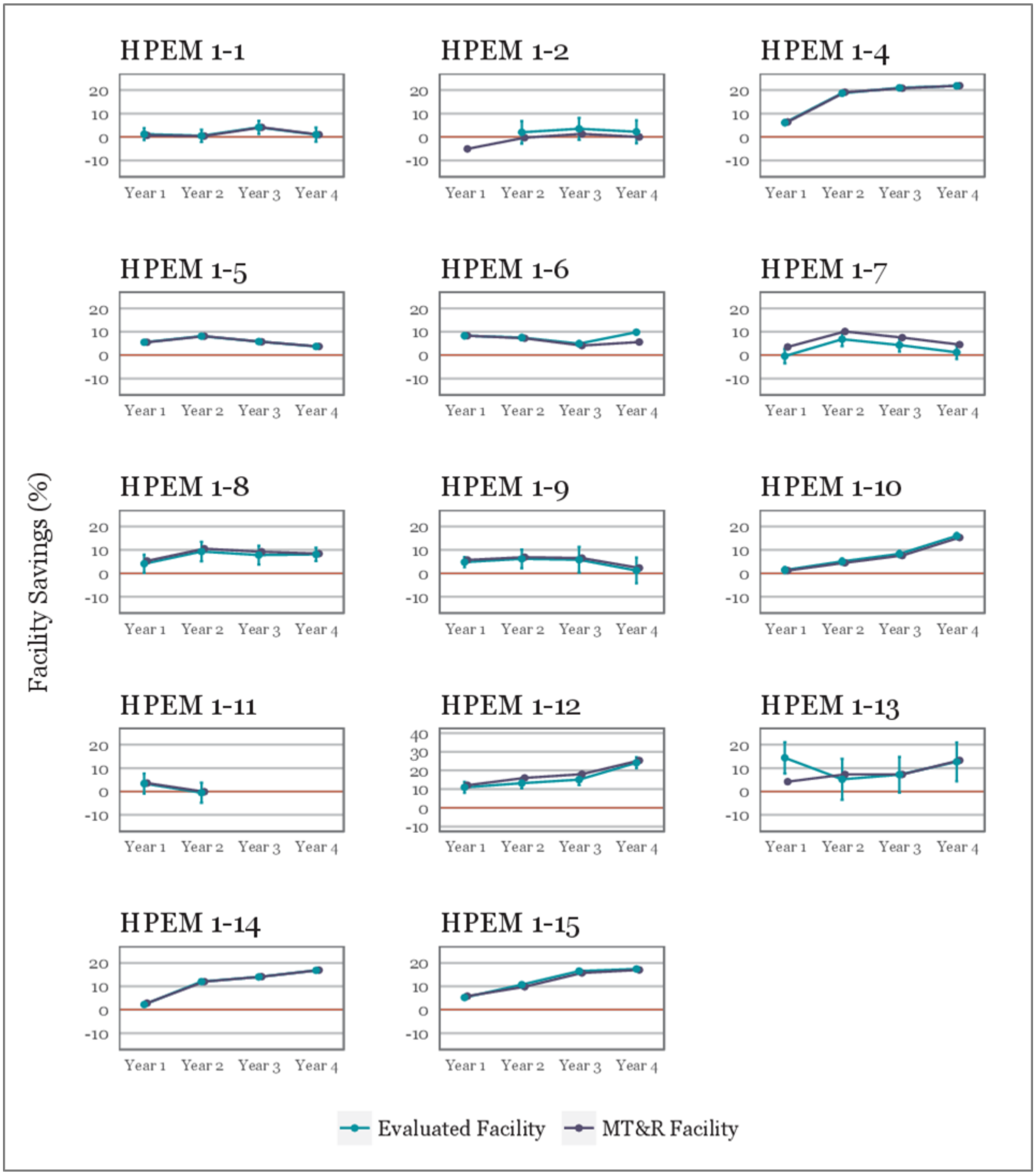
Note: Each dot represents the annual SEM savings for an individual facility in a program year.

4.4.1 Within-Facility Annual Facility Savings

There was also significant variation of annual savings for individual facilities. While many EM facilities increased facility or SEM savings each year, many facilities exhibited seesawing or downward trends in savings.

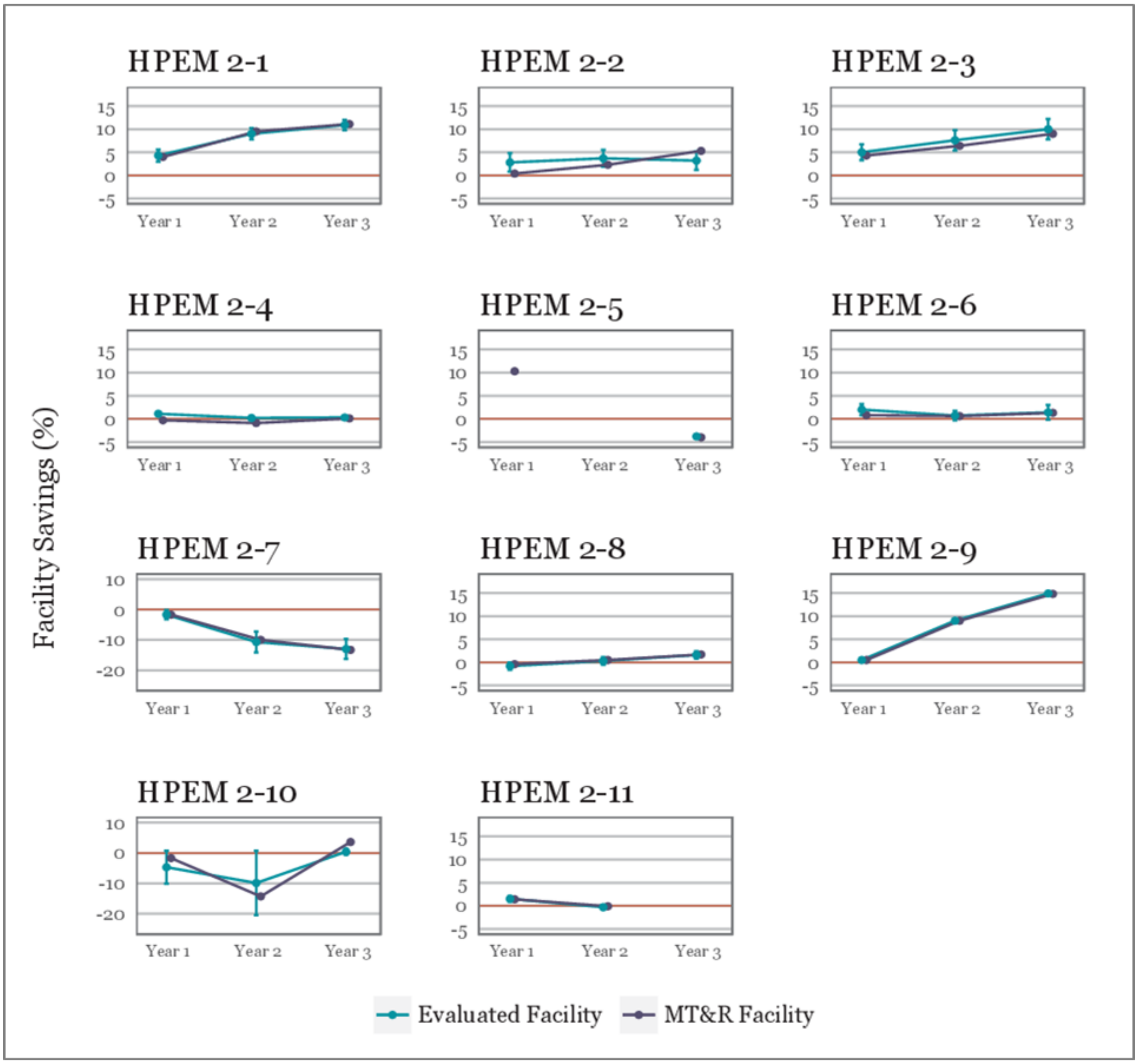
Figure 17, Figure 18, and Figure 19 display annual facility savings for each evaluated facility, by program component, showing the evaluation and MT&R facility percentage savings estimates for the HPEM cohort (HPEM 1 and 2) and T&T facilities. The blue lines represent the evaluated savings, and the green lines represent the MT&R savings. The vertical bars indicate 80% confidence intervals for the evaluated facility savings. Comparison of the lines shows the difference between evaluated and MT&R facility savings before removal of any capital projects savings.

Figure 17. HPEM 1 Cohort Evaluation Versus MT&R Percentage Savings Panel



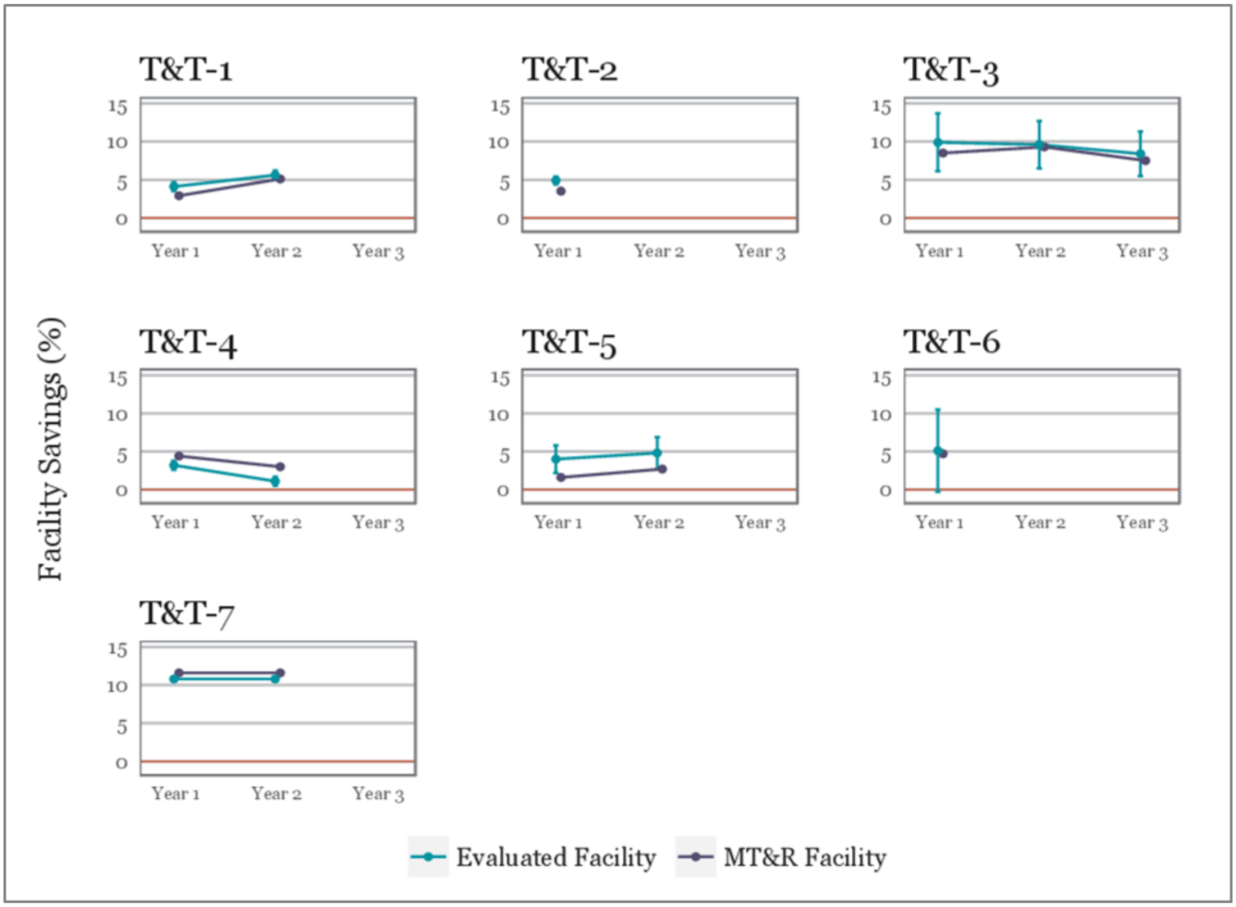
Note: Vertical axis scaling may differ between adjacent plots.

Figure 18. HPEM 2 Cohort Evaluation Versus MT&R Percentage Savings Panel



Notes: Vertical axis scaling may differ between adjacent plots. HPEM 2-5 was not evaluable until year 3.

Figure 19. T&T Facilities Evaluation Versus MT&R Percentage Savings Panel



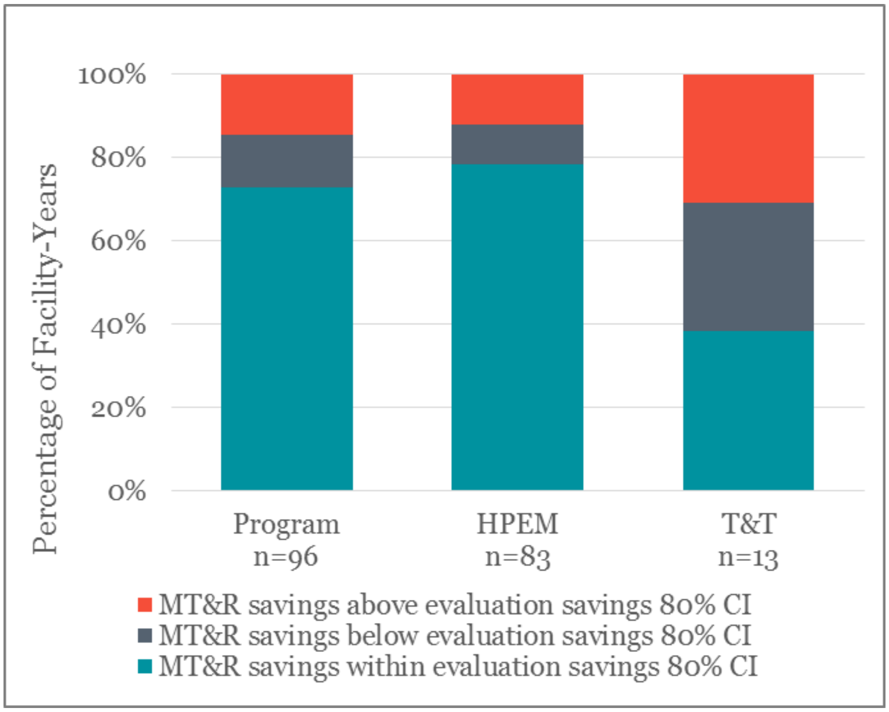
Notes: Vertical axis scaling may differ between adjacent plots. T&T-2 and T&T-6 only participated for one year during the evaluation period.

Out of 29 facilities with more than one year of evaluated savings, 38% increased percentage savings each year, 21% decreased percentage savings each year, and 41% had seesawing percentage savings that increased in some years and decreased in others. For example, Facility 7 of the HPEM 1 cohort experienced an increase in percentage savings between year 1 and year 2, then had successive decreases during the following two years. Decreases in percentage savings could reflect either that some facilities backslid in implementing SEM activities and achieved smaller kWh savings or that facility consumption increased relatively more than savings. It could also reflect uncertainty of the facility savings estimates.

In addition to showing variation of annual savings for individual facilities, Figure 17, Figure 18, and Figure 19 demonstrate that the EPT team and evaluation team obtained similar savings estimates for individual facilities. There were only small differences in estimated facility percentage savings, and the 80% confidence intervals for evaluated savings typically contained the MT&R savings estimate. Figure 20 shows that for the majority of facilities and program years, the evaluated facility savings estimate was not statistically different from the MT&R facility savings estimate.

In 73% of facility-years, the confidence interval for the evaluated savings contained the MT&R savings estimate.[36] In 15% of facility-years, the MT&R savings estimate was above the confidence interval, and in 12% of facility-years, the MT&R savings estimate was below the confidence interval. However, across sites and years, the difference in estimated savings averaged close to zero, as suggested by the MT&R savings realization rate of about 1.0. Appendix G shows the corresponding counts of facilities where the MT&R savings were within, above, or below the evaluation savings 80% confidence interval.

Figure 20. MT&R Estimates Relative to Evaluation Confidence Regions



4.4.2 Within-Facility Annual SEM Savings

The evaluation team estimated SEM savings as the difference between the estimated facility savings and the capital project savings. To show the effect of subtracting capital project savings from facility savings, Figure 21, Figure 22, and Figure 23 show evaluated facility and evaluated SEM MWh savings for individual facilities in, respectively, the HPEM 1, HPEM 2, and T&T components. The solid lines represent facility savings and the dashed lines represent SEM savings. The difference between these lines represents the

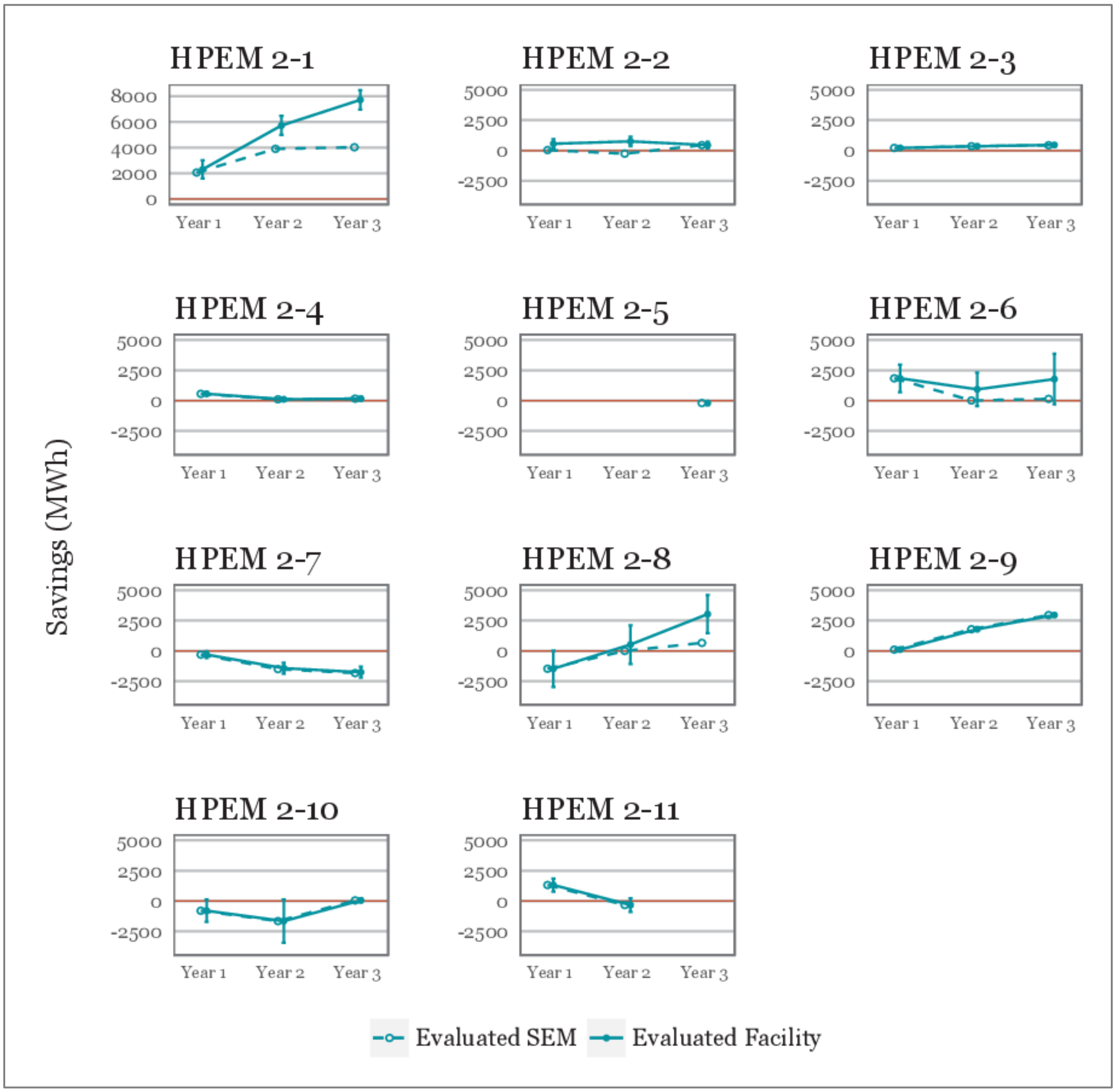
36 A facility-year savings estimate is the estimate of savings for a facility during one particular participation year.

estimated capital project savings. Facilities with overlapping lines did not implement any capital projects that were incentivized through other utility energy efficiency programs. These figures also show a wide variety of SEM savings trends, including upward, downward, flat, and seesawing savings.

Figure 21. HPEM 1 Cohort Evaluation Facility Versus SEM MWh Savings Panel

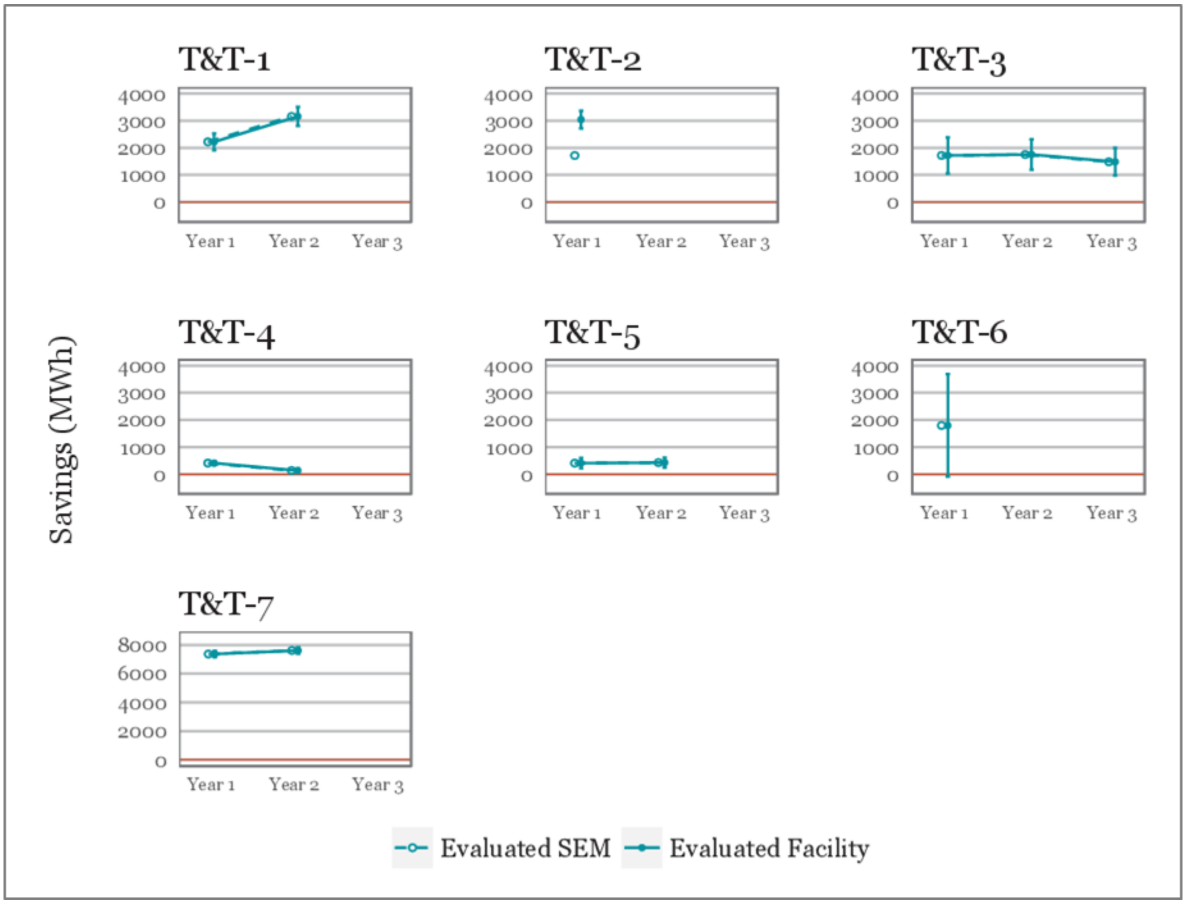
Note: Vertical axis scaling may differ between adjacent plots.

Figure 22. HPEM 2 Cohort Evaluation Facility Versus SEM MWh Savings Panel



Note: Vertical axis scaling may differ between adjacent plots.

Figure 23. T&T Evaluation Facility Versus SEM MWh Savings Panel

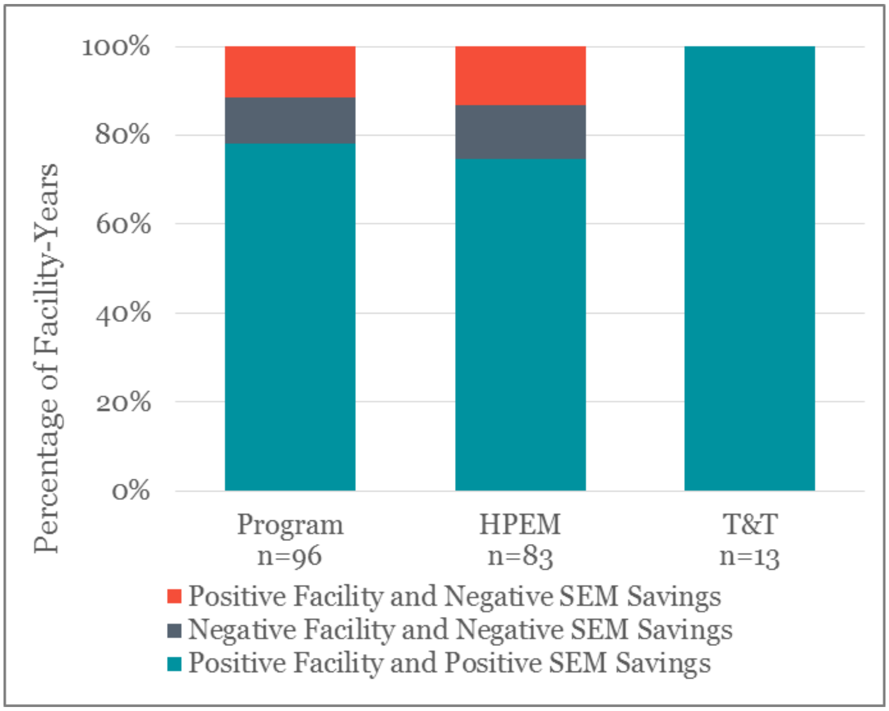


Notes: Vertical axis scaling may differ between adjacent plots. T&T-2 and T&T-6 only participated for one year during the evaluation period.

The SEM savings estimate was positive if the facility savings estimate exceeded the capital project savings and negative if the opposite were true. Although facilities exhibited a variety of MWh savings trends, the evaluated savings were positive in most years. As Figure 24 shows, in 78% of facility-years, the evaluation team estimated positive facility savings and positive SEM savings. In approximately 12% of facility-years, the evaluation team estimated positive facility savings but negative SEM savings after subtracting capital project savings. In the remaining 10% of facility-years, the team estimated that facility electricity consumption increased (i.e., the estimates of facility and SEM savings were negative).[37]

37 Appendix H shows counts of positive and negative SEM savings estimates.

Figure 24. Summary of Frequency of Facilities with Negative Savings Estimates



Estimated increases in consumption likely reflect difficulties in the measurement of savings because of omitted variables, degradation in capital equipment performance, or unaccounted for non-programmatic effects—not that the program caused consumption to increase. However, an increase in facility consumption (e.g., because of a program implementation error) cannot be ruled out. As there is no accepted method for differentiating between omitted variables and a program causal effect, the evaluation results included estimated consumption increases.

Appendix K discusses the evaluation team's treatment of negative savings estimates in greater depth.

4.5 Incremental Annual SEM Savings

The evaluation team calculated incremental annual savings for comparison with the BPA EM reporting system, as described in Section 3.2: Reported Incremental Annual SEM Savings. For the years included in this evaluation (2010 through 2014), the BPA EM reporting system tracked incremental annual SEM savings. However, BPA also tracks annual savings (which do not subtract the previous year's savings), and the evaluation team focused this report on the annual savings. We calculated realization rates based on both average annual savings and incremental savings so that BPA can apply realization rates retrospectively or prospectively.

Table 9 shows the total incremental annual savings from the BPA EM reporting system. The evaluation team calculated incremental savings based on the MT&R savings and evaluated savings. MT&R savings differ from reported savings because BPA does not report estimates of increased consumption (negative savings). The realization rate comparing the evaluated to MT&R incremental savings was 1.02. The realization rate comparing the evaluated to reported incremental savings was 0.81.

Table 9. Incremental Annual SEM Savings and Realization Rates by Program Component

Component	Reported Total Incremental SEM Savings (MWh)*	MT&R Total Incremental SEM Savings (MWh)	Evaluated Total Incremental SEM Savings (MWh)	Realization Rate (Evaluated/ Reported)	Realization Rate (Evaluated/ MT&R)
HPEM	27,237	15,671	15,649	0.57	1.00
T&T	12,658	15,776	16,496	1.30	1.05
Total	**39,895**	**31,447**	**32,144**	**0.81**	**1.02**

Note: Savings include the busbar adjustment, accounting for line losses.

* Source: BPA EE reporting system.

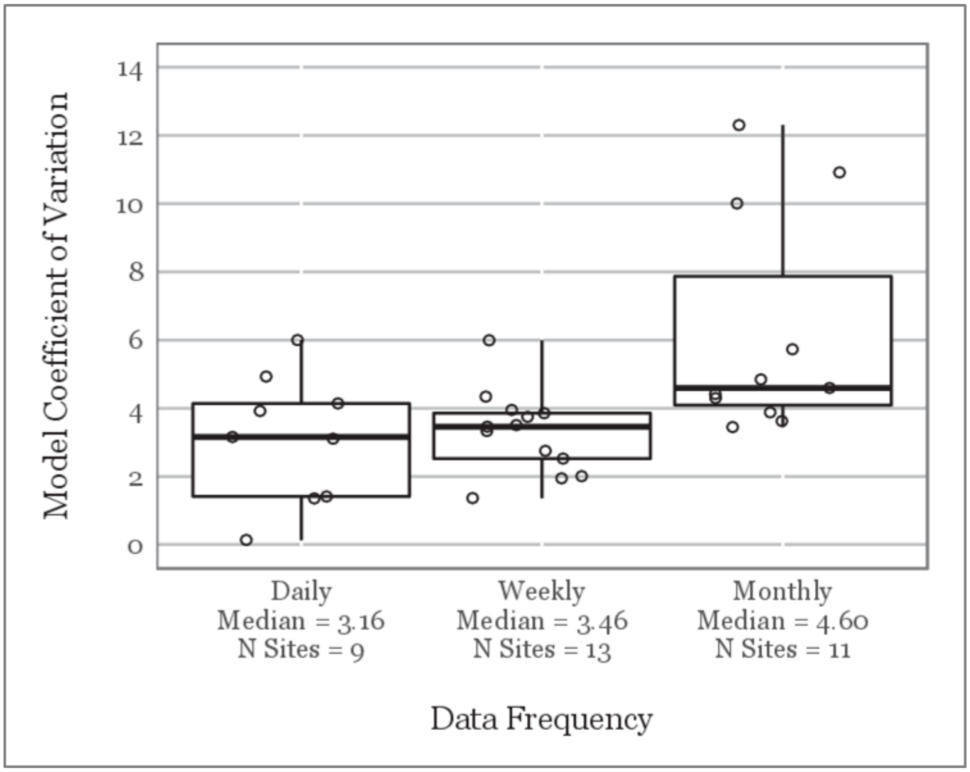
4.6 Model Uncertainty and Data Frequency

The evaluation team studied whether the precision of the energy savings estimates improved when energy consumption data were available at higher frequencies (i.e., daily or weekly rather than monthly). For each facility, the evaluation team calculated the regression coefficient of variation (CV), which is the ratio of the model root mean square error to mean response.[38] A large CV indicates a model with high prediction uncertainty. A low regression CV indicates that the model can explain more of the variation in facility energy consumption. When a model explains most of the variation in a facility's energy consumption, there is greater likelihood of detecting savings statistically. The evaluation team computed all model CVs from regressions estimated with baseline period data.

Figure 25 shows boxplots of the model CV for evaluated facilities by the frequency of the facility energy consumption data. The boxplot shows the quartiles, where the middle band represents the median. There were 9 facilities with daily energy consumption data, 13 facilities with weekly energy consumption data, and 11 facilities with monthly or bi-monthly data. The median CV was 3.2 for daily models and 3.5 for weekly models; both of which were much lower than the median CV for monthly models of 4.6. This suggests that daily and weekly models may better explain facility energy consumption.

[38] The regression model CV is a unit-less measure of model variability. CV for a regression model is calculated as $100 \times \frac{RMSE}{KWH_{avg}}$, where $RMSE$ is the root mean squared error of the regression model and KWH_{avg} is the average energy usage across all periods used in the model.

Figure 25. Model Coefficient of Variation by Frequency of Data



Note: The 25th, 50th (median), and 75th savings percentiles are the top, middle, and bottom horizontal lines of the box, respectively. The 10th and 90th savings percentiles are represented by the endpoints of the vertical lines.

These results suggest that program managers and evaluators should attempt to collect high frequency energy consumption data whenever possible. Sometimes, however, production data will be the limiting factor, as they may only be available at lower frequencies.

5 SEM Adoption

The evaluation team conducted an SEM adoption assessment to determine the extent to which the EM Program participants implemented the minimum SEM activities, as defined by the Consortium for Energy Efficiency (CEE). Though BPA designed the program before CEE defined the minimum elements of SEM, the CEE definition is a useful standard for comparing SEM programs with different implementation strategies and objectives. Comparing SEM programs to a common standard can reveal whether SEM activities are related to achieved savings or savings persistence.

The evaluation team conducted phone interviews with facility managers to help the EPT team assess which SEM elements were less frequently implemented and why. The EPT team can weigh the importance of those activities and determine whether they should be emphasized in the future.

5.1 SEM Adoption Methodology

We assessed the SEM adoption level at each evaluated facility by designing and administering a survey based on the CEE definition of minimum SEM elements:

- **Customer commitment** consists of developing and communicating energy goals, establishing an energy team, and having regular team meetings.
- **Planning and implementation** is measured by the use of energy maps, energy management assessments, employee engagement, and reassessment of goals and regular updates to the opportunity register or tune-up action item list.
- **Systems for measuring and reporting** energy performance criteria, including energy measurement and tracking techniques, updates with the SEM advisor, and frequent communication of progress to others.

Appendix L provides the survey guide. The evaluation team assigned a *full SEM adoption* score to participants who implemented all of the CEE's minimum SEM activities, and a *some SEM adoption* score to participants who implemented some activities. Appendix M provides the detailed methodology we used for scoring SEM adoption from the participant survey responses.

The evaluation team worked with BPA and the utilities to improve the likelihood that facilities would participate in the survey. The utilities contacted their customers participating in BPA's program to inform them of the study and to ensure they remained receptive to the request. The team then received permission from the utilities to contact all but one facility.

Twenty-four of 31 facilities responded to the survey, as shown in Table 10. We attempted to reach the other seven facilities at least six times each.

Table 10. Survey Response Disposition

Status	HPEM 1	HPEM 2	T&T	Total
Population	14	11	7	**32**
Available to Call	14	10	7	**31**
Completed Survey	**12**	**6**	**6**	**24**
Refused (utility or site)	0	0	0	**0**
Did not reach (answering machine, no answer, not available)	2	4	1	**7**

5.2 SEM Adoption Findings

The evaluation team surveyed 24 of 32 HPEM and T&T participants to assess their adoption levels of different SEM elements, based on CEE's definition of the minimum SEM elements. The team analyzed survey question responses to determine which SEM elements were adopted. Figure 26 and Figure 27 show the results for each element.

Figure 26. Percentage of Respondents with Full Adoption of SEM Elements

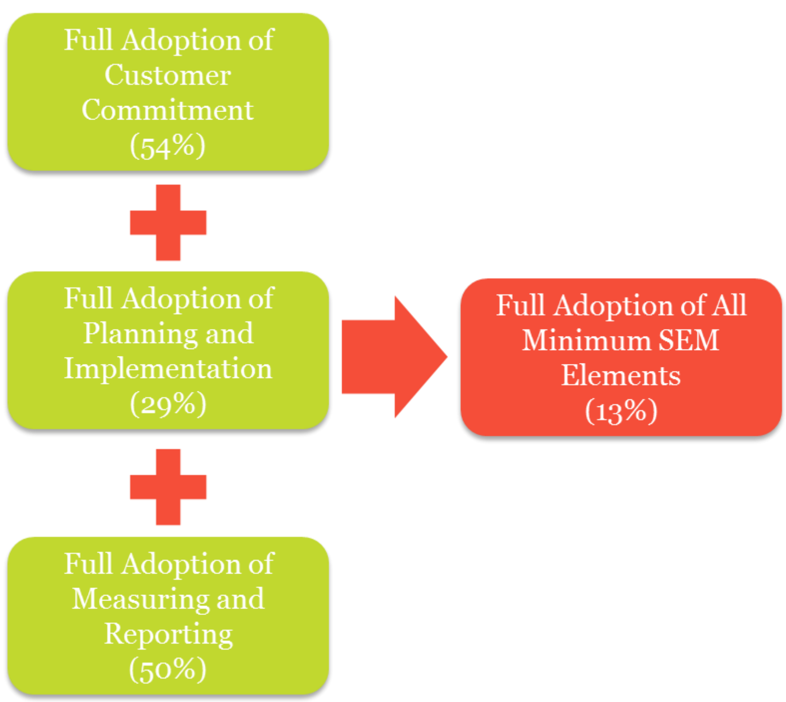
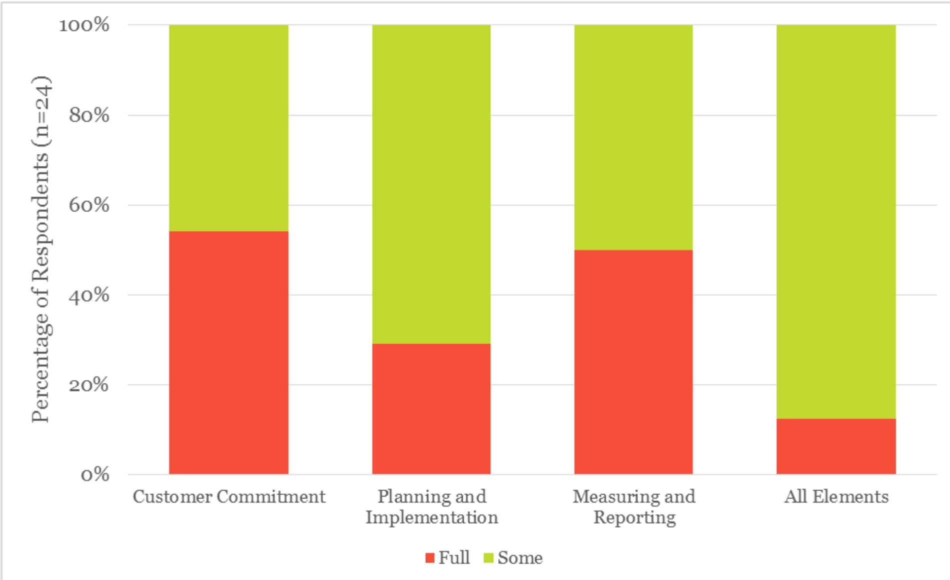


Figure 27. Overall SEM Adoption Level Results



Overall, three of 24 (13%) facilities implemented all of the minimum SEM elements, and all other facilities implemented some SEM aspects. Thirteen facilities (54%) met the customer commitment criteria, seven (29%) met the planning and implementation criteria, and 12 (50%) met the system for measuring and reporting energy performance criteria. The sections below detail results for each category.

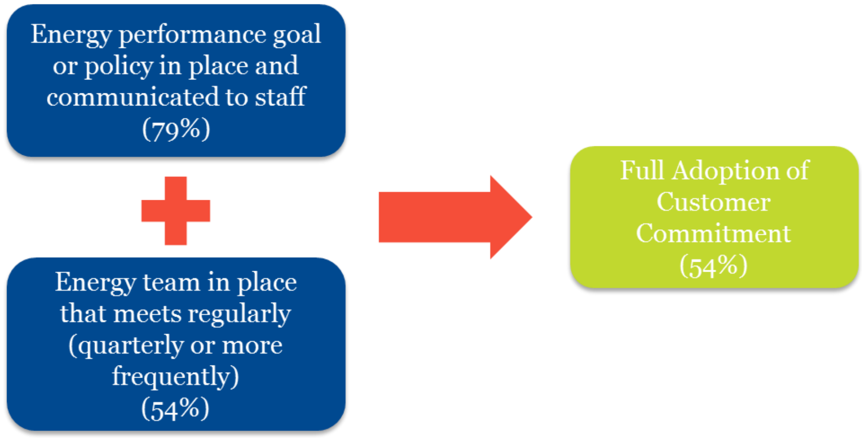
5.2.1 Customer Commitment

Customer commitment consists of meeting the following criteria:

- Employ an energy performance goal or policy and communicate this to staff
- Employ an energy team that meets regularly (quarterly or more frequently)

Figure 28 shows the percentage of respondents that met each criteria.

Figure 28. Customer Commitment Criteria Results



Thirteen of the 24 (54%) surveyed facilities met all customer commitment criteria. Eleven facilities did not meet one or more areas, including two that did not have an energy performance goal or policy in place and three with a goal or policy that was not communicated to staff. Four facilities did not have an energy team. Of those that did, four met less often than quarterly, and three met as needed but did not provide a frequency, so the evaluation team could not determine if they met this criterion.

5.2.2 Planning and Implementation

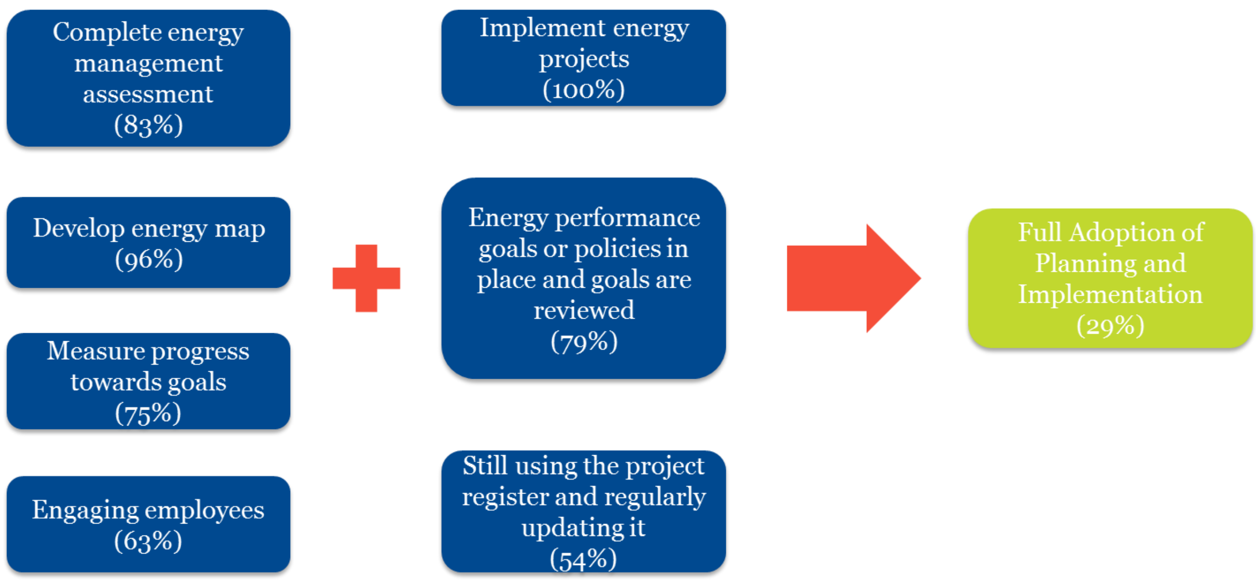
Planning and implementation consists of meeting the following criteria:

- Complete an energy management assessment
- Develop an energy map
- Establish metrics and goals, and measure progress towards goals
- Develop and use a project register
- Engage employees
- Implement energy projects
- Review goals to ensure they align with business and energy performance priorities, and regularly update the project register

Facilities met many of these criteria through engagement with HPEM or T&T, as shown in Figure 29. As part of HPEM, participants conducted an energy management assessment and developed an energy map, though this was not part of T&T. The evaluation team asked HPEM participants to confirm that these activities had been completed.

As all HPEM and T&T participants had an energy model, the team assessed whether participants used the model to measure progress towards their goals. HPEM and T&T participants developed a project register (HPEM called this an opportunity register; T&T called it a tune-up action item list), so we asked whether they still used this register. All HPEM and T&T participants implemented energy projects, as documented in the completion reports that BPA provided to the team for the evaluation.

Figure 29. Planning and Implementation Criteria Results



Seven of 24 (29%) surveyed facilities met all of these criteria. All facilities implemented energy projects. Facilities most commonly did not engage employees (nine of 24) and did not use (six of 24) or update (six of 19) the project register.

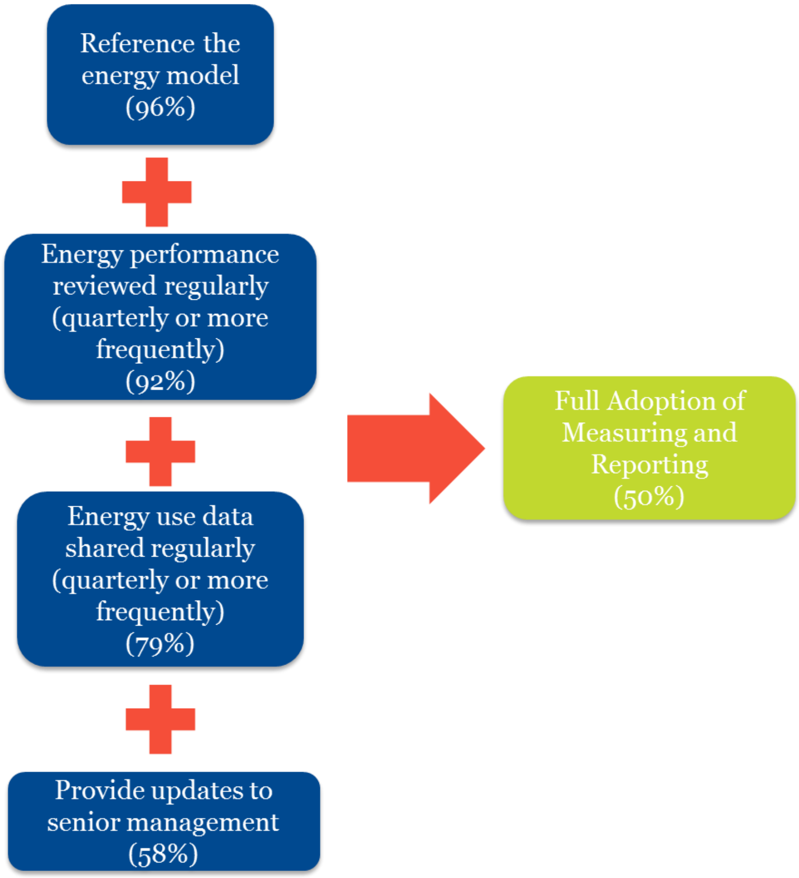
5.2.3 System for Measuring and Reporting Energy Performance

The criteria for measuring and reporting energy performance were as follows:

- Reference the energy model developed through HPEM or T&T to track energy performance at least quarterly
- Regularly provide updates to senior management
- Share energy consumption data with others in the organization (at least quarterly)

Figure 30 shows the percentage of respondents that met each criteria.

Figure 30. Measurement and Reporting Criteria Results



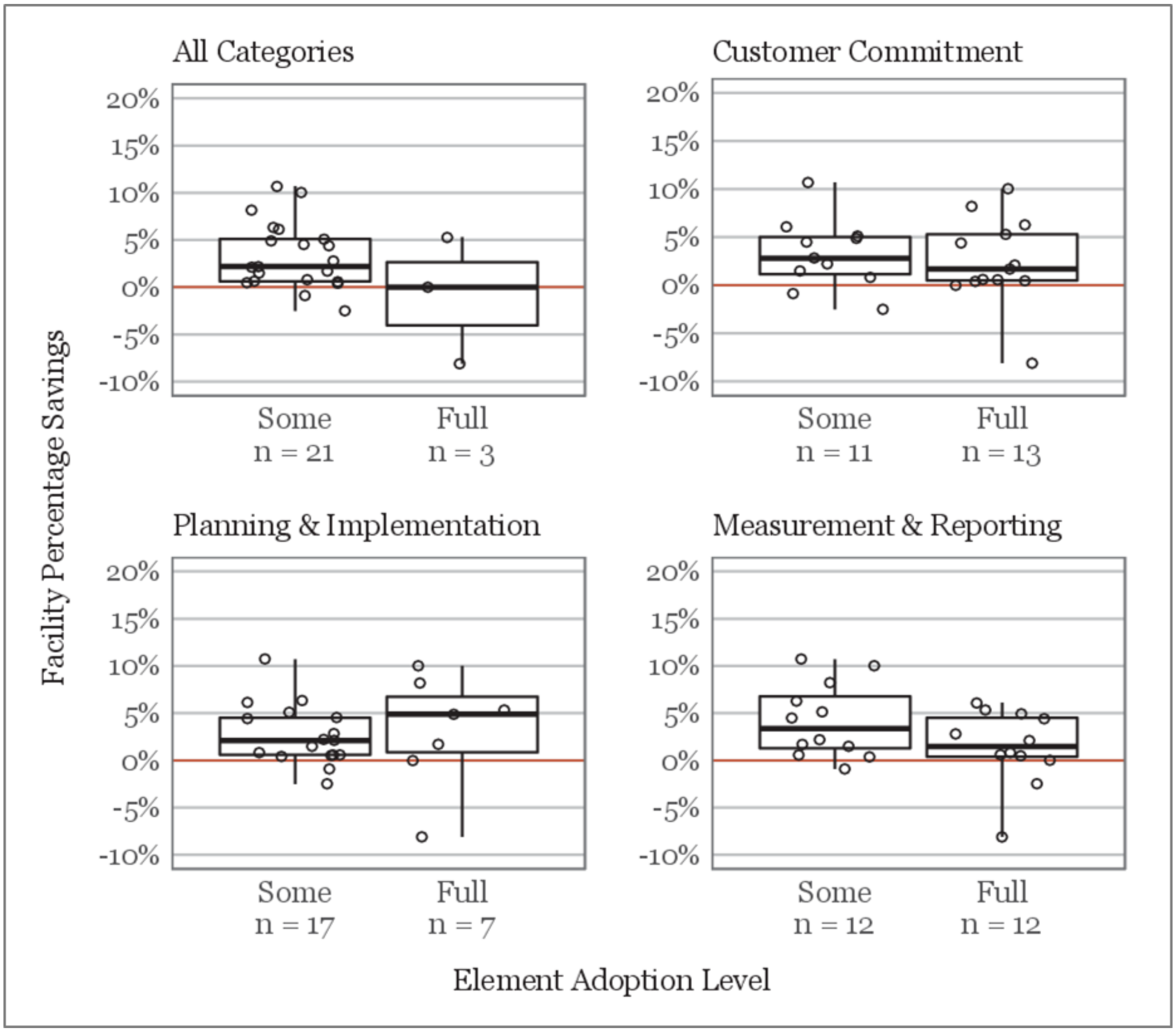
Twelve of the 24 (50%) surveyed facilities met all of the measurement and reporting criteria. Twelve facilities did not meet the criteria in one or more areas, including one facility that did not reference the energy model, and two facilities that reviewed energy performance data less often than quarterly. Ten facilities did not provide updates to senior management, and five did not share energy consumption data within their organization quarterly or more frequently; though all five reported sharing data, two shared data twice a year and three shared data annually.

5.2.4 SEM Adoption Correlation with Energy Savings

We reviewed whether facilities with higher SEM adoption also showed larger facility energy savings, or whether the adoption of certain elements could explain larger facility energy savings.

Figure 31 shows the adoption level overall and for each minimum element on the x-axis versus the evaluated facility energy savings on the y-axis. The box plot shows the quartiles, with the median represented by the middle band within the box. The points represent individual facility evaluated SEM savings results. Appendix N shows similar box plots for each sub-element.

Figure 31. Adoption Level of SEM Elements and Evaluated Facility Percentage Savings



No pattern emerged between the evaluated percentage facility energy savings and the overall SEM adoption level, nor between evaluated facility savings and adoption of the customer commitment, planning, and implementation, or adoption of a system for measuring and reporting energy savings. In fact, one facility with an estimated consumption increase (i.e., a negative savings estimate) had full adoption of some SEM elements.

These results, based on 24 respondents, do not indicate that full adoption of the minimum SEM elements is correlated with the amount of energy savings achieved across facilities. This null finding may indicate measurement issues rather than that the adoption of SEM elements does not correlate with higher savings. First, savings may depend on how well these activities are carried out and the energy saving measures that are implemented as a result. Second, there may be significant heterogeneity in facility savings that are not explained by implementing the SEM elements.

For example, percentage savings may vary by the types of industrial process used by participants. Third, the savings measurements contain uncertainty.

Program administrators can use SEM adoption data to provide valuable feedback to participants and to track a facility's progress with implementing SEM. In addition, tracking SEM adoption and savings annually throughout each facility's engagement may allow evaluators to better correlate the adoption of SEM minimum elements with energy savings.

6 Overall Findings and Recommendations

Before presenting recommendations, we summarize the most important evaluation findings.

6.1 Overall Findings

The important findings from the team's evaluation of the EM Program were as follows:

Finding 1. The EPT team carefully documented the program implementation and collected the data required for evaluation. Overall, the EPT team's EM Program data collection and documentation can serve as an industry standard for SEM programs. The EPT team's ongoing communication with participants through several program years resulted in the collection of high-quality data for the evaluation. The evaluation team was able to estimate savings for most facilities because the EPT team had thoroughly documented the program's implementation. For each facility and year, the EPT team prepared a project completion report, which described the facility operations and energy consumption, documented implemented SEM activities, and provided an estimate of the SEM energy savings. In addition, the EPT team collected energy consumption data and production data required for evaluating participating facilities.

Finding 2. SEM saved 2.3% of facility electricity consumption. The evaluation team estimated that, across all years, sampled EM Program facilities saved 4.1% of electricity consumption from the combination of SEM and capital projects, for an annual average savings of 3.8 average megawatts (aMW). Capital project savings equaled 1.8% of electricity consumption.[39] SEM savings equaled 2.3% of electricity consumption, or an average of 2.1 aMW per year.

Finding 3. SEM savings varied by Energy Management Program component. Sampled T&T facilities saved the most energy as a percentage of consumption, with total facility savings of 7.1% and SEM savings of 6.8% (average of 1.1 aMW). Sampled HPEM participants achieved facility savings of 3.7% and SEM savings of 1.6% (an average of 1.3 aMW).

Finding 4. SEM savings persisted during the participation period. The evaluation team tracked the energy savings of sampled HPEM facilities that participated for three or four years. Facility savings increased throughout the participation period and SEM savings (dashed lines) persisted after the first year and increased slightly in the last year. This persistence of savings

39 Capital project savings were not evaluated in this study. The evaluation team obtained these savings from original M&V estimates, contained in the MT&R reports.

suggests that facilities continued to practice energy management activities throughout the engagement.

Finding 5. Individual facility savings were variable. There was significant variation in savings between facilities and from year-to-year for individual facilities. The percentage savings coefficient of variation (the ratio of the sample standard deviation to the sample mean) was 201%. This variation in annual savings likely reflected differences in SEM implementation, changes in electricity consumption, and uncertainty of the savings estimates.

Finding 6. Some facilities had estimated consumption increases. In the majority (78%) of facility program years, evaluated SEM savings estimates were positive. However, in 22% of facility program years, the SEM savings estimate was negative. This includes 10% of cases where both facility and SEM savings were negative, as well as 12% of cases when the facility savings estimate was positive but the SEM savings estimate was negative after subtracting capital project savings.

Estimated increases in consumption likely reflected difficulties in the measurement of savings because of omitted variables, degradation in capital equipment performance, or unaccounted for non-programmatic effects—not that the program caused consumption to increase. However, an increase in facility consumption (e.g., because of a program implementation error) cannot be ruled out. As there is no accepted method for differentiating between omitted variables and a program causal effect, the evaluation results included estimated consumption increases.

Finding 7. The adoption of SEM elements was not correlated with SEM percentage savings. The Consortium for Energy Efficiency identified 13 management practices, called "elements," for facilities to continuously improve their energy performance. The evaluation team surveyed 24 EM Program participants in both program components to assess their adoption of these elements. We analyzed whether facilities that implemented a larger number of SEM elements or that adopted specific elements saved more energy. The results in Appendix N show no pattern of specific SEM elements. This may be due to the small sample size, unexplained variation in percentage savings between facilities, or because savings depended on factors outside this survey (such as how well participants implemented the SEM practices).

Finding 8. The evaluation team verified the MT&R SEM savings estimates. The evaluation team's estimate of SEM savings (2.3% of consumption) was slightly higher than the EPT team's MT&R SEM savings estimate (2.2% of consumption). The MT&R SEM savings realization rate—the ratio of evaluated to MT&R savings—was 1.06.[40] The MT&R realization rates

40 The realization rate was the ratio of evaluation savings to either the MT&R or reported savings. Realization rates greater than 1.0 indicate that the evaluation team calculated more savings than the EPT team estimated or BPA reported.

were 1.05 for T&T and 1.08 for HPEM. The MT&R and evaluation savings estimates for individual facilities were also similar: in 73% of facility-years, the evaluated savings and the MT&R savings estimates were not statistically different.[41] The evaluation savings estimate was statistically different and higher than the MT&R estimate in 12% of facility-years and statistically different and lower than the MT&R in 15% of facility-years.[42]

Finding 9. The evaluation team estimated lower SEM savings than BPA reported due to BPA's reporting practices. BPA reported program SEM energy savings of 2.7% (average of 2.4 aMW per year). The evaluation team estimated savings of 2.3% (average of 2.1 aMW per year), or 12% less. The reported SEM savings realization rate was 0.88. The reported savings realization rates were 1.05 for T&T and 0.79 for HPEM.

The evaluated savings were less than the reported savings because of BPA's practice of reporting zero savings for facilities with negative savings estimates. BPA reasoned that an increase in facility electrical consumption was not likely to have been caused by SEM implementation. Also, because incentives are based on savings, this convention mitigates a change in payment policies.

However, this reporting convention treats negative and positive savings estimates inconsistently. Positive savings estimates were just as likely to exhibit error as negative savings estimates, and the sign of the savings estimate should not be the reason for accepting or rejecting it. Reporting zero savings for negative facility savings biases the estimates of overall program savings upwards. Appendix K discusses the issue of negative SEM savings estimates.

Finding 10. More research about estimating SEM savings is needed. This evaluation led to new insights about the reliability of different SEM savings estimation methods, estimation of SEM savings uncertainty, causes of negative savings estimates, and ways of controlling for significant, non-programmatic changes in facility operations and energy consumption (non-routine adjustments). Nevertheless, more research is needed in each of these areas.

6.2 Key Recommendations for EM Program M&V

The evaluation team makes the following recommendations for performing measurement and verification of the EM savings:

- The EPT team should continue using statistical analysis of facility consumption to estimate savings. Specifically, the team should employ the forecast savings estimation approach on a site-specific basis. This

41 The savings estimates were not statistically different when the 80% confidence interval around the evaluated facility savings included the MT&R facility savings.

42 Facility-year savings were savings for a facility during a participation year.

approach is widely accepted, familiar to program participants, and expected to produce accurate savings estimates.

- The EPT team should continue documenting non-routine adjustments to support model specification or re-baselining and to inform evaluation efforts.
- The EPT team should continue to collect high-frequency consumption data when possible, rather than monthly billing data, since facilities with higher frequency energy model data (i.e., daily or weekly rather than monthly) had a smaller regression coefficient of variation.
- The EPT team should continue to report energy consumption increases in the MT&R model workbooks and to document the application of any non-routine adjustments.
- The EPT team should have discretion about whether to calculate and report uncertainty of the MT&R facility savings estimates. Estimation of savings uncertainty might provide some value to the program team, but it is not essential for M&V.
- The EPT team should routinely test for the statistical significance of weather variables in the MT&R energy consumption regression model and include these variables in the model if they are significant.
- BPA should attempt to improve the accuracy of the reported SEM savings by recording negative SEM savings estimates or making program-level adjustments to savings.
- The EPT team should review and, if necessary, update guidelines for when it is appropriate to choose a new consumption baseline for a facility. Section 5.0 of BPA's ESI MT&R Reference Guide provides guidance about re-baselining.

If BPA wants to conduct additional research into specific topics, we recommend the following:

- To improve the accuracy of SEM savings estimates in the long run at facilities with custom capital projects, BPA could investigate how the persistence of capital project savings can impact the accuracy of SEM savings estimates.
- To understand whether participation in an SEM program increases the number of capital projects implemented, BPA could compare the number of implemented capital projects in participant and non-participant facilities. BPA could also investigate whether SEM program participation impacts the persistence of capital project savings.
- To support an assessment of program cost-effectiveness, BPA should collect data on participant facilities' costs of implementing SEM and savings from other fuels.

- To study the persistence of savings after a facility finishes its engagement with the program, BPA should continue to collect data from participant facilities after engagement ends. Collection of such data would help BPA to better assess the program's long-term value and cost-effectiveness.

6.3 SEM Adoption Recommendations

The evaluation team did not find a relationship between the number of SEM activities adopted and the magnitude of facility energy savings. However, promoting these activities may lead to greater persistence of energy management practices and to sustained energy savings after participants graduate from the program, though this has yet to be demonstrated. We have the following recommendations for BPA to consider:

- To further understand the relationship between savings and adoption of specific SEM elements, BPA could conduct the energy management assessment annually to update participants' progress in implementing SEM.
- The EPT team should encourage energy teams to schedule regular meetings, at least quarterly. Twenty (of 24) facilities reported using an energy team, but seven of those teams did not meet regularly.
- The EPT team should encourage energy teams to develop methods to engage other employees in efforts to improve energy performance. Nine (of 24) facilities reported not conducting employee engagement activities.
- The EPT team should encourage energy managers or teams to regularly update senior management. All facilities reported sharing energy consumption data within their company, but 10 facilities reported that senior management did not require regular updates. The energy team should review these data at least annually with senior management to highlight accomplishments, so senior management continues to recognize the value of those efforts.

6.4 Recommendations for Future Evaluations

In summary, the evaluation team offers the following recommendations to BPA for conducting future evaluations:

- In general, evaluators can choose from a number of different statistical regression methods to estimate savings. These methods, which are reviewed in Appendix B, are expected to produce accurate savings estimates. However, in selecting a method, evaluators should consider the potential benefits of aligning their approach with that used by the program.

- In situations when there was a significant, non-programmatic change to facility operations and energy consumption, one estimation method may produce a more accurate savings estimate than another. Evaluators should consider the relative merits of different savings estimation approaches in these circumstances.
- Evaluators should consider employing automated variable selection methods in building baseline regression models. These methods provide an objective and cost-efficient way of identifying relevant independent variables, as well as higher-order terms of and interactions between relevant variables.

Although this evaluation has broken new ground in many areas, there are still several topics that BPA or other national evaluators of SEM programs could explore further:

- Evaluate the energy savings of the newest EM projects, which were not considered in this evaluation. Such an evaluation would show whether the newest participants achieved savings similar to that of the facilities included in this evaluation.
- Assess the effect of BPA's new policy of establishing a new baseline for participant facilities every two years on savings realization rates.
- Conduct a process evaluation to understand why HPEM cohorts performed differently and to gain insights about the relationship between savings and implementation of specific SEM activities.
- Estimate the persistence of SEM savings after a facility's engagement with the program ends in order to evaluate program cost-effectiveness.
- Investigate the feasibility and reliability of evaluating savings for a sample of SEM participants instead of the population.
- Study how uncertainty of capital project savings estimates affects SEM savings estimates.

APPENDICES

A. Sampling Simulation Study Results Summary

The evaluation team performed a small simulation study using the reported and verified electricity savings observed from 15 projects in the previous program evaluation.[43] We defined the certainty stratum to include projects that contributed to the top 65% of annualized reported savings, resulting in a sample size of six projects in the certainty stratum. From the remaining nine projects, we randomly sampled five in the sample stratum to reach a target sample size of 11. The team performed the random sampling procedure 10 times and calculated the resulting realization rates, verified total savings, and precision at the study level. The results are provided in Table 11. The simulation results can be compared to the results from verifying a census of the HPEM 1 cohort sites, which gave a realization rate of 94% with a total verified savings of 9.9 MWh.

Table 11. Simulation of HPEM 1 Cohort EM Program Evaluation Results

HPEM 1 Cohort	Realization Rate	Study Level Estimated Verified Savings (kWh)	Relative Precision	Does Confidence Interval Contain Census Realization Rate Result?
Simulation 1	91%	9,650,107	±13%	Yes
Simulation 2	83%	8,820,294	±14%	Yes
Simulation 3	90%	9,483,861	±12%	Yes
Simulation 4	86%	9,070,269	±12%	Yes
Simulation 5	100%	10,567,872	±1%	No
Simulation 6	86%	9,115,166	±13%	Yes
Simulation 7	88%	9,356,357	±12%	Yes
Simulation 8	86%	9,129,230	±14%	Yes
Simulation 9	86%	9,116,314	±13%	Yes
Simulation 10	108%	11,402,619	±3%	No

The resulting realization rates and total verified savings estimates vary widely, with realization rates between 83% and 108%, and savings results between 8.8 MWh and 11.4 MWh. Eight of the 10 simulations yielded a confidence interval range that included the realization rate result from the census analysis.

---

43 Cadmus. "Energy Management Pilot Impact Evaluation." Prepared for Bonneville Power Administration. February 1, 2013. Available online: http://www.bpa.gov/EE/Utility/research-archive/Documents/BPA_Energy_Management_Impact_Evaluation_Final_Report_with_Cover.pdf

The relative precision ranges between 1% and 14%; however, with such a wide range of estimated savings, this tight precision may create the false impression that there is little uncertainty about the true savings. In fact, there is tight precision around a result that is potentially very far off from the true savings.

B. Overview of Savings Estimation Methods

This section provides an overview of the forecast, backcast, and pre-post methods for calculating facility savings is given below. Each of these methods is described, followed by a discussion of the main differences between the approaches and their advantages and disadvantages. The EPT Team used the forecast method exclusively, while the evaluation team used the forecast method as the default but employed the pre-post method in select cases. Appendix D describes the evaluation team's logic for selecting which of these methods to use.

Forecast Method

The forecast approach is prescribed in the forthcoming DOE's Uniform Methods Project protocols and the SEP M&V protocol and adheres to IPMVP Option C. This method analyzes individual facility energy consumption and compares metered energy consumption with adjusted baseline energy consumption, which is an estimate of what facility energy consumption would have been if the facility had not implemented efficiency measures. This method is illustrated in Figure 32.

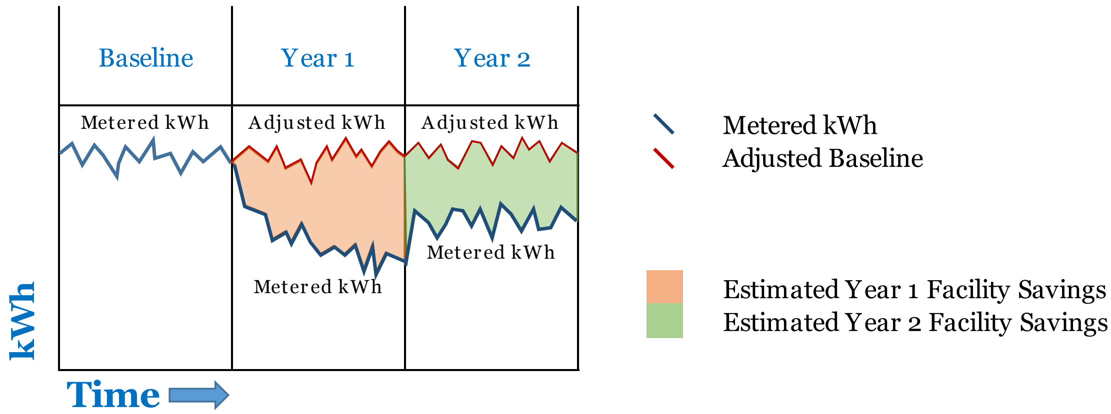
First, a regression model is estimated using baseline period data. Then the regression model is used to predict reference energy consumption (shown by the baseline model predicted kWh). Specifically, for each time interval during the reporting period, the estimated model coefficients and independent variables measured during the reporting period are used to estimate what energy consumption would have been if SEM had not been implemented (i.e., if facility output would have remained as in the reporting period but baseline period operating conditions had persisted during the performance period). Finally, facility energy savings are then calculated as the difference between the adjusted baseline and metered energy consumption. Facility savings during the reporting period equal the area between the adjusted baseline and metered energy consumption.

Similar to most national programs, the EPT team's MT&R models uses the forecast approach to estimate facility savings.[44]

The evaluation team tested the forecast approach during the exploratory analysis case studies, which are discussed further in Appendix I.

44 SEM program implementers commonly use the forecast approach to calculate facility savings.

Figure 32. Forecast Approach



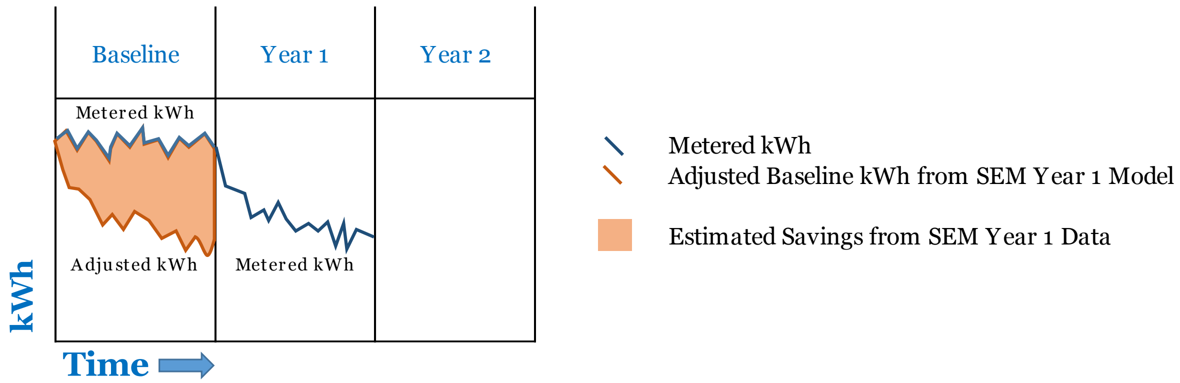
Backcast Method

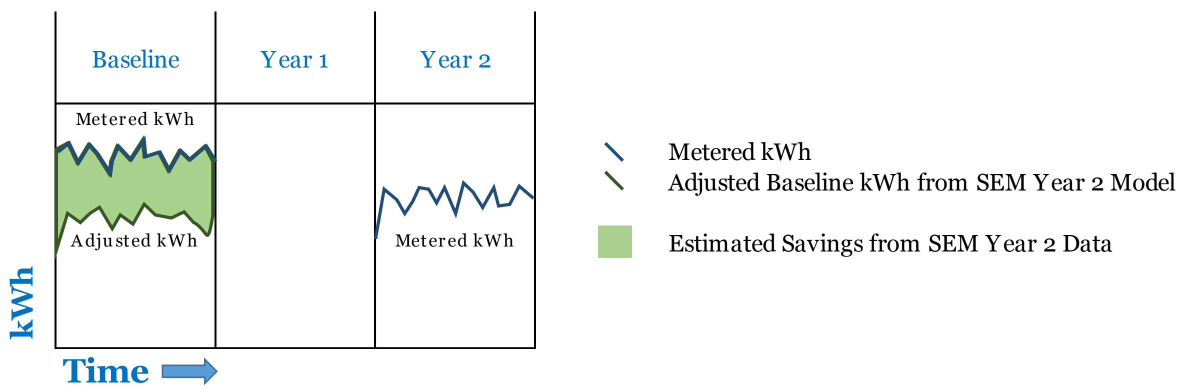
The backcast method is outlined in DOE's SEP M&V protocol and is similar to the forecast method in the way it compares metered energy consumption to an adjusted baseline. The primary difference is that using the backcast modeling technique, this concept is applied retrospectively. While the forecast method uses metered energy usage during the baseline period to predict reporting period energy consumption under baseline operating conditions, the backcast method uses reporting period meter data to predict baseline period energy consumption under baseline period conditions assuming efficiency measures were in place. Figure 33 illustrates the approach.

For this method, the evaluator first estimates energy consumption regression model using reporting period data. Next, they use the regression model specification to predict what energy consumption would have been during the baseline period if SEM had been implemented at that time. Then the evaluator estimates savings as the difference between metered energy consumption and the backcasted adjusted baseline. Separate regression models need to be built for each program year. In Figure 33, savings equal the area between metered energy consumption and the adjusted baseline for each year. The backcast method produces a savings estimate for the baseline period, that is, the facility energy savings that would have occurred if SEM had been implemented during the baseline period. This is different than the forecast method measures and therefore the backcast and forecast energy savings estimates may differ.

The backcast savings can be expressed relative to baseline period metered energy consumption to estimate SEM savings as a percentage of consumption. Then the evaluator can apply the percentage savings to the performance period energy consumption to estimate performance period SEM savings.

Figure 33. Backcast Approach





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.

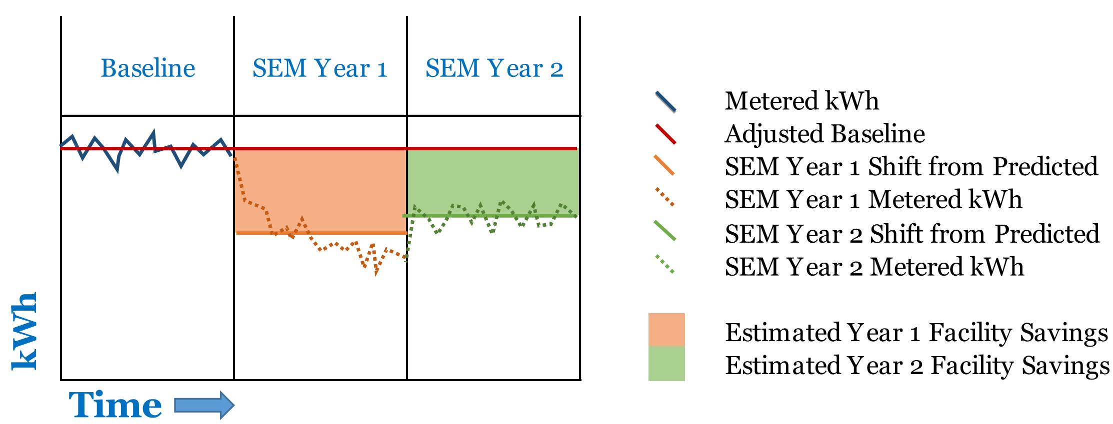
The evaluation team tested the during the exploratory analysis case studies, which are discussed further in Appendix I.

Pre-Post Method

The pre-post method is used extensively in program evaluation and was applied to industrial SEM evaluation by Luneski (2011), who directly estimated the facility average energy savings per time interval using a regression of baseline period and performance period energy consumption.[45] The coefficient on an indicator variable for SEM activity in the model represents the average facility savings per time interval. The SEM variable can stand alone in the model or be interacted with other model independent variables, such as output or weather. If the indicator variable stands alone, the model implies that SEM had a level shift effect on energy consumption, as shown in Figure 34. If the SEM activity indicator variable interacts with other variables, the model implies that SEM savings depend on the other variables.

Also, the evaluator can include SEM indicator variables for periods of less than one year to measure savings over a shorter period, such as one month or three months; For example, by adding separate SEM indicator variables for each month, the evaluator can estimate any ramping of savings during the first program year.

Figure 34. Pre-Post Approach



Evaluation Selection of Methods

As stated earlier, the EPT team used the forecast method to estimate facility energy savings. The evaluation originally intended to use the pre-post method because it was expected to produce accurate savings estimates and would have simplified the uncertainty calculations. In the end, however, the evaluation team decided to use the forecast method as a default to align with the EPT team's approach. The evaluation team made this decision after

45 For applications of pre-post model to program evaluation, see: Imbens, Guido W. and J. M. Wooldridge. "Recent Developments in the Econometrics of Program Evaluation." Journal of Economic Literature (2009): 47. pp. 5-86.

extensive testing of the accuracy of both approaches. These tests are presented in Appendix I and Appendix J.

The evaluation team 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.[46] The team developed decision logic to determine when to apply the pre-post method, which we present and discuss in Appendix D.

[46] When the evaluation team employed the pre-post method, we checked the sensitivity of the pre-post model savings estimates by estimating the model with and without interaction variables between the SEM activity indicator and other variables (such as output and weather). The team obtained very similar savings estimates both with and without these interaction variables, suggesting that the savings estimates were not sensitive to being modeled as a level shift or as a function of output.

C. Evaluation Methodology To Estimate Energy Savings: Additional Details

This appendix provides further details of the evaluation team's method for estimating savings including selecting the baseline model (evaluation method step 2).

Step 1: Define the Baseline and Program Periods

The evaluation team reviewed the MT&R model definitions of the baseline period and program period and adopted the same definitions for nearly all of the evaluation models. For some facilities, the EPT team determined that a redefinition of the baseline was required due to changes at the facility unrelated to the EM Program. The EPT team documented these facility changes in the annual completion reports and noted when they decided the changes warranted re-baselining.

The evaluation team used different baseline period definitions for four facilities based on a careful review of these data and documentation. We first attempted to account for changes to facility energy consumption and operations unrelated to SEM by including new variables in the regression model. If that effort was unsuccessful, we specified a new baseline and estimated a regression model using data from the new baseline. We attempted to select the baseline period to be free of program implementation activities and to have conditions that were otherwise representative of those during the reporting period.

Step 2: Build the Facility Baseline Consumption Model

For each facility, the team estimated several regression model specifications with different functional relationships between energy consumption and different independent variables. We used the EPT team's MT&R model as a starting point for building the evaluation regression model. These prior modeling efforts significantly reduced the evaluation team's time to build an energy consumption model and improved the quality of the final model.

The model selection process involved applying both engineering knowledge about a facility's energy consumption and automated variable selection methods. The specific steps were:

- **Step 2a.** Identify a candidate set of explanatory variables including output, weather, and facility closures and production shutdowns. The evaluation team identified candidate variables for the baseline energy consumption model based on an engineering description of the facility in the annual participant report and MT&R data. We collected weather data from the NOAA weather station closest to the facility.

- **Step 2b.** Identify significant drivers of facility energy consumption using stepwise selection procedures that consider the candidate variables above, as well as interactions between and higher order terms of these variables. An automated variable selection process can identify variables that affect facility energy consumption that cannot be identified through engineering analysis. We carefully reviewed the model specification selected through the automated procedure, then added or removed variables as necessary based on our knowledge of the site type and the site production. In most cases, the model we selected was very similar to the model selected by the EPT team.
- **Step 2c.** Select the final baseline model. In selecting the best model, the team followed a step-by-step process of diagnostics analysis, variable selection, and model selection, with specific attention given to the signs and statistical significance of the estimated parameters, the joint significance of the parameters, prediction accuracy, and model comparison tools such as AIC (Akaike's information criterion), BIC (Bayes' information criterion), and R^2 (coefficient of determination).

The final model selected to estimate a facility's savings took the following general form:

$$kWh_t = \alpha + f(output_t, \beta) + g(other_t, \gamma) + \varepsilon_t$$

with model variables defined as follows:

kWh_t	=	Electricity consumption at facility during the "t" time interval (could be a day, week, or month)
α	=	Intercept indicating the average facility base load energy consumption per interval
output_t	=	The vector of different outputs produced at the facility during the "t" time interval; the model might contain several different outputs, with linear or nonlinear relationships to electricity consumption
β	=	The coefficient vector that defines the relationship between outputs and energy usage, defined as the average energy usage per unit of output
$other_t$	=	The vector of additional explanatory variables and/or indicators related to electricity consumption at the facility during the "t" time interval; this may contain weather variables, indicators of facility shutdowns or closures, or indicators for changes in input quality
γ	=	The coefficient vector that defines the relationship between the additional explanatory variables (other than output) and electricity consumption, defined as the average electricity consumption per unit

ε_t = The model error term representing unobservable influences on electricity consumption during time interval "t"

Step 3: Calculate Adjusted Baseline Energy Consumption

The evaluation team estimated the facility regression model using either ordinary least squares (OLS) or, if tests revealed that energy consumption was autocorrelated, we used the Yule-Walker (feasible generalized least squares) estimator. In the presence of autocorrelation, OLS estimation yields unbiased and consistent coefficient estimates, but the coefficient standard errors and inferences based on the standard errors would be incorrect.

Then, for each interval of the reporting period, the evaluation team used the estimated Equation 1 to calculate the adjusted baseline:

$$\widehat{e_t} = \hat{\alpha} + f\left(output_t, \hat{\beta}\right) + g(other_t, \hat{\gamma})$$

where:

$\widehat{e_t}$ = The adjusted baseline energy consumption for time interval "t" of the reporting period

ˆ = Denotes an estimate of a coefficient

The other variables are defined as above.

We evaluated the adjusted baseline using the reporting period values of output, outside temperature, and other variables.

Step 4: Estimate Facility Savings

Facility energy savings were from all energy efficiency projects and behavior changes undertaken by the facility during the reporting period. Facility savings included savings from changes in operations, maintenance procedures, and employee behaviors, as well as from capital projects. Some capital projects may have received funding from other efficiency programs, and therefore their savings were subtracted from the facility savings (Step 5) so as to not double-count these savings across two programs.

The evaluation team estimated each facility's energy savings s during interval "t" of the reporting period as:

$$s_t = \widehat{e_t} - e_t$$

Facility energy savings during the reporting period equaled the sum of savings over the intervals of the reporting period:

$$\mathrm{S} = \sum_t \widehat{s_t}$$

The evaluation team calculated standard errors and 80% confidence intervals for the annual facility savings. The team estimated the standard error of the estimated savings as follows:

$$\text{standard error(S)} = \sqrt{Var(\sum_{t=1}^{T^P} \hat{\alpha} + f(output_t, \hat{\beta}) + g(other_t, \hat{\gamma})) + T^P \widehat{\sigma^2}}$$

Where $\widehat{\sigma^2}$ is the regression standard error (i.e., the estimate of the error variance σ^2 from the baseline period regression model). The first term in the formula is the variance of the adjusted baseline consumption. The second term in the standard error formula, $T^P \widehat{\sigma^2}$, is an estimate of the metered energy-use variance during the reporting period. This may be estimated using the regression standard error (i.e., the regression root mean square error) of the baseline regression, assuming the error variance during the baseline and reporting periods is equal. The methodology for calculating the facility savings confidence intervals is described in Appendix E.

It was not possible to calculate confidence intervals for the SEM savings because estimates of uncertainty for capital measure savings were unavailable. The team based the capital measure savings estimates on engineering algorithms, and quantifying the associated uncertainty would have been difficult and was beyond the scope of this evaluation.

Step 5: Estimate SEM Savings

The evaluation team estimated SEM savings for each facility by subtracting any savings from capital project incentivized through other BPA or utility programs (S_K) during the reporting period from S:

$$\text{SEM Savings} = S - S_K$$

The team obtained estimates of the facility's annual capital project savings from the facility's annual project completion report. When a capital project occurred midway through a year, we prorated the annual savings.

D. Logic Flow for Applying Non-Routine Adjustments

This appendix presents the logic that the evaluation team applied in making non-routine adjustment to facility consumption. A non-routine adjustment is an adjustment to metered energy consumption that accounts for a non-programmatic change at the facility.

The factors we used to determine how to make the non-routine adjustment for a particular facility and program year were:

- Existence of a change to facility energy consumption requiring a non-routine adjustment.
- Whether an engineering estimate of the change in facility energy consumption was available.
- Whether the change requiring the non-routine adjustment occurred during the baseline or reporting period.
- Whether it was possible to estimate the separate impacts of the non-routine adjustment and SEM savings using the pre-post regression model.
- Whether the non-routine adjustment was small relative to the expected SEM savings.

Figure 35 shows the logic flow chart. The chart has three main paths for making a non-routine adjustment defined by whether a non-routine adjustment was necessary and the availability of an engineering estimate, each described in more detail below:

- A non-routine adjustment was not necessary
- A non-routine adjustment was necessary and an engineering estimate was available
- A non-routine adjustment was necessary and an engineering estimate is not available

Path A: A non-routine adjustment was not necessary
When a non-routine adjustment was not required, the evaluation team employed the default forecast method.

Path B: A non-routine adjustment was necessary and an engineering estimate was available
When a non-routine adjustment was required and an engineering estimate was available, our strategy for making the non-routine adjustment depended on whether the non-programmatic change in the facility's consumption occurred during the baseline or performance period. If the change occurred during the baseline period, the evaluation team applied the non-routine adjustment to the baseline period consumption and employed

the forecast method. The evaluation team determined that the forecast method could account for the uncertainty of the non-routine adjustment when the non-routine adjustment was applied to the baseline period consumption.

If the change occurred during the performance period, our strategy for making the non-routine adjustment depended on whether the pre-post method could have been used to make the adjustment. The evaluation team chose the pre-post method if the pre-post method would have yielded separate estimates of the non-routine adjustment and SEM savings impacts. In general, the pre-post method would yield estimates of separate impacts if high frequency data (weekly or daily) were available, and the non-routine adjustment and the reporting period did not coincide too closely. The evaluation team would use the pre-post method even if an engineering estimate of the non-routine adjustment was available. The evaluation team preferred this path over making the non-routine adjustment to the performance period data and applying the forecast model because the pre-post method was expected to produce valid estimate of both savings and savings uncertainty.

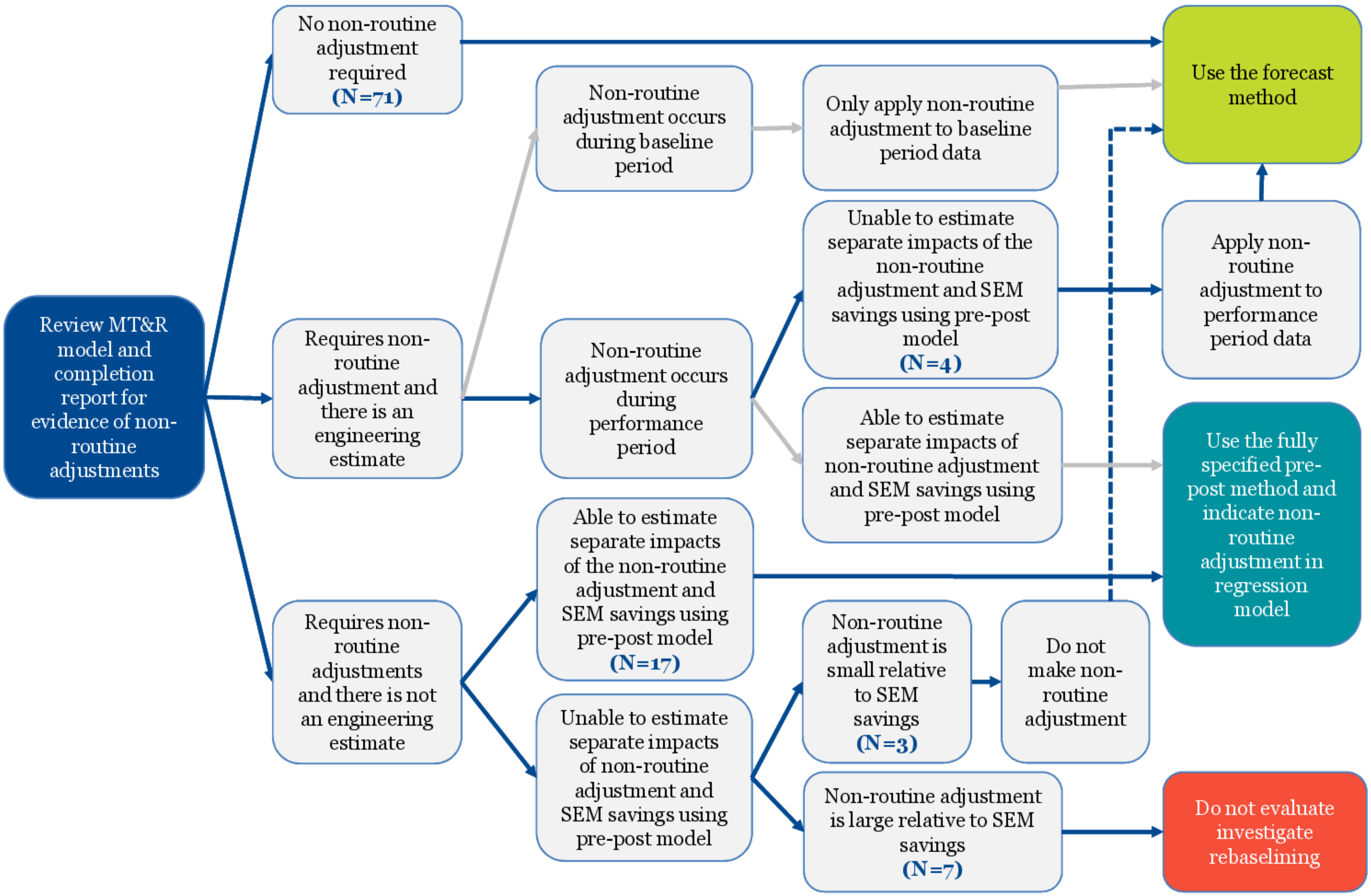
The final pathway corresponded to a situation in which an engineering estimate for the non-routine adjustment was available, but it was not possible to use the pre-post method to estimate the separate impacts of the change and the SEM savings. In this case, the evaluation team made the non-routine adjustment using the engineering estimate and applied the forecast method. This approach produced an accurate savings estimate (assuming the engineering estimate was accurate), but it would not produce a valid estimate of the SEM savings uncertainty unless an estimate of the uncertainty of the non-routine adjustment was available.

Path C: A non-routine adjustment was necessary and an engineering estimate was not available

When an engineering estimate for the non-routine adjustment was not available, our strategy for making the non-routine adjustment depended on whether it was possible to apply the pre-post model. The evaluation team applied the pre-post method when the pre-post model would yield valid estimates of the impacts of the non-routine adjustment and the SEM savings. Again, the pre-post model would produce a valid estimate if high frequency data (weekly or daily) were available, and the non-routine adjustment and the reporting period did not coincide too closely.

When the evaluation team could not apply the pre-post method, our strategy depended on whether the non-routine adjustment was small enough relative to the SEM savings to ignore. If the non-routine adjustment was relatively small, the team did not account for the non-routine adjustment and applied the forecast method. But if the non-routine adjustment was large relative to SEM savings, the evaluation team could not obtain a valid savings estimate by ignoring the non-routine adjustment. In this case, we concluded that it was not possible to evaluate the facility savings and investigated the possibility of redefining the baseline period to include the change in facility consumption necessitating the non-routine adjustment.

Figure 35. Evaluation Flow Chart



E. Uncertainty Calculation Methodology for Forecast Model Savings Estimates

This appendix describes the recommended approach for estimating the standard errors of forecast model savings estimates. The evaluation team requires an approach for calculating the standard errors to compare the precision of pre-post model and forecast model savings estimates. In addition, BPA requested technical guidance about the calculation of forecast savings standard errors for future program evaluations.

The analytic formula that the evaluation team recommends captures two sources of uncertainty: the variance of the adjusted baseline consumption and the variance of metered energy use during the performance period. It is necessary to account for both components to obtain an accurate estimate of the forecast model savings standard error.

The first section of this appendix presents a framework for deriving the standard error of the forecast model savings estimates. It presents a simple model of facility energy use and defines SEM savings. The second section proves that under the assumptions of the classical linear regression model, the pre-post method and the forecast method are both expected to yield unbiased SEM savings estimates. The third section of the appendix presents the formulas for the pre-post model savings and forecast model savings standard errors. The fourth section recommends that the standard error of the forecast model savings be estimated using this formula.

Definition of SEM Savings

Consider an SEM program facility. The period preceding the start of participation is the baseline and the period following is the performance period. Suppose the following regression model describes facility electricity use per interval kWh_t in the baseline period:

$$kWh_t = \alpha + \beta x_t + \varepsilon_t \qquad \text{(Equation 1)}$$

where x_t is an explanatory variable such as output for interval t and α and β are coefficients to be estimated. α can be interpreted as baseload energy use per interval and β can be interpreted as the energy use per unit of output. The error term εt is normally, independently, and identically distributed with conditional mean zero and variance σ^2.

During the SEM performance period, the facility implements changes to improve efficiency of baseload energy use and energy use per unit of output. After implementation, facility electricity use per interval of the SEM performance period (P) is given by:

$$kWh_t = \alpha^P + \beta^P x_t + \varepsilon_t^P \qquad \text{(Equation 2)}$$

where P denotes the performance period, kWh_t and x_t are energy use and output for interval t and α^P and β^P are coefficients to be estimated. α^P is baseload energy use per interval and β^P is the energy use per unit of output after implementation of SEM. The error term ε_t^P is normally, independently, and identically distributed with mean zero and variance σ^2_P. The variance of ε_t and ε_t^P may differ.

For interval t of the performance period with facility output x_t^P, SEM energy savings s_t equals the difference between expected energy use conditional on x_t^P under baseline conditions and expected energy use conditional on x_t^P under performance period conditions:

$$s_t = E[kWh_t | x_t^P, \alpha, \beta] - E[kWh_t | x_t^P, \alpha^P, \beta^P]$$

$$= \alpha + \beta x_t^P - \alpha^P - \beta^P x_t^P$$

$$= (\alpha - \alpha^P) + (\beta - \beta^P) * x_t$$

where E is the expectation operator and | denotes "conditional on." In the last equation, the first term is the baseline energy savings per interval and the second term is the energy savings per unit of output multiplied by the amount of output in interval t. Note that by assumption ε_t^P represents random influences on facility energy use during the performance period and therefore does not enter the saving definition. As we show below, defining savings as a difference in conditional expected energy use can help to explain surprising results such as when estimated savings are negative.

$$S = (\alpha - \alpha^P)*T^P + (\beta - \beta^P) * \sum_{t=1}^{T^P} x_t$$

$$= \alpha^\Delta * T^P + \beta^\Delta * \sum_{t=1}^{T^P} x_t$$

where:

$$\alpha^\Delta = \alpha - \alpha^P; \text{ and}$$

$$\beta^\Delta = \beta - \beta^P$$

Savings Estimation Approaches

We can estimate the performance period energy savings S using either the pre-post method or the forecast method. This section shows that the pre-post and forecast methods both yield unbiased estimates of S.

Pre-Post Method

In the first approach, we nest both Equation 1 and Equation 2 in a single model and estimate the resulting pre-post model:

kWh_t = baseline energy use - savings + error

$= \alpha - \alpha^{\Delta}*Post_t + \beta x_t - \beta^{\Delta} x_t*Post_t + \varepsilon_t + (\varepsilon_t^{P} - \varepsilon_t)*Post_t$ (Equation 3)

where

$Post_t$ = 1 for intervals during the performance period and = 0, otherwise;

Note that if Post=0 the model reduces to Equation 1, and if Post=1, the model reduces to Equation 2. When the models are nested in a single model, the model includes a full set of interactions between Post and all variables affecting energy use during the baseline period.

The model is estimated by OLS and we obtain an estimate of performance period savings $\widehat{S_t}$:

$$\widehat{S} = T^{p} * a^{\Delta} + b^{\Delta} * \sum_{t=1}^{T^P} x_t$$

where a^{Δ} is the OLS estimate of α^{Δ} and b^{Δ} is the unbiased estimate of β^{Δ}. Under the assumptions of the classical linear regression model, OLS will yield an unbiased estimates of α, α^{Δ}, β, and β^{Δ}, and therefore $\widehat{S}$ is an unbiased estimate of S.

Forecast Method

A second approach for estimating savings is the forecast method. Using data from t=1, 2, ..., T periods during the baseline period, the researcher estimates Equation 1 by OLS and obtains estimates of α, β, and error variance σ^2, denoted a, b, and $\widehat{\sigma^2}$.[47]

Next, the researcher uses the model $\widehat{kWh_t} = a + b\,x_t$ to predict expected energy use in the performance period (P) under the assumption that SEM had not been implemented. For each of the t=1, 2, ... , T^P intervals during the performance period, we observe both kWh_t^P and x_t^P.

Energy savings in interval t of the performance period are estimated as:

$$\widehat{s_t} = \widehat{kWh_t^P} - kWh_t^P$$

$$= a + bx_t^P - kWh_t^P$$

$$= a + b\,x_t^P - \alpha^P - \beta^P x_t^P - \varepsilon_t^P$$

47 Let e_t be the residual of the regression in period t. $\widehat{\sigma^2}$ is estimated as the sum of squared residuals divided by T-k, that is, $\Sigma_{t=1}^{T} e_t^2/(T\text{-}k)$, where k is the number of coefficients to be estimated in the regression.

where $\widehat{kWh}_t^P$ is an estimate of the expected energy use under baseline conditions (the adjusted baseline) during the performance period and kWh_t^P is metered energy use during the performance period. Note that in accordance with Equation 2, kWh_t^P can be expressed as the sum of the expected value of kWh_t^P conditional on x_t^P plus an error, that is, $kWh_t^P = E[kWh_t | x_t^P, \alpha^P, \beta^P] + \varepsilon_t^P$. We will use this fact below in calculating the variance of forecast savings.

Performance period savings equals:

$$\hat{S} = \sum_{t=1}^{T^P} \hat{s}_t$$

$$= \sum_{t=1}^{T^P} a + b\, x_t^P - \alpha^P - \beta^P x_t^P - \varepsilon_t^P \qquad \text{Equation (3)}$$

Taking expectations (E[ ]) of both sides,

$$E[\hat{S}] = E[a + b\, x_t^P - \alpha^P - \beta^P x_t^P - \varepsilon_t^P]$$

$$= (\alpha - \alpha^P)*T^P + (\beta - \beta^P) * \sum_{t=1}^{T^P} x_t^P$$

$$= \alpha^\Delta * T^P + \beta^\Delta * \sum_{t=1}^{T^P} x_t^P$$

$$= S$$

The second equality follows because under the assumptions of Equation 1, OLS yields an unbiased estimate of the model parameters, $E[a] = \alpha$ and $E[b] = \beta$. Therefore, $\hat{S}$ is an unbiased estimate of the pilot savings, and both the forecast method and the pre-post method are expected to provide unbiased estimates of S.

The forecast method is the same method that IPMVP Option C (2012) and ASHRAE Guideline 14 (2014) recommend for conducting whole facility savings estimation.

Estimation of Savings Uncertainty

This section presents formulas for estimating the uncertainty of the pre-post model and forecast model savings estimates.

Standard Errors of Forecast Method Savings

Next, we derive the formula for the standard error of savings during interval t of the performance period.

$$\mathrm{Var}(\hat{s}_t) = \mathrm{var}(\widehat{kWh}_t^P - kWh_t^P)$$

$$= \mathrm{var}\,(a + b\, x_t^P - \alpha^P - \beta^P x_t^P - \varepsilon_t^P)$$

$$= \mathrm{Var}\,(a + b\, x_t^P) + \mathrm{Var}(\varepsilon_t^P)$$

$$= \widehat{\sigma^2} x_t^{P\prime}(x'x)^{-1} x_t^P + \sigma_P^2$$

where $\mathbf{x_t}^P$ is a 2 x 1 vector with first element equal to 1 and the second element equal to x_t^P . (Note the 2 columns of $\mathbf{x_t}^P$ correspond to the 2 parameters of Equation 1 (α and β)). X is a T x 2 matrix with ones in the first column and the values of x_t in the second column for the t=1, 2, ...T intervals of the baseline period.

The third equality follows because α^P and β^P are unknown but fixed parameters and the error ε_t^P is independent. Note that the variance of the savings estimate for interval t depends on both $\mathbf{x_t}^{P\prime}(X'X)^{-1}\mathbf{x_t}^P$, the variance of the *adjusted baseline* conditional on x_t^P, and the variance of energy use during the performance period $\hat{\sigma}_p^2$. It is necessary to account for the variance of performance period energy use because this energy use depends on random factors not affected by SEM.[48] The standard error is obtained by taking the square root of the variance.

To calculate the variance of the performance period savings estimate $\widehat{S}$, we take the variance of both sides of Equation 3.

$$\mathrm{Var}(\widehat{S}) = Var\left(\sum_{t=1}^{T^P} a + b\, x_t^P - \alpha^P - \beta^P x_t^P - \varepsilon_t^P\right)$$

$$= \mathrm{Var}\left(\sum_{t=1}^{T^P} a + b\, x_t^P - \varepsilon_t^P\right)$$

$$= \mathrm{Var}\left(\sum_{t=1}^{T^P} a + b\, x_t^P\right) + \mathrm{Var}\left(\sum_{t=1}^{T^P} \varepsilon_t^P\right)$$

$$= \widehat{\sigma^2}\, \mathbf{x}^{Psum\prime}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}^{Psum} + T^P\, \hat{\sigma}_p^2 \qquad \text{(Equation 4)}$$

where $\mathbf{x}^{Psum}$ is a 2 x 1 vector with first element equal to T^P and the second element equal to $\sum_{t=1}^{T^P} x_t^P$.

In Equation 4, if we make the simplifying assumption that the variance of the errors in the baseline and performance periods are equal, i.e., $\hat{\sigma}_p^2 = \widehat{\sigma^2}$, then the variance of the performance period savings equals:

$$\mathrm{Var}(\widehat{S}) = \widehat{\sigma^2}\, \mathbf{x}^{Psum\prime}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{Psum} + T^P\, \widehat{\sigma^2}$$

$$= \widehat{\sigma^2}\left(\mathbf{x}^{Psum\prime}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}^{Psum} + T^P\right) \qquad \text{(Equation 5)}$$

Forecast Model Savings Uncertainty with Autoregressive Errors

Auto-correlated errors arise when random, unobservable factors affecting facility energy use in one interval affect energy use in future intervals. For example, autocorrelation may occur in a facility that has an inventory of non-energy production inputs (e.g., timber) and that uses the highest quality inputs first. Assuming production requires less energy when

[48] This follows from the definition of savings presented above. According to the definition, savings are the difference in expected energy use conditional on x_t^P. This implies that $\widehat{kWh_t^P}$ should be interpreted as the expected value of kWh conditional on x_t^P under baseline conditions, i.e., E[kWh_t| x_t^P, α, β]; and kWh_t^P should be interpreted as the expected value kWh conditional on x_t^P under SEM conditions *plus an error*. When taking the variance of kWh_t^P , it is necessary to account for the variance of ε_t^P.

processing high quality inputs, energy use per unit of output will increase with time since the last restocking of inventory. Unless this facility's practice of using the highest quality inputs first is explicitly modeled, the model error term will exhibit autocorrelation.

With auto-correlated errors, savings is estimated as the difference between the adjusted baseline and metered energy use, just as with a forecast model that satisfies the classical linear regression model assumptions. However, when estimating forecast model savings with auto-correlated errors, it is necessary to account for the autocorrelation in calculating the adjusted baseline and in estimating the standard error of the savings.

Suppose that the error term of facility energy use during the baseline period follows an autoregressive (AR) process of order 1:[49]

$$kWh_t = \alpha + \beta x_t + \varepsilon_t \qquad \text{(Equation 6)}$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t$$

The error term, ε_t, is a function of the error in period t-1, ε_{t-1}, and an independent and identically distributed disturbance for period t, μ_t. The coefficient ρ is the autocorrelation coefficient and determines the extent to which disturbances in an interval carry over to the next. If $\rho=0$, the AR model reduces to classical OLS model.

The adjusted baseline $\widehat{kWh}$ for period t of the reporting period estimated with a forecast model with auto-correlated errors is given by:[50]

$$\widehat{kWh}_t^{AR} = a_{GLS} + b_{GLS}x_t + \hat{\rho}\,(kWh_{t-1} - a_{GLS} - b_{GLS}x_{\boldsymbol{t-1}})$$

(Equation 7)

where:

a^{GLS} and b^{GLS}	=	estimates of α and β from two-stage Generalized Least Squares (GLS) estimation of Equation 3.[51]
$\hat{\rho}$	=	the estimate of the autocorrelation coefficient ρ.

The coefficients a^{GLS}, b^{GLS}, and $\hat{\rho}$ may be obtained from the two-stage GLS or maximum likelihood estimation of Equation 6.[52] In Equation 7, it is evident that through $\hat{\rho}$, random disturbances to energy use in interval t-1 (estimated

49 The calculation of savings and estimation of the standard errors would proceed analogously for a forecast model with a higher order autoregressive error process. More details about higher order AR processes can be found in Johnston and DiNardo (1997) or other standard econometrics texts.

50 See Johnston and DiNardo (1997). Econometric Methods, p. 192.

51 The two-stage GLS coefficient estimates are sometimes referred to as Yule-Walker estimates. Instead of GLS estimation, the estimates may also be obtained through full-information maximum likelihood estimation.

52 In the first stage of the GLS estimation, the model is estimated by OLS and an estimate of ☐ is obtained. In the second stage, the data are transformed using $\hat{\rho}$, and then the model is estimated by OLS using the transformed data.

as kWh_{t-1} – a^{GLS} – $b^{GLS}x_{t-1}$) are carried forward into interval t and future intervals.

Savings for interval t are estimated as:

$$s_t^{AR} = \widehat{kWh}_t^{AR} - kWh_t$$

It is not possible to estimate the variance of s_t^{AR} analytically, because there is no closed-form (analytic) expression for $\widehat{kWh}_t$ (Johnston and DiNardo, 1997, p. 193). However, under the assumption that the autocorrelation coefficient ρ is known and not estimated, it is possible to approximate the variance of savings for interval t as:

$$var(s_t^{AR}) = \sigma_{AR}^2 \left(1 + x_{t,*}^{P\prime}(X_{*,}^{\prime}X_{*,})^{-1} x_{t,*}^{P}\right)$$

where:

X_*	=	the T^P x 2 matrix of transformed ones and x_t's. The first column is the transformed vector of ones (1- $\hat{\rho}$) and the second column is the transformed vector x $_{t,}$ $x_{t,*} = x_t - \hat{\rho}\,(x_{t-1})$.[53]
x_{t*}^P	=	the 2 x 1 vector of the transformed one and x_t for interval t of the reporting period. The first column is the transformed constant (1- $\hat{\rho}$) and the second column is the transformed scalar $x_{t,}$ $x_{t,*} = x_t - \hat{\rho}\,(x_{t-1})$.
σ_{AR}^2	=	the mean squared error of the GLS regression model (i.e., the regression standard error) and is estimated as:

$$\widehat{\sigma_{AR}^2} = \frac{\sum_{t=1}^{T}(y_{t,*} - a_{GLS}(1-\hat{\rho}) - b_{GLS}x_{t,*})^2}{T - 2}$$

where:

y_{t*}	=	the transformed energy use for interval t of the baseline period equal to $y_t - \hat{\rho}\,(y_{t-1})$.
x_{t*}	=	the transformed energy use for interval t of the baseline period equal to $x_t - \hat{\rho}\,(x_{t-1})$.

Following the same steps as for the regression model that satisfies the classical assumptions, the variance of performance-period savings may be approximated as:

$$var(S^{AR}) = \widehat{\sigma_{AR}^2}\left(T^P + x_{t,*}^{Psum\prime}(X_{*,}^{\prime}X_{*,})^{-1} x_{t,*}^{Psum}\right) \quad \text{(Equation 8)}$$

where:

53 This assumes that the first observation in the data set is dropped. See Johnston and DiNardo (1997, p. 190) for matrix expression if the first observation is retained.

x_{t*}^{PSum} = the 2 x 1 vector of the sum of the transformed ones and xt's for intervals of the reporting period.

Summary

This memo demonstrates that forecast savings estimate has two sources of uncertainty: the first is the variance of the adjusted baseline and the second is the variance of metered energy use. Both components should be accounted for to obtain an accurate estimate of the variance of the savings estimate.

In addition to providing a more accurate estimate of the variance, accounting for the error of metered energy use can help to explain unexpected results such as a negative savings estimates. For example, suppose that a facility experiences a random disturbance during the performance period that causes the facility's energy use to increase significantly and the estimated savings to become negative. Since this disturbance was large, it is important the standard error reflect the magnitude of the disturbance; otherwise, the standard error may be under-estimated, the savings estimate may be reported as statistically significant when it was not, and the evaluator may wrongly conclude that the program caused consumption to increase. Accounting for the error of metered energy use reduces the likelihood that the evaluator will find savings when there were none and can explain why savings were negative.

Recommendation

The evaluation team recommends estimating the variance of the forecast savings estimate using Equation 5, which assumes $var(\varepsilon_t) = var(\varepsilon_t^P)$. We do not recommend that evaluators estimate Equation 4 to obtain an estimate of $var(\varepsilon_t^P)$, because the likely gain in accuracy will not be worth the additional modeling effort. In cases of auto-correlated errors, the evaluation team recommends estimating the variance using Equation 8.

F. MT&R, Reported, and Evaluated Savings by Year

Table 12. All Program Components MT&R, Reported, and Evaluated Savings by Year

All Program Components	Year 1 (MT&R n=32 Evaluation n=30)*				Year 2 (n=29)				Year 3 (n=24)				Year 4 (n=13)				Average Annual Program Savings			
	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar
MT&R GWh Savings	28.6	6.5	22.1	24.2	36.8	15.8	21.0	22.9	37.1	21.5	15.6	17.0	24.8	13.3	11.5	12.6	31.8	14.2	17.6	19.1
MT&R % Savings	2.6%	0.6%	2.0%	2.0%	3.4%	1.5%	1.9%	1.9%	4.8%	2.8%	2.0%	2.0%	10.3 %	5.5%	4.8%	4.8%	4.0%	1.8%	2.2%	2.2%
Reported GWh Savings	N/A	N/A	24.3	26.5	N/A	N/A	27.5	30.0	N/A	N/A	19.6	21.4	N/A	N/A	13.6	14.9	N/A	N/A	21.3	23.2
Reported % Savings	N/A	N/A	2.2%	2.2%	N/A	N/A	2.6%	2.6%	N/A	N/A	2.5%	2.5%	N/A	N/A	5.7%	5.7%	N/A	N/A	2.7%	2.7%
Evaluation GWh Savings	33.4	6.4	27.0	29.5	36.2	15.8	20.4	22.2	36.5	21.5	15.0	16.4	25.6	13.3	12.3	13.5	32.9	14.2	18.7	20.4
Evaluation % Savings	3.1%	0.6%	2.5%	2.5%	3.4%	1.5%	1.9%	1.9%	4.7%	2.8%	1.9%	1.9%	10.6 %	5.5%	5.1%	5.1%	4.1%	1.8%	2.3%	2.3%
80% Confidence Interval (GWh)	± 4.2	N/A	N/A	N/A	± 4.4	N/A	N/A	N/A	± 3.7	N/A	N/A	N/A	± 2.4	N/A	N/A	N/A	± 2.8	N/A	N/A	N/A
80% Confidence Interval (%)	± 0.4%	N/A	N/A	N/A	± 0.4%	N/A	N/A	N/A	± 0.5%	N/A	N/A	N/A	± 1.0%	N/A	N/A	N/A	± 0.4%	N/A	N/A	N/A
Realization Rate																				
Evaluation / MT&R	1.17	N/A	1.22	1.22	0.98	N/A	0.97	0.97	0.98	N/A	0.96	0.96	1.03	N/A	1.07	1.07	1.04	N/A	1.06	1.06
Evaluation / Reported	N/A	N/A	1.11	1.11	N/A	N/A	0.74	0.74	N/A	N/A	0.77	0.77	N/A	N/A	0.90	0.90	N/A	N/A	0.88	0.88

* The different values for n in Year 1 (the number models contributing to the year estimates) are a result of the evaluation team's determination that HPEM facilities 1-2 and 2-5 were not evaluable during this year.

Table 13. HPEM 1 and HPEM 2 Cohorts MT&R, Reported, and Evaluated Savings by Year

HPEM 1 and 2	Year 1 (MT&R n=25 Evaluation n=23)*				Year 2 (n=24)				Year 3 (n=23)				Year 4 (n=13)				Average Annual Savings			
	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar
MT&R GWh Savings	13.1	5.1	8.0	8.7	23.4	15.8	7.6	8.3	35.8	21.5	14.3	15.6	24.8	13.3	11.5	12.6	24.3	13.9	10.3	11.3
MT&R % Savings	1.6%	0.6%	1.0%	1.0%	2.6%	1.7%	0.8%	0.8%	4.7%	2.8%	1.9%	1.9%	10.3%	5.5%	4.8%	4.8%	3.5%	2.1%	1.5%	1.5%
Reported GWh Savings	N/A	N/A	10.2	11.1	N/A	N/A	14.1	15.4	N/A	N/A	18.3	19.9	N/A	N/A	13.6	14.9	N/A	N/A	14.1	15.3
Reported % Savings	N/A	N/A	1.2%	1.2%	N/A	N/A	1.5%	1.5%	N/A	N/A	2.4%	2.4%	N/A	N/A	5.7%	5.7%	N/A	N/A	2.0%	2.0%
Evaluation GWh Savings	16.5	5.1	11.4	12.4	23.1	15.8	7.3	8.0	35.0	21.5	13.5	14.8	25.6	13.3	12.3	13.5	25.0	13.9	11.1	12.1
Evaluation % Savings	2.0%	0.6%	1.4%	1.4%	2.5%	1.7%	0.8%	0.8%	4.6%	2.9%	1.8%	1.8%	10.6%	5.5%	5.1%	5.1%	3.7%	2.1%	1.6%	1.6%
80% Confidence Interval (GWh)	± 3.7	N/A	N/A	N/A	± 4.3	N/A	N/A	N/A	± 3.7	N/A	N/A	N/A	± 2.4	N/A	N/A	N/A	± 2.8	N/A	N/A	N/A
80% Confidence Interval (%)	± 0.4%	N/A	N/A	N/A	± 0.5%	N/A	N/A	N/A	± 0.5%	N/A	N/A	N/A	± 1.0%	N/A	N/A	N/A	± 0.5%	N/A	N/A	N/A
Realization Rate																				
Evaluation / MT&R	1.26	N/A	1.43	1.43	0.99	N/A	0.96	0.96	0.98	N/A	0.95	0.95	1.03	N/A	1.07	1.07	1.03	N/A	1.08	1.08
Evaluation / Reported	N/A	N/A	1.12	1.12	N/A	N/A	0.52	0.52	N/A	N/A	0.74	0.74	N/A	N/A	0.90	0.90	N/A	N/A	0.79	0.79

* The different values for n in Year 1 (the number models contributing to the year estimates) are a result of the evaluation team's determination that HPEM facilities 1-2 and 2-5 were not evaluable during this year.

Table 14. T&T MT&R, Reported, and Evaluated Savings by Year

T&T	Year 1 (n = 7)				Year 2 (n = 5)				Year 3 (n = 1)				Year 4 (n = 0)				Average Annual Savings			
	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar	Facility	Capital	SEM	SEM with Busbar
MT&R GWh Savings	15.5	1.3	14.2	15.5	13.4	0.0	13.4	14.6	1.3	0.0	1.3	1.4					10.1	0.4	9.6	10.5
MT&R % Savings	6.0%	0.5%	5.5%	5.5%	8.1%	0.0%	8.1%	8.1%	7.5%	0.0%	7.5%	7.5%					6.8%	0.3%	6.5%	6.5%
Reported GWh Savings	N/A	N/A	14.2	15.5	N/A	N/A	13.4	14.6	N/A	N/A	1.3	1.4					N/A	N/A	9.6	10.5
Reported % Savings	N/A	N/A	5.5%	5.5%	N/A	N/A	8.1%	8.1%	N/A	N/A	7.5%	7.5%					N/A	N/A	6.5%	6.5%
Evaluation GWh Savings	17.0	1.3	15.6	17.1	13.1	0.0	13.1	14.3	1.5	0.0	1.5	1.6					10.5	0.4	10.1	11.0
Evaluation % Savings	6.5%	0.5%	6.0%	6.0%	7.9%	0.0%	7.9%	7.9%	8.4%	0.0%	8.4%	8.4%					7.1%	0.3%	6.8%	6.8%
80% Confidence Interval (GWh)	± 2.1	N/A	N/A	N/A	± 0.7	N/A	N/A	N/A	± 0.5	N/A	N/A	N/A					± 0.5	N/A	N/A	N/A
80% Confidence Interval (%)	± 0.8%	N/A	N/A	N/A	± 0.4%	N/A	N/A	N/A	± 2.9%	N/A	N/A	N/A					± 0.4%	N/A	N/A	N/A
Realization Rate																				
Evaluation / MT&R	1.09	N/A	1.10	1.10	0.98	N/A	0.98	0.98	1.13	N/A	1.13	1.13					1.04	N/A	1.05	1.05
Evaluation / Reported	N/A	N/A	1.10	1.10	N/A	N/A	0.98	0.98	N/A	N/A	1.13	1.13					N/A	N/A	1.05	1.05

G. MT&R Savings Relative to Evaluation Savings

This appendix summarizes the number of occurrences where the MT&R facility savings were above, below, or within the evaluation facility savings 80% confidence interval.

Table 15. MT&R Savings Relative to Evaluation Savings

	Year 1	Year 2	Year 3	Year 4
MT&R savings above evaluation savings 80% CI	5	4	4	1
MT&R savings within evaluation savings 80% CI	19	22	19	10
MT&R savings below evaluation savings 80% CI	6	3	1	2
Total	**30**	**29**	**24**	**13**

H. Positive and Negative Evaluation Facility and SEM Savings

This appendix summarizes the instances where negative savings estimates occurred.

Table 16. Counts of Positive and Negative Evaluation Facility and SEM Savings Estimates by Program Year

	Year 1	Year 2	Year 3	Year 4
Positive Facility and SEM Savings	25	21	20	9
Negative Facility and SEM Savings	4	4	2	0
Positive Facility and Negative SEM Savings	1	4	2	4
Total	**30**	**29**	**24**	**13**

I. Exploratory Statistical Analysis Results

The evaluation team performed two types of exploratory analysis: (1) a model specification sensitivity analysis and (2) a comparison of model estimation methods. The first analysis took the form of a step-by-step walkthrough of incremental changes to the model specifications, demonstrating how each change affected the energy savings estimates. In the second exploratory analysis, we compared the accuracy and precision of savings estimates from the forecast and pre-post methods.

Model Specification Sensitivity Analysis

The evaluation team undertook a model specification sensitivity analysis to understand the factors affecting the evaluation savings estimates and differences between our estimates and those from the EPT team. Table 17 presents the objectives of the exploratory analysis and the activities taken to address each objective.

Table 17. Exploratory Analysis Primary Objectives

Objective	Activities
Identify differences between MT&R and evaluation facility savings estimation.	Performed a sensitivity analysis on regression models to identify the main drivers in savings differences.
Explore the advantages and disadvantages of alternative modeling approaches.	Computed facility savings estimates using forecast, backcast, and pre-post methods.
Identify potential improvements to modeling facility savings from SEM.	Discussed the results and implications with the EPT team and developed recommendations for future modeling.

The EPT team and evaluation team selected several facilities as candidates for the model case studies based on the following criteria:

- Evidence of unexplained variation in energy consumption
- Re-baselining for one or more performance years or major changes in production that occurred during the performance years
- Potential to explore the effects of adding weather variables to a model
- Negative or non-significant savings
- Possible unaccounted-for interactions between variables
- Potential for non-linear transformations of independent variables
- Low adjusted R^2 value

After investigating each candidate facility, the EPT team selected three facilities, denoted as Facility A, Facility B, and Facility C to protect the confidentiality of the participants for further exploratory analysis. The following factors influenced the facility selection:

- Replicability of the MT&R baseline model by the evaluation team;
- Similarity between the MT&R model and the evaluation model specifications;
- Consistency of yearly and total savings estimates between MT&R and evaluation results; and
- Statistical significance of differences in yearly and total savings estimates between the MT&R and evaluation models.

The evaluation team and EPT team agreed on a framework for conducting the analysis. For the model development process, the evaluation team followed the procedure outlined in 2.4: Savings Estimation, which consists of the following steps:

Replicate the MT&R baseline model documented in the project completion report.

Add performance-year indicator variables to the MT&R baseline model, and calculate annual savings using the resulting coefficients on the performance-year indicator variables.

Use baseline period data to select predictor variables for an initial evaluation model.

Use forward stepwise selection to select the model variables.

Test for autocorrelation.

Add performance-year indicator variables, then calculate annual savings estimates using the selected model.

If necessary, revise the evaluation model specification. Document the effect of each change to the model on the savings estimate.

Findings

Specific results for each of the three case study facilities are presented in Appendix J. This section presents the main findings of the sensitivity analysis.

Weather. Weather was an important determinant of energy consumption at industrial facilities, and should be accounted for when estimating savings. However, the specific functional form of weather or which weather variables to include in the models may not be important.

Interaction Variables and Automated Variable Selection. In two of the case studies, the evaluation team found that interaction variables significantly explained facility energy consumption, and that including these variables significantly improved the model fit. The evaluation team identified the interaction variables using automated variable selection methods. Evaluators should employ automated methods to identify important correlates of facility energy consumption.

Backcast Approach. Backcasting can be employed when program period conditions are inclusive of baseline conditions but not vice versa. However, the backcast approach yielded savings estimates that were not robust for two facilities. The evaluation team recommends conducting more research to determine whether backcasting is a valid approach.

Re-Baselining. Two of the three case study facilities experienced significant changes in production during the reporting period. It was not feasible to account for these changes using the baseline regression model. In cases such as these, the evaluation team recommends that evaluators establish a new baseline that incorporates the significant changes. Re-baselining is an appropriate method for capturing the impacts of the production changes on facility energy consumption.

Comparison of Forecast and Pre-Post Savings Estimation Methods

Either the forecast method or the pre-post method can be used to estimate facility savings, and both methods are expected to produce the same savings estimate. Appendix E, which shows the standard errors for savings estimated using the forecast method, proves this mathematically.

The evaluation team also compared the two methods in practice. We estimated facility energy savings using the forecast method and the pre-post method for facilities in the HPEM 1, HPEM 2, and T&T facilities. We compared the forecast method to two types of pre-post models:

Simple pre-post model: this model included a stand-alone SEM indicator variable for each program year (e.g., if there were four program years, the model would include four indicator variables, one for each program year). The coefficient on the SEM indicator for the j^{th} year would indicate the average savings per interval during the j^{th} year.

Fully-specified pre-post model: this model includes a stand-alone indicator variable for each program year, plus all statistically significant interaction variables between the program year indicators and independent variables in the regression model used to calculate the forecasted adjusted baseline. For example, if the forecast regression model included daily output and daily average temperature as explanatory variables, the fully-specified pre-post model would include any statistically significant interaction between these independent variables and the program year indicator variables.

For each of the facilities, we compared the point estimates of facility savings and the estimated standard error of savings from the forecast method and the two pre-post models.

Comparison Results

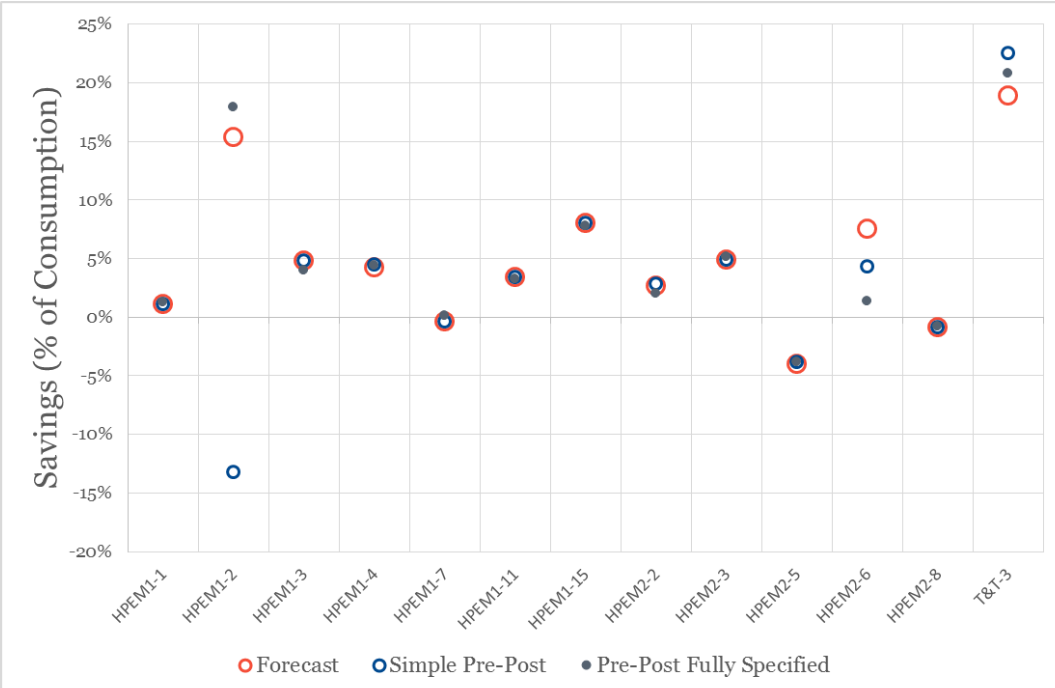
Figure 36 shows a comparison of savings estimates as a percentage of consumption for 13 facilities using the forecast method, simple pre-post method, and fully-specified pre-post method. For 10 of 13 facilities, the three approaches produced similar savings estimates, as expected. However, for three facilities, the models produced different estimates. The evaluation team reviewed the documentation and determined that the following situations led to the disparity between estimates:

An operational change occurred at the facility at the same time the reporting period began. The evaluation team attempted to add an indicator variable to the pre-post models to account for this change; however, the pre-post model was not able to estimate the impacts of both the SEM activity and the operational change. Year 1 for this site was determined to be non-evaluable.

The evaluation team found documentation that the facility was highly sensitive to changes in outside temperature. Temperatures in the reporting period were outside of the range of those experienced during the baseline period.

One of the meters that recorded facility consumption was inoperable for part of the reporting period. The evaluation team added an indicator variable to the pre-post models to account for this change; however, the forecast model did not reflect the change.

Figure 36. Model Specification Comparison



Significant differences in savings between the three facilities arose because the simple pre-post and the pre-post fully specified models included control variables to account for the operational changes while the forecast model could not.

J. Case Study Analysis Results

This appendix has results from three facility case studies that compared savings estimation methods and performed sensitivity analysis to determine if changes in model specification affected the savings estimates.

Case Study 1

Overview of Facility 1

The first facility selected for an in-depth analysis had three primary production outputs and a secondary output. The evaluation team selected this facility in order to investigate a disparity between MT&R and evaluated savings, assess the impact of a major change in facility equipment, and explore the impact of adding weather variables. The facility data included weekly energy consumption and production data, covering the baseline period of October 30, 2010, through October 29, 2011, and the three following SEM engagement years beginning on October 30, 2011, and ending on October 25, 2014. The facility had incentivized outdoor lighting and compressor upgrade projects installed in SEM years 1 and 2. Additionally, in SEM year 3, a facility equipment rebuild was started. This resulted in an increase in energy usage at the facility due to the use of less efficient replacements.

Exploratory Approach for Facility 1

The evaluation team tested a total of seven models of facility energy consumption:

- A pre-post version of the MT&R model
- Four evaluation pre-post models
- A forecast model
- A backcast model

We determined model specifications based on the following considerations:

Weather Variables. A description of the facility production process led the evaluation team to hypothesize that weather was an important explanatory variable. We added HDD, CDD, and mean weekly temperature to the stepwise selection process. The model fit was optimal with the combination of both HDD and CDD variables.

Equipment Rebuild. The completion reports indicated that an equipment rebuild was started in SEM year 3, necessitating that the facility temporarily use less efficient replacement equipment. In accordance with the evaluation team's defined methodology, we created and added an indicator variable to the evaluation model for the date range of the rebuild. The significance of this term may suggest that the less efficient temporary equipment had an effect on energy consumption at the facility.

Variable Interactions. In addition to including an indicator for the equipment rebuild period, the evaluation team tested for an interaction between mean weekly temperature and the rebuild indicator, hypothesizing that the equipment operation was weather dependent. This represents the variability in energy consumption relating to temperature changes that occur specifically during the rebuild period.

Other Explorations. In addition to the above, the evaluation team tested various combinations of HDD/CDD for sensitivity, examined the model residuals for serial correlation, and specified the forecast and backcast models for comparison to the pre-post models.

Findings of Exploratory Analysis for Facility 1

Replication of MT&R Model Results. The evaluation team conducted a regression analysis using the MT&R data and was able to replicate the coefficients in the MT&R model specification. The team then applied the DOE SEP Measurement and Verification Protocol forecast model methodology, and was able to replicate the MT&R savings results for years 2 and 3; however, the team was not able to replicate the MT&R savings results for year 1, instead calculating more than double the MT&R savings.

Evaluation Model Specification. The evaluation team selected a final evaluation pre-post model consisting of the set of variables selected for the MT&R model along with HDD, CDD, equipment rebuild indicator, and the equipment rebuild and mean weekly temperature interaction. Though the latter two variables were non-significant, we opted to retain them based on information about the equipment rebuild in the completion reports and improvements in the model selection criteria.

Sensitivity Analysis. As hypothesized, including weather variables led to an improvement in model fit, regardless of which combination of weather variables was used. However, savings estimates varied widely depending on which weather variables were included in the model. While a model including CDD provided the highest savings estimate, it was among the worst performing according to model selection criteria. This suggested that savings estimates for this facility were highly dependent on how weather was modeled. The evaluation team selected a model with both HDD and CDD, as indicated by the model fit criteria.

Aside from weather, as expected, the addition of variables related to the equipment rebuild resulted in an increase in estimated savings at the facility.

Backcast Model. The evaluation team doubted that the backcast modeling approach was necessary for this facility. Despite the equipment rebuild, there were no dramatic changes in energy consumption at this facility during the program. When we applied the backcast methodology, the estimated savings decreased dramatically in SEM years 1 and 3, and produced the second lowest savings estimates in SEM year 2.

Table 18 presents the percentage savings and R^2 associated with each of the models considered by the evaluation team in the exploratory analysis.

Table 18. Case Study 1 Specifications, Savings, and R2

Model	Description	Year 1	Year 2	Year 3	R2
Forecast	MT&R model	0.61%	0.60%	1.20%	0.9540
Pre-post 1	MT&R model with added SEM year indicators	0.41%	1.33%	0.92%	0.9863
Pre-post 2	Initial automated stepwise selection	-0.40%	0.09%	-0.43%	0.9897
Pre-post 3	Pre-post 2 with HDD/CDD replacing mean temp	-0.21%	0.46%	-0.15%	0.9898
Pre-post 4	Pre-post 3 with rebuild indicator	-0.16%	0.44%	0.11%	0.9898
Pre-post 5	Pre-post 4 with rebuild/temperature interaction	-0.17%	0.45%	0.12%	0.9898
Backcast	As per SEP, separate model for each year	-0.76%	0.40%	-2.65%	N/A

Conclusions from Case Study 1. The evaluation team and EPT team arrived at two primary conclusions based on the above findings. First, the model showed an improvement in selection criteria when weather variables were included, implying that weather is an important predictor of consumption. Savings decreased when weather was added to the models, suggesting that omitting weather may bias the savings estimates.

Second, it was important to indicate changes in facility energy consumption unrelated to EM Program in the model. A rise in savings for year 3 (-0.15% to 0.11%) was evident after the evaluation team included an indicator for the equipment rebuild in the model. The indicator for the rebuild accounted for the rise in consumption in the third program year. The estimated savings would have been downwardly biased if the rebuild variable had not been included.

Case Study 2

Overview of Facility 2

The second facility selected for exploratory analysis had one primary output. A percentage of the facility output was produced using a second process that demanded greater energy usage. The facility data included weekly summaries for consumption, production, and the percentage of the specialized output requiring additional processing. The baseline period for SEM year 1 and year 2 began on October 31, 2010, and ended on October 1, 2011. The year 1 and year 2 SEM engagement began on November 1, 2011, and continued through October 31, 2013.

During SEM year 2, the production of specialized output began to increase. By SEM year 3, this output had increased considerably, to the extent that specialized production during year 3 was outside the baseline period range. As a result, the EPT team decided to re-baseline (i.e., select a new baseline period), to better represent baseline conditions. This new baseline period began on August 5, 2012, and ended on March 1, 2014, and covered parts of the year 1 and all of year 2. Year 3 of SEM engagement ran from March 2, 2014, to October 25, 2014.

The revised baseline period gave the EPT team and evaluation team an opportunity to explore the implications of re-baselining and the effect of using multiple energy consumption models for a single facility. Additionally, as with the first case study, the selection of this facility for a deeper statistical analysis allowed the evaluation and EPT teams to further explore differences between evaluation and MT&R savings estimates and the impacts of adding a variety of weather variables to the energy consumption models.

Exploratory Approach for Facility 2

The evaluation team tested 14 facility energy consumption models:

- SEM years 1 and 2, and SEM year 3 (after re-baselining)
  - A pre-post version of the MT&R model
  - Four evaluation pre-post models
  - A forecast model
  - A backcast model

We determined the model specifications based on the following considerations:

Weather Variables. The evaluation and EPT teams were interested in determining if weather was a significant driver of energy consumption at industrial facilities and should routinely be included as an explanatory variable in energy consumption models. Based on results from the first case study, both teams agreed that evaluators should consider not just HDD and CDD, but also average temperature as explanatory variables in future energy consumption analyses, therefore an average temperature variable was tested for this case study.

Percentage of Output That Was Specialized. Specialized output began to increase in SEM year 2 and continued to increase in SEM year 3. The maximum specialized output in the original baseline period was 3.5% of total output. In SEM years 1 and 2 the range widened, reaching as high as 5.7% and 6.5%, respectively. In SEM year 3, the minimum specialized production was 5.7% of total output and the maximum rose to 13.4%. The evaluation team agreed with the EPT team's determination that establishing a new baseline was justified since the original baseline range for specialized output was not representative of the SEM year 3 range.

The evaluation team tested three variable forms of the specialized output: the raw percentage value, an indicator of whether specialized output was present (i.e., that specialized output was greater than 0%), and an interaction between total output and the percentage of total output that is specialized. We added each of these variables to the set of candidate variables for stepwise selection.

Non-Linear Relationship of Production and Consumption. The energy consumption model selected by the evaluation team prior to the exploratory analysis did not explain a considerable amount of energy consumption based on model residual plots. The evaluation team tested various non-linear transformations of production in an attempt to reduce the unexplained variation.

Findings of Exploratory Analysis for Facility 2

Replication of MT&R Model Results. Using the MT&R data, the evaluation team was able to replicate the estimated MT&R models. The team then applied the forecast methodology and was able to replicate the MT&R savings results for year 3; however, we were not able to replicate the MT&R savings results for years 1 and 2. Instead, the team found more than double the MT&R savings for year 1 and approximately 75% of the MT&R savings for year 2, and was unable to determine the cause of these disparities.

Sensitivity Analysis. The evaluation team conducted a sensitivity analysis to investigate the following:

- The impacts of including HDDs and CDDs instead of mean temperature as explanatory variables
- How to best model the specialized output
- Whether there is a non-linear relationship between output and energy consumption

The evaluation team found that weather was an important variable for this facility. Accounting for weather as some combination of HDD, CDD, or mean temperature improved model fit considerably, though the specific choice of variable had little effect on model fit criteria and savings estimates. The evaluation team selected HDD as the optimal explanatory variable for both evaluation models, based on model fit criteria.

The stepwise selection process chose different specialized output variables for the SEM years 1 and 2 model versus the SEM year 3 model. For the SEM years 1 and 2 model, stepwise selection chose a model that included an interaction term between regular output and the percentage of total output that was specialized. For the SEM year 3 model, stepwise selection chose a model that included both the total output and the raw percentage of output that was specialized.

Finally, the evaluation team tested quadratic and cubic transformations of the production variable to account for unexplained variation in the model residuals. In SEM years 1 and 2, the non-linear transformations led to poor model performance, with low-scoring model fit criteria and savings estimates reduced to negative values. Conversely, in SEM year 3, the non-linear transformation greatly increased savings estimates, provided an optimal fit (as determined by all model fit criteria), and reduced unexplained variation in the model residuals. The evaluation team concluded from this that as specialized output increases, the production-energy usage relationship strays from being linear.

Backcast Model. The facility chosen for the second case study had a significant change in specialized production in SEM year 3. The evaluation team found that for all SEM years, the backcast models provided savings estimates that were similar to the evaluation savings estimates.

Savings Estimates. Table 19 and Table 20 present the percentage savings and R^2 associated with each of the models considered by the evaluation team in the exploratory analysis. Table 19 presents results for the SEM years 1 and 2 models, and Table 20 presents results for the SEM year 3 models.

Table 19. Case Study 2 SEM Years 1 and 2 Specifications, Savings, and Adjusted R2

Years 1 & 2 Model	Description	Year 1	Year 2	Adjusted R2
Forecast	MT&R model	0.39%	2.19%	0.9349
Pre-post 1	MT&R model with added SEM year indicators	0.93%	1.56%	0.9349
Pre-post 2	Initial automated stepwise selection	1.66%	2.16%	0.9456
Pre-post 3	Pre-post 2 with added raw % specialized production	1.45%	2.20%	0.9451
Pre-post 4	Pre-post 2 change HDD to mean temperature	1.73%	2.15%	0.9383
Pre-post 5	Pre-post 2 with added quadratic and cubic production and removed production-specialized production interaction	-0.76%	0.40%	*0.9537
Backcast	As per SEP, separate model for each year	1.11%	1.66%	N/A

*Pre-post 5 has the highest R^2 of the candidate models but performed the worst in both AIC and BIC model fit criteria.

Table 20. Case Study 2 SEM Year 3 Specifications, Savings, and Adjusted R^2

Year 3 Model	Description	Year 3	Adjusted R2
Forecast	MT&R model	5.33%	0.9000
Pre-post 1	MT&R model with added SEM year indicators	5.20%	0.8999
Pre-post 2	Initial automated stepwise selection	0.59%	0.9021
Pre-post 3	Pre-post 2 with removed production-% specialized production interaction	0.53%	0.9033
Pre-post 4	Revised automated stepwise candidate variables with mean temperature added for consideration	1.77%	0.9017
Pre-post 5	Pre-post 3 with added quadratic and cubic production, removed specialized production indicator, and added raw % specialized production	2.59%	0.9568
Backcast 1	As per SEP, separate model for each year, on baseline 1	1.61%	N/A
Backcast 2	As per SEP, separate model for each year, on baseline 2	1.69%	N/A

Conclusions from Case Study 2. The evaluation and EPT teams found that weather was an important determinant of facility energy consumption. However, in contrast to the case study 1 exploratory analysis, the specific variable chosen for modeling weather in the SEM years 1 and 2 did not have a significant impact on savings estimates.

The evaluation team found that the estimated savings depended on how specialized production was modeled. The SEM years 1 and 2 model used a different approach to incorporating specialized production than the SEM year 3 model. In SEM years 1 and 2, the interaction between production and the percentage of output that was specialized provided the optimal fit. In SEM year 3, the selected model included the raw percentage of production that was specialized. The evaluation team found that non-linear transformations of production had dramatic effects for this facility that were specific to the years modeled. The non-linearity in production is likely related to the increased specialized production in SEM year 3.

The evaluation team found that estimates of energy savings and model performance for this facility varied based on which variables were selected for each model. Scenarios in which many potential transformations of explanatory variables are possible make a case for the use of automated procedures or other machine learning methods for variable selection when the engineering relationship is unknown.

Based on changes in production of the specialized output for this facility, the evaluation team expected that the backcast model would be an appropriate evaluation approach. The backcast model and pre-post model yielded very similar savings estimates for SEM year 3. The evaluation team tested the backcast model on both the original and revised baselines, obtaining nearly identical results (at 1.69% and 1.61% savings, respectively). These results suggest that the backcast approach may be a viable evaluation approach for facilities that experience significant changes in the level of production between the baseline and program periods.

Case Study 3

Overview of Facility 3

The third case study was a facility that split production of output across four different floors. The facility data included weekly energy consumption and output production for floors 1 and 2 combined and separately for floors 3 and 4. The baseline period for SEM years 1 and 2 began on October 31, 2010, and ended on October 1, 2011. The data covered SEM years 1 and 2, beginning on November 1, 2011, and continuing through October 31, 2013. The facility shut down for one or more days during a number of weeks. The facility data included shut-down days as a variable, and the evaluation team included this variable in all models.

Similar to the case study 2 facility, this facility had a substantial increase in production of output during the SEM engagement period, beginning in SEM year 1. By SEM year 3, production had risen substantially. As a result, the EPT team decided to establish a new baseline for year 3, hoping to better represent baseline conditions under increased production. The new baseline period began on November 6, 2011, and ended on February 1, 2014. SEM year 3 ran from February 2, 2014, to October 25, 2014.

The revised baseline gave the EPT and evaluation teams a second opportunity to explore the modeling implications of re-baselining, the use of multiple energy consumption models for a single facility, and fundamental changes in a facility's output and energy usage during SEM engagement. Additionally, the evaluation and EPT teams continued to explore differences between the evaluation and MT&R savings estimates and the impacts of including a variety of weather variables in the energy consumption models.

Exploratory Approach for Facility 3

The evaluation team tested a total of 19 models:

- SEM years 1 and 2
  - A pre-post version of the MT&R model
  - 6 evaluation pre-post models, including one autoregressive AR(1) model
  - A forecast model
  - A backcast model
- SEM year 3
  - A pre-post version of the MT&R model
  - 7 evaluation pre-post models, including one autoregressive AR(1) model
  - A forecast model
  - A backcast model

The evaluation team determined the model specifications based on the following considerations:

Weather. The data provided by the EPT team included two weather variables: average dry bulb temperature and average dry bulb temperature with a change point. The evaluation team tested model specifications with HDD and CDD in addition to the weather variables provided by the EPT team.

Increased Facility Production. Overall, the facility nearly doubled production over the course of the SEM engagement. The increased production did not affect all production floors equally. Production floor 2 did not change substantially and floor 1 only had a slight increase. Production floor 3 had a much more noticeable increase in production, while floor 4 had the most dramatic increase in production, with very little production in SEM year 1 that rose to a level similar to that of the other three floors.

Weather-Production Interaction. The evaluation team found that the energy consumption model could not explain a significant portion of energy consumption in year 3. The model had a tendency to over-predict before production increased and under-predict after production increased. As the EPT team recommended, the evaluation team also tested for interactions between weather and production.

Serial Correlation. The evaluation team's diagnostic tests revealed that there might be serial correlation in the data; this was based on examining model residuals, as well as autocorrelation and partial autocorrelation plots.

Findings of Exploratory Analysis for Facility 3

Replication of MT&R Model Results. The evaluation team was able to replicate the MT&R model results and savings estimate for SEM year 3. The evaluation team was able to replicate neither the coefficients nor the savings estimates for the SEM years 1 and 2 models.

Backcast Model. For SEM years 1 and 2, the facility experienced relatively small changes in production on floors 1, 2, and 3. In these years, the backcast model produced much larger estimates of energy savings than the MT&R forecast and pre-post models, though estimated savings were still negative. Applying the backcast methodology to year 3, when production increased the most, led to different results depending on which baseline was used: backcasting a SEM year 3 model onto the original baseline produced a savings estimate of -5.08%, while backcasting onto the revised baseline produced a percentage savings estimate of 0.41%. The percentage savings estimate calculated from applying the year 3 backcast model to the original baseline may be biased due to extrapolation. The range of weekly energy usage in the year 3 had no overlap with the range from the original baseline, so predictions of energy usage were made outside of the data set used to specify the model.

Sensitivity Analysis. The evaluation team found that energy usage for the case study 3 facility was sensitive to the selection of production floor, form of the weather variables, and autocorrelation. For the SEM years 1 and 2 models, the automated stepwise process did not select the fourth production floor variable. For the SEM year 3 model, the automated stepwise process selected floor 1 production, floor 2 production, and combined floors 3 and 4 production. The evaluation and EPT teams decided to model all production floors individually (with the exception of the floor 1 and 2 production being summed). The fits of the SEM years 1 and 2 models and the SEM year 3 models improved by separating production into its components. Additionally, for SEM years 1 and 2, separating production increased savings estimates in both years, while for SEM year 3, savings estimates slightly decreased.

The model fit statistics for each of the models suggest that the specific weather variable selected was less important than ensuring that a weather variable was included. Model fit improved regardless of the weather variable included. For both models, the specific weather variable included as a regressor had very little effect on the estimated savings.

The evaluation team found that for all years, R^2 was improved by accounting for serial correlation in the model error. For the SEM years 1 and 2 models, the AR(1) model resulted in a slight increase in estimated savings, while for the SEM year 3 models, there was a substantial decrease in savings.

Savings Estimates. Table 21 and Table 22 present the percentage savings and R^2 associated with each of the models considered by the evaluation team in the exploratory analysis. Table 21 displays results for the SEM years 1 and 2 models, showing all negative savings estimates, indicating that the facility did not achieve EM Program-related energy savings in these years.

Table 21. Case Study 3 SEM Years 1 and 2 Specifications, Savings, and A Adjusted R2

Years 1 & 2 Model	Description	Year 1	Year 2	Adjusted R^2
Forecast	MT&R model	-1.68%	-14.34%	0.7500
Pre-post 1	MT&R model with added SEM year indicators	-1.11%	-2.49%	0.6238
Pre-post 2	Initial automated stepwise selection	-3.34%	-4.45%	0.6982
Pre-post 3	Pre-post 2 and remove all temperature variables	-3.16%	-4.36%	0.3952
Pre-post 4	Pre-post 3 with added mean temperature	-3.52%	-4.52%	0.4445
Pre-post 5	Pre-post 3 with added mean temperature with change point	-3.17%	-4.38%	0.6242
Pre-post 6	Pre-post 2 with added production floor 4	-2.77%	-3.80%	0.6949
Pre-post 7	Pre-post 6 with AR(1)	-2.54%	-3.77%	0.7533
Backcast	As per SEP, separate model for each year	-0.87%	-1.31%	N/A

Table 22 displays results for the SEM year 3 models.

Table 22. Case Study 3 SEM Year 3 Specifications, Savings, and Adjusted R2

Year 3 Model	Description	Year 3	Adjusted R^2
Forecast	MT&R model	3.57%	0.8580
Pre-post 1	MT&R model with added SEM year indicators	2.45%	0.8577
Pre-post 2	Initial automated stepwise selection	2.43%	0.8739
Pre-post 3	Pre-post 2 and remove all temperature variables	1.77%	0.7624
Pre-post 4	Pre-post 3 with added mean temperature	2.51%	0.8572
Pre-post 5	Pre-post 3 with added mean temperature with change point	2.33%	0.8740
Pre-post 6	Pre-post 5 with separated production floors	2.28%	0.8738
Pre-post 7	Pre-post 6 with AR(1)	-1.12%	0.9329
Pre-post 8	Pre-post 7 with temperature-production interaction	-0.58%	0.9352
Backcast 1	As per SEP, separate model for each year, baseline 1	-5.08%	N/A
Backcast 2	As per SEP, separate model for each year, baseline 2	0.41%	N/A

Conclusions for Case Study 3. As with the first two case studies, the evaluation team concluded that weather played an important role in this facility's energy consumption. This conclusion is also shown by the two EPT team models, which both include weather as an explanatory variable. The evaluation team found that for all three years, while the energy savings estimates did not depend on the specific weather variable selected for the model, the savings estimates were sensitive to the inclusion or omission of a weather variable.

The evaluation team used backcasting to estimate this facility's savings. The facility significantly increased production during the year 3, so baseline period conditions were not inclusive of engagement period conditions. The backcast savings estimates for the SEM year 3 model were not robust. The evaluation team obtained different results depending on which baseline was used with the estimated backcast model.

Accounting for serial correlation in the energy consumption model estimation improved the model fit. The evaluation team found evidence of autocorrelation in the SEM years 1 and 2 models and SEM year 3 models. It is important for evaluators to test for autocorrelation and if there is evidence of autocorrelation, to control for it.

K. Negative Savings Details

This appendix describes the evaluation team's treatment of negative SEM savings estimates for some EM program facilities. It first describes the scenarios that can lead to negative savings estimates. Then it presents and compares BPA's and the evaluation team's conventions for reporting negative SEM savings.

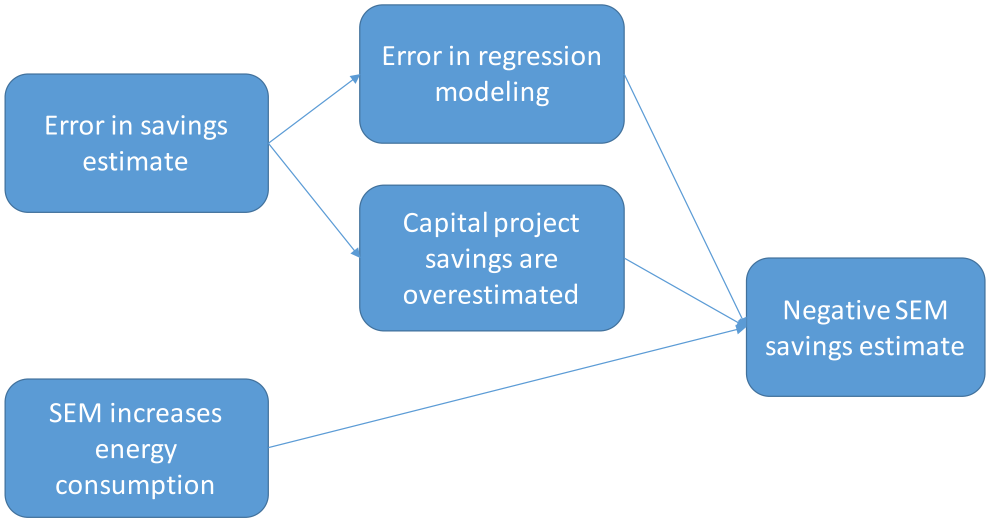
The BPA EPT team and the evaluation team estimated SEM savings by taking the difference between the regression-based estimate of facility savings and the engineering-based capital project savings:

SEM savings = Regression-Based Facility Savings – Capital Project Savings

When the estimate of the facility savings is negative or the capital project savings exceeds the facility savings estimate, the estimated SEM savings will be negative.

Negative SEM savings may occur for three reasons, as shown in Figure 37. 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. Each scenario is discussed below.

Figure 37. Sources of Negative SEM Savings Estimates



Capital Project Savings Are Overestimated

Overestimation of the capital project savings will lead to underestimation of SEM savings. If the capital project savings are sufficiently overestimated, so that they are larger than the estimated facility savings, then the SEM savings are negative. For example, if the true capital project savings are 1.5%, but the estimate of capital project savings is 2.5%, the estimated SEM savings would be negative if the facility savings estimate is less than 2.5%.

The Regression-Based Facility Savings Estimate is Erroneous

Electricity consumption in industrial facilities is often very complex. The largest known energy drivers (e.g., facility production) are typically measured and used as inputs in the regression model. Over the course of a multi-year engagement, non-programmatic effects (e.g., product changes or facility expansions) may take place, and need to be accurately reflected in the model specification. Additionally, some factors affecting consumption may be unmeasured and omitted. If these non-programmatic effects or omitted factors are correlated with SEM implementation, the SEM savings estimates may be biased.

SEM Caused Energy Consumption to Increase

SEM implementation could cause facility energy consumption to increase. For this to occur, the facility would have to intensify its use of energy in the production process. Energy consumption intensity could increase if an efficiency strategy was implemented incorrectly or a strategy was implemented with an incorrect understanding of the facility production process.

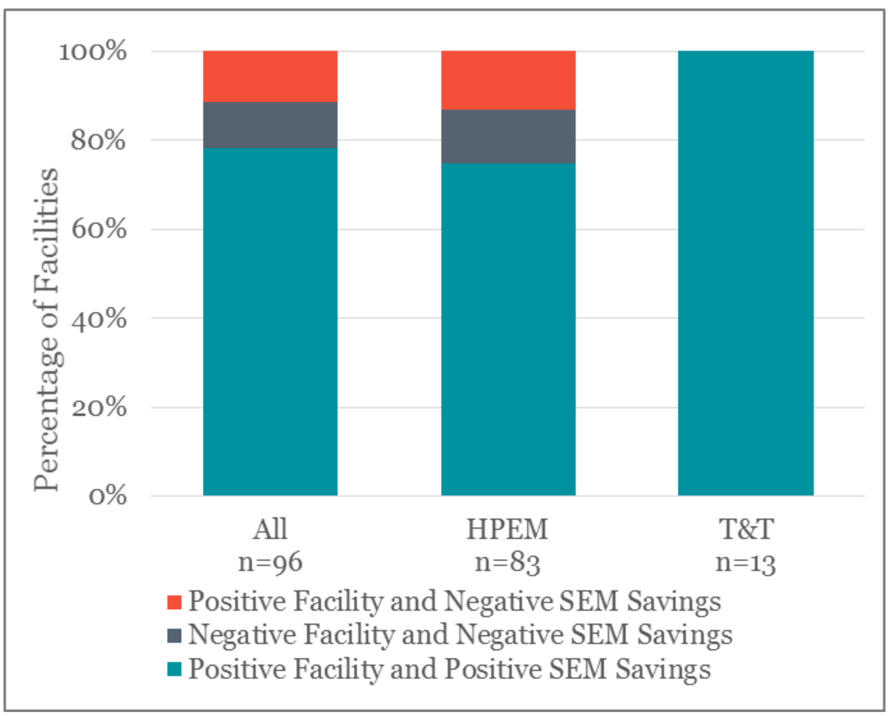
Situations in which implementation of SEM leads to an increase in energy intensity are expected to occur rarely. When negative savings estimates occur, it is more likely that error in regression modeling or in the capital project savings estimate is responsible.

How Significant of an Issue Was Negative Savings Estimates?

Both BPA's EPT team and the evaluation team estimated negative SEM savings for some facilities and years. In 78% of all facilities and years, both facility and SEM savings were positive, that is, the SEM savings estimate was positive after subtracting savings from capital improvements. In 10% of facility-years, the facility savings estimate was negative, and in 12% of facility-years, the facility savings estimate was positive but smaller than the capital project savings estimate. The sum of negative SEM savings estimates for all facilities equaled -0.3% of consumption.

Figure 38 presents the distribution of facilities by sign of estimated facility savings and SEM savings for each program year. In 63% of facilities and years, a facility had capital projects savings, and in 18% of those cases (11 of 60), the capital project savings estimate was larger than the facility savings estimate.

Figure 38. Percent of Facilities with each Savings Scenario



Reporting of Negative SEM Savings Estimates

BPA and the evaluation team employed different conventions for reporting negative SEM savings estimates. BPA reported negative SEM savings estimates as zero, reasoning that it was unlikely that the ESI Energy Management Program could have caused an increase in energy use intensity. Any increase in energy consumption after controlling for changes in output, weather, and other variables was likely caused by other changes at the facility that were not measured and therefore unaccounted for in the energy consumption regression model.

In contrast, the evaluation team reported the unadjusted negative savings estimate in the estimating program savings.

Evaluation Team Assessment of Reporting Conventions

Although it is more likely that a negative SEM savings estimate reflects error in modeling consumption or capital project savings, it is not possible to differentiate between negative savings estimates that arise because of modeling error and those that arise because of actual increases in energy consumption intensity. As there is no valid, auditable basis for identifying the causes of negative savings estimates, facilities with negative savings estimates should not be excluded from the analysis sample and their savings estimates should not be modified.

Another important reason for preserving negative savings estimates is that error in modeling consumption or capital project savings can affect facilities with either positive or negative savings. Large positive savings estimates may entail positive modelling errors, but these facilities are not being flagged for exclusion or censoring.

Furthermore, best practice in impact evaluation requires choosing an impact evaluation methodology and applying that methodology consistently to the observations in the analysis sample.[54] Sample selection must occur before conducting the analysis and not be based on the estimates of the outcomes that the evaluation is measuring. When BPA reports negative SEM savings estimates as zero, this approach effectively excludes some observations based on the outcome and this does not conform to evaluation best practices.[55]

---

54 When evaluators use regression analysis of individual building consumption to estimate savings in other sectors, the industry standard is to accept both positive and negative savings results for individual sites. The biggest body of evidence is in the residential sector, where regression-based billing analysis is used frequently. The results are often expressed as average savings, but the underlying distribution of savings almost always has some percentage of cases where estimated savings were negative. Recent examples include: weatherization, ductless heat pumps, and behavior savings.

55 According to Greene (2012, p. 141), “In principle, an ‘outlier’ is an observation that appears to be outside the reach of the model, perhaps because it arise from a different generating process… Unusual residuals are an obvious

The evaluation team also discussed this issue with outside experts. Four independent experts were informally asked whether they might treat negative savings estimates as zero, and all agreed that negative savings estimates should not be excluded.[56] The BPA project manager also discussed with external stakeholders, including another consulting firm, an evaluation colleague at a regional entity conducting evaluation in this area, RTF staff and Council staff. All indicated that the exclusion of negative savings would not be appropriate.

In summary, facilities with negative savings estimates should be left in the analysis sample unless it can be demonstrated that the baseline is invalid because it cannot account for one or more factors affecting energy use. It should also be demonstrated that any test used to exclude facilities does not have a bias towards removing facilities with negative modelling errors.

Conclusion

The evaluation team understands BPA's reasoning for reporting negative savings estimates as zero savings. However, there is no rigorous way to differentiate between negative savings estimates that arise because of modeling error and negative savings estimates that reflect actual increases in energy consumption intensity. Accordingly, negative facility savings estimates should be reported. Reporting negative savings estimates as zero will cause upward bias in the program savings.

References

Greene, William, 2012. Econometric Analysis (7th International Edition). New York: Pearson.

Reichmuth, Howard, 2013. Independent M&V Report. Puget Sound Energy Manufactured Home Duct Sealing (MHDS) Program. http://rtf.nwcouncil.org/meetings/2015/08/PSE_%20MHDS_%20Analysis-Final_12%205%2012%20Rev1.docx

choice [for identifying outliers.] But, since the distribution of disturbances would anticipate a certain small percentage of extreme observations in any event, simply singling out observations with large residuals is actually a dubious exercise." Similarly, it would appear that a facility with negative savings cannot be reconciled with *a priori* beliefs about SEM program effects.

[56] IPMVP committee members were informally questioned as to whether they might exclude facilities with negative savings estimates, such as some facilities within an Energy Savings Performance Contract portfolio. Four responses were received, and all respondents stated that negative savings have to be included. Respondent 1: "…If some of the sites have negative savings, they have to be taken into account and subtracted from other savings to assess the overall performance of the project… And that's consistent with what I see in Federal ESPC projects." Respondent 2: "…I'd be reluctant to "discount" any results…" Respondent 3: "Negative savings cannot be ignored unless you have verified non-routine adjustments to account for them. We have done several projects where we are aggregating savings from multiple sites, and all the negative savings sites have been included." Respondent 4: "Unfortunately, sometimes some sites indeed have negative savings. Of course we never ignore these results…"

L. Participant Survey

Researchable Topics	Item
Company and contact information	Section A
Customer Commitment – Policy and Goals	Section C
Customer Commitment – Resources	Section B
Planning and Implementation – Energy Assessment and Energy Map	Section D
Planning and Implementation – Action Plan, Employee Engagement, and Implementation	Section 0
Planning and Implementation – Metrics and Goals, Measurement and Reporting	Section F

CEE Minimum Element	Data Source for HPEM	Data Source for T&T
1a. Customer Commitment – Goals	C1, C2	C1, C2
1b. Customer Commitment – Resources	B1, B2	B1, B2
2a. Planning & Implementation – Energy Management Assessment	D1	D1
2b. Planning & Implementation – Energy Map	D2	D2
2c. Planning & Implementation – Metrics and Goals	F2	F2
2d. Planning & Implementation – Project Register	E1	E1
2e. Planning & Implementation – Employee Engagement	E3	E3
2f. Planning & Implementation – Implementation	Completion Reports	Completion Reports
2g. Planning & Implementation – Reassessment	C3, E2	C3, E2
3a. System for Measuring and Reporting Energy Performance - Measurement	F1, F3	F1, F3
3b. System for Measuring and Reporting Energy Performance – Data Collection and Analysis		
3c. System for Measuring and Reporting Energy Performance – Analysis		
3d. System for Measuring and Reporting Energy Performance – Reporting	B3, F4	B3, F4

Target Quota = 32 (Census)

Variables to be pulled into interview

- Contact Name
- Facility/company name
- EM Type (HPEM 1, HPEM 2, or T&T)
- Year began EM
- Utility name / Program name for HPEM 2 facilities
- Were they contacted as part of the capital measure evaluation?
    - If HPEM 2, utility program name they participated in

General Instructions

- Interviewer instructions are in green [LIKE THIS].
- CATI programming instructions are in red [LIKE THIS].
- Items that should not be read by the interviewer are in parentheses like this ( ).

A. Introduction

A1. May I speak with [CONTACT NAME]? [IF THAT PERSON IS NOT AT THIS PHONE NUMBER, ASK FOR NAME AND PHONE NUMBER AND START AGAIN]

1. (Yes)
2. (No, person is not able to come to phone) [GET NAME, PHONE NUMBER, AND SCHEDULE CALLBACK]
3. (No, person no longer works there) [ASK FOR THE CONTACT NAME AND PHONE NUMBER FOR THE PERSON MOST FAMILIAR WITH PARTICIPATING IN {SEM TYPE} IN {SEM YEAR}]

98. (Don't know) [ASK TO SPEAK WITH SOMEONE WHO KNOWS AND BEGIN AGAIN]
99. (Refused) [THANK AND TERMINATE]

A2. Hello, I'm [INSERT NAME] calling from Cadmus on behalf of [IF HPEM COHORT 1 OR T&T "Bonneville Power Administration Energy Smart Industrial" OR IF HPEM COHORT 2 "UTILITY NAME/ PROGRAM NAME"]. We are conducting an important study with participants in industrial energy management programs to understand the program's influence in helping facilities adopt strategic energy management practices, often referred to as SEM. It is our understanding that you are the energy champion at [facility name]. Is this correct?

1. (Yes)
2. (No) [ASK FOR THE CONTACT NAME AND PHONE NUMBER FOR THE ENERGY CHAMPION]

98. (Don't know) [ASK TO SPEAK WITH SOMEONE WHO KNOWS AND BEGIN AGAIN]
99. (Refused) [THANK AND TERMINATE]

A3. How long have you had this role of energy champion?

A4. One objective of our study is to compare SEM practices at facilities who are receiving support from SEM programs to those who are not receiving support. We have a standard set of questions for both groups, and we recognize that many of the activities we will ask you about are part of your work with Cascade. However, we need to ask these questions to retain consistency, and so all we need is verification from you that these activities have been implemented.

READ IF PART OF CAPITAL EVALUATION SAMPLE "We are aware that you were already contacted for the evaluation of equipment or custom projects which received rebates through the [IF HPEM 1 OR T&T "Energy Smart Industrial program" OR IF HPEM 2 "Utility / program name"]. That was a related study, but focused on capital projects instead of practices."

We expect this call to take about 5 minutes. Is this a good time?
IF NOT A GOOD TIME, ASK TO SCHEDULE A TIME TO CALL BACK

Before we get started, I'd like to note that your responses are confidential and will only be publicly reported in aggregate. Individual facility responses will not be identified in public documents. [IF NEEDED: individual responses will be reported anonymously as part of a group. We will not publicly report any identifying information]

B. *Customer Commitment - Resources*

B1. Do you have an **energy team** [dedicated staff for energy and energy efficiency] at your facility?

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

[ASK B2 IF B1= YES]

B2. How frequently does the energy team meet?

1. (Daily)
2. (Weekly)
3. (Monthly)
4. (Quarterly)
5. (Twice a year)
6. (Annually)
7. (Other) [SPECIFY]
8. (We don't meet)
98. (Don't know)
99. (Refused)

B3. Does your senior management require regular updates from the energy team?

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

C. Customer Commitment - Energy Policies & Goals

C1. Does your company or facility currently have goals or action item plans to improve energy performance?

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

[IF C1=YES, ASK C2 THROUGH C3]

C2. Have the energy performance goals or policies been communicated to staff?

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

C3. Have you reviewed the goals since they were set to ensure they still align with business and energy performance priorities?

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

D. Planning & Implementation - Energy Management Assessment and Energy Map

D1. [IF HPEM COHORT 1 OR 2] Our records show that an energy management assessment was conducted as part of your participation in HPEM. Is that correct?
[IF T&T] Has your company completed an energy management assessment? [IF NEEDED: This is an assessment of the energy management structure that identifies how management can better support energy efficiency efforts.]

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

D2. [IF HPEM COHORT 1 OR 2] Our records show that an energy map was developed as part of your participation in HPEM. Is that correct?
[IF T&T] Has your company identified the key energy drivers or largest energy consumers? [IF NEEDED: This is a breakdown of energy end uses broken down by facility/processes either by estimated energy use or % of facility energy use.]

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

E. Planning & Implementation - Project Register, Implementation, and Employee Engagement

When you first started with [EM TYPE], an [IF HPEM: "OPPORTUNITY REGISTER" OR IF T&T: "TUNE UP ACTION ITEM LIST"] was developed listing potential energy-efficiency projects and activities at your facility.

E1. Are you still using the [IF HPEM: "OPPORTUNITY REGISTER" OR IF T&T: "TUNE UP ACTION ITEM LIST"]?

1. (Still using it)
2. (No longer using it)
98. (Don't know)
99. (Refused)

E2. Do you regularly update the [IF HPEM: "OPPORTUNITY REGISTER" OR IF T&T: "TUNE UP ACTION ITEM LIST"]? [IF NEEDED: INCLUDES ADDING NEW PROJECTS OR UPDATING PROJECTS ALREADY IN THE PLAN]

1. (Update regularly)
2. (Update occasionally)
3. (Almost never update it)
98. (Don't know)
99. (Refused)

E3. Has the energy team conducted any specific employee engagement activities? [IF NEEDED: INCLUDES ANY ACTIVITIES THAT INVOLVE STAFF OUTSIDE THE ENERGY TEAM, SUCH AS ENGAGING STAFF TO TURNING OFF EQUIPMENT WHEN NOT USED, AWARENESS CAMPAIGNS, ETC.]

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

F. MT&R Model

As part of HPEM or T&T, an energy model was developed and is periodically updated to track your energy usage and energy performance over time.

F2. Does the energy model use energy performance indicators to measure progress towards goals? [READ IF NEEDED: For example, an energy performance indicator could be energy consumption per unit of production]

1. (Yes)
2. (No)
98. (Don't know)
99. (Refused)

F3. How frequently is energy performance reviewed?"

1. (Daily)
2. (Weekly)
3. (Monthly)
4. (Quarterly)
5. (Twice a year)
6. (Annually)
7. (Continuously)
8. (Other) [SPECIFY]
98. (Don't know)
99. (Refused)

F4. How often is energy use data shared with others in your organization?

1. (Daily)
2. (Weekly)
3. (Monthly)
4. (Quarterly)
5. (Twice a year)
6. (Annually)
7. (Other) [SPECIFY]
98. (Don't know)
99. (Refused)

G. Closing

Those are all my questions. Thank you very much for your time and for your support of this important study. Have a great day!

M. SEM Adoption Scoring Methodology

Table 23. SEM Adoption Scoring Method

SEM Element	Survey Question(s)	Level of SEM Implementation: Full	Some	None
1a. Policy and Goals	▪ Does your company or facility currently have goals or action item plans to improve energy performance? ▪ Have the energy performance goals or policies been communicated to staff?	Have goals or action item plans, and these have been communicated to staff	Any other response combination	Don't have goals or action item plans (or DK)
1b.Resources	▪ Do you have an energy team [dedicated staff for energy and energy efficiency] at your facility? ▪ How frequently does the energy team meet?	Have an energy team that meets quarterly or more frequently	Any other response combination	No energy team (or DK)
2a.Energy Management Assessment	▪ [IF HPEM COHORT 1 OR 2] Our records show that an energy management assessment was conducted as part of your participation in HPEM. Is that correct? ▪ [IF T&T] Has your company completed an energy management assessment?	[IF HPEM 1 OR 2] Revisited or updated assessment [IF T&T] Completed an assessment	Any other response combination	[IF HPEM 1 OR 2] Did not revisit or update assessment (or DK) [IF T&T] Did not complete an assessment
2b. Energy Map	▪ [IF HPEM COHORT 1 OR 2] Our records show that an energy map was developed as part of your participation in HPEM. Is that correct? ▪ [IF T&T] Has your company identified the key energy drivers or largest energy consumers?	[IF HPEM 1 OR 2] Use/reference energy map developed through SEM [IF T&T] Completed an energy map	Any other response combination	[IF HPEM 1 OR 2] Do not use/reference energy map developed through SEM [IF T&T] Did not complete an energy map
2c. Metrics and Goals	▪ Does the energy model use energy performance indicators to measure progress towards goals?	Energy model has performance indicators to measure progress towards goals	Any other response combination	Energy model does not have performance indicators to measure progress towards goals (or DK)

SEM Element	Survey Question(s)	Level of SEM Implementation		
		Full	**Some**	**None**
2d. Project Register	▪ Are you still using the [IF HPEM: "OPPORTUNITY REGISTER" OR IF T&T: "TUNE UP ACTION ITEM LIST"]?	Still using opportunity register or action item list	Any other response combination	Not using opportunity register or action item list (or DK)
2e. Employee Engagement	▪ Has the energy team conducted any specific employee engagement activities?	Have conducted specific employee engagement opportunities	Any other response combination	Did not conduct specific employee engagement opportunities (or DK)
2f. Implementation	▪ Reviewed documentation (no questions in survey for this element)	Completed one or more projects	Any other response combination	Did not complete any projects
2g. Reassessment	▪ Have you reviewed the goals since they were set to ensure they still align with business and energy performance priorities? ▪ Do you regularly update the [IF HPEM: "OPPORTUNITY REGISTER" OR IF T&T: "TUNE UP ACTION ITEM LIST"]?	Update goals and update the opportunity register or tune up action item list regularly or occasionally	Any other response combination	Do not update goals (or DK), and almost never or never update the opportunity register or tune up action item list (or DK)
3a. Measurement **3b. Data Collection and Availability** **3c. Analysis**	▪ Do you reference the energy model developed through [HPEM or T&T] to track your energy performance? ▪ How frequently is energy performance reviewed?	Reference the energy model quarterly or more frequently	Any other response combination	Do not reference the energy model
3d. Reporting	▪ Does your senior management require regular updates from the energy team? ▪ How often is energy consumption data shared with others in your organization?	Senior management requires regular updates and shares energy consumption data with others in the organization quarterly or more often	Any other response combination	Management does not require regular updates (or DK), energy consumption data are shared with others in organization less often than quarterly (or DK)

N. SEM Sub-Element Adoption Scores and Energy Savings

Figure 39 shows the adoption level overall and for each minimum subelement on the x-axis versus the evaluated facility energy savings on the y-axis. The box plot shows the quartiles, with the median represented by the middle band within the box. The points represent individual facility evaluated SEM savings results.

Figure 39. Adoption Level of SEM Sub-Elements and Percentage Savings