Report **Group D – D01.02**

Workplan for 2018 Industrial Strategic Energy Management (SEM) Evaluation

Submitted to
California Public Utilities Commission
505 Van Ness Avenue
San Francisco, CA 94102

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

In association with
BuildingMetrics Inc.
Energy350
Opinion Dynamics
Ridge & Associates

Revised July 2, 2019

ENERGY • WATER • EFFICIENCY

Acknowledgements

We want to acknowledge the critical contributions from Lisa Paulo and Peter Lai, from Commissions Energy Division staff, who guided the development of the work plan and provided many clarifications regarding Commission policies and relevant practices, in addition to their critical review of this document.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Acronym/ Abbreviation	Definition
BRO	Behavioral, retro-commissioning and operational based energy savings
CMPA	Custom Measure and Project Archive
CPUC	California Public Utilities Commission
CUSUM Graph	Cumulative sum of differences graph – This is a graph that plots the cumulative difference between model-predicted energy use and actual energy use.
Custom Project Review	The custom project process estimates the potential energy savings for an energy efficient measure before it is installed based on predictions of typical operating conditions and baseline usage. The process is the basis for utilities to claim savings for their energy efficiency portfolios. The custom project process covers deemed measures (DEER and non-DEER), measures that have well-known and consistent performance characteristics, and custom measures.
ESPI	Efficiency savings and performance incentive
EUL	Effective useful life – The average time over which an energy efficiency measure results in energy savings, including the effects of equipment failure, removal, and cessation of use.
Hypothesis Model	Baseline NMEC model— This determines the independent variables and model coefficients. See NMEC.
IOUs	Investor owned utilities
kW	Kilowatts
kWh	Kilowatt-hours
M&V	Measurement and verification
MMBtu	Millions of British thermal units
NMEC	Normalized metered energy consumption – This is a means by which savings are quantified at the meter level through a regression approach that normalizes energy use by independent variables such as weather, production, etc.
NTGR	Net-to-gross ratio
Opportunity Register	A document of identified energy saving opportunities intended to be a living document that assigns actions to responsible parties and includes due dates for activities, notes, etc.
PA	Program administrators – For industrial SEM programs, all program administrators are investor owned utilities.
POE	Preponderance of evidence
PG&E	Pacific Gas and Electric

Acronym/ Abbreviation	Definition
RFP	Request for proposals
RUL	Remaining useful life – The number of remaining years that an item, component, or system is estimated to be able to function
SCE	Southern California Edison
SCG	Southern California Gas
SDG&E	San Diego Gas and Electric
SEM	Strategic energy management — A strategy for generating energy savings from large energy users by incorporating energy management into their internal management systems
SEM BRO	SEM savings that are quantified at a whole facility level using an NMEC approach consistent with the statewide SEM M&V Guide — While it is anticipated that this will be largely behavioral, retro-commissioning and operational measures, nothing precludes capital measures from being included. All non-SEM BRO project savings should be netted out of the SEM model, including SEM Custom, deemed measures, and non-incentivized projects that do not arise from SEM.
SEM Custom	A custom capital project occurring at an SEM site— Projects with this designation follow the custom capital project process with savings estimated using engineering analysis.
SEM Participant Tracking Report	Report containing key dates of customer participation and progress with a short summary of issues and activities
SEM Workshop Summary Report	Report detailing attendees and activities of each workshop
SOP	Standard operating procedure
Standard – Very Large Project NTGR	A version of the nonresidential self-report NTGR survey that involves the most detailed level of analysis that is typically applied to the largest and most complex projects— The estimation of the NTGR is often based on data obtained from program files, interviews with key decision-makers, vendors, and utility program staff, as well as reviews of other research finding such as studies of industry standard practice. The data is then analyzed using the preponderance-of-evidence approach to produce an estimate of the NTGR.
Treasure Hunt	A one-day facility scan conducted by a team of facility staff and SEM implementers who search for energy efficiency opportunities— Opportunities found are populated in the Opportunity Register. See opportunity register.
Workshops	Training sessions that gather SEM participants to deliver curriculum defined in the SEM statewide design guide— The guide calls for eight workshops; each covering different topics.

LIST OF STATISTICS TERMS

Term	Definition
Coefficients	Constants multiplied by the independent variable in a model— Once coefficients are determined, this (along with the y intercept) serves as the basis for the model.
Collinear	When one variable has high correlation with another variable— For example, raw material input versus output tend to be collinear and would generally not both be used in a model.
Independent Variable	Variables that affect energy use (energy drivers)
Measurement Boundary	Defines the boundary within which energy use is quantified— Small outbuildings or separately metered areas that are not a source of energy savings may often be excluded.
Non-Routine Event	An event that changes the energy use profile of a facility, triggering the need for an adjustment to the NMEC model— Examples include the addition of significant process equipment, changes in product mix, unplanned outages, etc.
Weighted Regression	Prediction model that reduces model errors by placing higher impact on certain data ranges— An example of weighted regression is using degree days instead of average temperature.

Executive Summary

The objective of this workplan is to plan and document our review and evaluation process and methodology regarding SEM evaluation activities.

Background

Strategic energy management (SEM) is a strategy for generating energy savings from large energy users by incorporating energy management into their internal management systems. The Consortium for Energy Efficiency (CEE) defines SEM as a holistic approach to managing energy use in order to continuously improve energy performance by achieving persistent energy and cost savings over the long term. It focuses on changing business practices by senior management through shop-floor staff, affecting organizational culture, and reducing energy waste.[1]

California Public Utilities Commission (CPUC) Decision 16-08-019[2] directed the investor-owned utilities (IOUs) to implement strategic-energy management programs. The decision defines SEM as a holistic, whole-facility approach that uses normalized metered energy consumption (NMEC) and a dynamic baseline model to determine savings from all program activities at the facility, including behavioral, retro-commissioning and operational (BRO) and custom projects. The decision calls for IOUs to administer their programs based on a consistent, statewide program design.

The California IOUs contracted with Sergio Dias Consulting to develop the Statewide Industrial SEM Design Guide and the Measurement and Verification (M&V) Guide. These documents serve as the framework by which each program administrator (PA) is directed to design and implement their SEM programs. The design guide calls for a two-year, cohort-based engagement with M&V periods that end at the end of years one and two.

Since the completion of the SEM design and M&V guides, the PAs have issued requests for proposals (RFPs), selected implementers, and written implementation plans for their industrial SEM programs. The PAs have recruited a total of 37 participants. The programs launched in 2018, and the PAs plan to begin claiming savings in 2019. While each PA's program has subtle differences, all generally follow the statewide design and M&V guides.

[1] Consortium for Energy Efficiency. *CEE Strategic Energy Management Minimum Elements.* https://library.cee1.org/system/files/library/11283/SEM_Minimum_Elements.pdf

[2] Public Utilities Commission of the State of California. 2016. Order Instituting Rulemaking Concerning Energy Efficiency Rolling Portfolios, Policies, Programs, Evaluation and Related Issues. http://ccag.ca.gov/wp-content/uploads/2016/10/6.5-Attachment-ALJ-Decision-16-08-019-081816.pdf

Scope of Work

This evaluation workplan outlines our approach for evaluating SEM, including the following key deliverables:

- Sample Design
- Early Feedback (Deliverable 13)
  - Participant Level Review
  - PA Level Review
  - Statewide Design & M&V Guide Review
- Custom Project Review (Deliverable 17)
  - SEM Custom Project Review
- Ex Post Evaluation (Deliverables 9 & 10)
  - Ex Post Gross Savings Evaluation (Deliverable 9)
  - Ex Post Net Savings Evaluation (Deliverable 10)
  - Cost Effectiveness Analysis
  - SEM Portfolio Impact Estimates

Coordination of SEM Model Review Activities

There are several touchpoints during the various SEM-evaluation activities that must be carefully coordinated, particularly with respect to reviewing models. The model-review touchpoints (depicted in Figure 1) are:

- **Early Feedback** – Early feedback will include a hypothesis (baseline) model review, which will be conducted shortly after the hypothesis models have been completed by the program implementers and reviewed by the PAs. Additionally, early feedback will include site visits to allow for activities such as review of identified and implemented opportunities, data collection and site staff interviews.
- **Custom Project Review** – We will conduct custom project reviews of all SEM custom projects at early feedback sites, following the CPUC procedures for reviewing custom projects. Additionally, we will conduct custom project reviews of the final models for all early feedback sites. We expect that, for early-feedback sites, PAs do not issue performance-based incentives until a custom project review of the final model has been conducted to ensure that incentives issued are consistent with the custom project review. This will evaluate the technical validity of the models and energy savings as well as the PA due-diligence process. Custom project reviews will result in dispositions that will inform efficiency savings and performance incentives (ESPIs).

- **Ex Post Evaluation** – The ex post evaluation will include reviews of the final model and M&V reports, site visits, interviews with facility staff, additional data collection, and M&V. This will determine the sample- and population-level net and gross energy savings and realization rates.

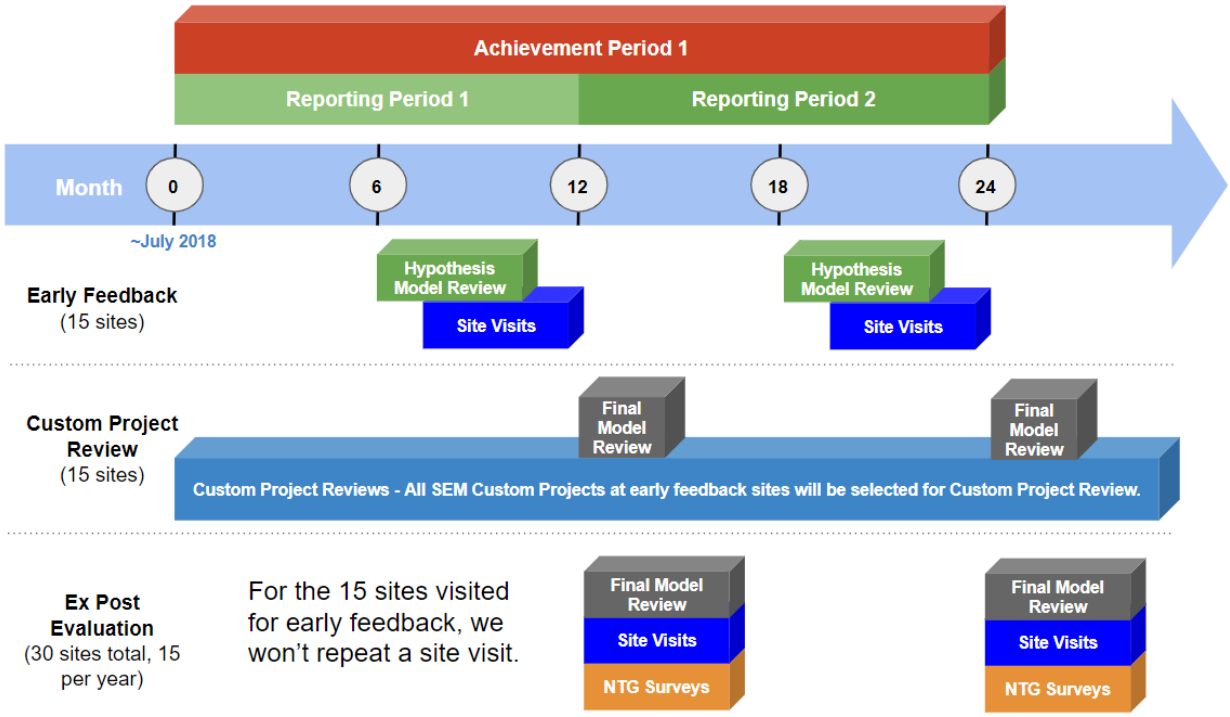


Figure 1: Model Review Touch Points

Sample Design

Given the interconnectedness of the SEM-related deliverables, we have crafted an approach that considers all SEM-related work together. This is especially critical with respect to sampling for several deliverables from a small population of only 37 SEM participants. We separated participants into two groups—those that receive early feedback, and those that do not. We then sampled 15 in each group using a stratified-random-sample approach. Table 1 shows that the sample represents the program with a 1.4% precision at 90% confidence.

Table 1: SEM Sample Summary

Early Feedback	Participants	Sample	Total Net MMBtu	Sample Estimate of Total Net MMBtu	Confidence Level	Relative Precision
No	18	15	116,348	119,095	90%	2.6%
Yes	19	15	141,565	139,505	90%	1.6%
All	37	30	257,913	258,600	90%	1.4%

Timeline

Figure 2 shows our schedule for completing this work. Note that all our activities will be completed twice—once for each of two reporting periods.

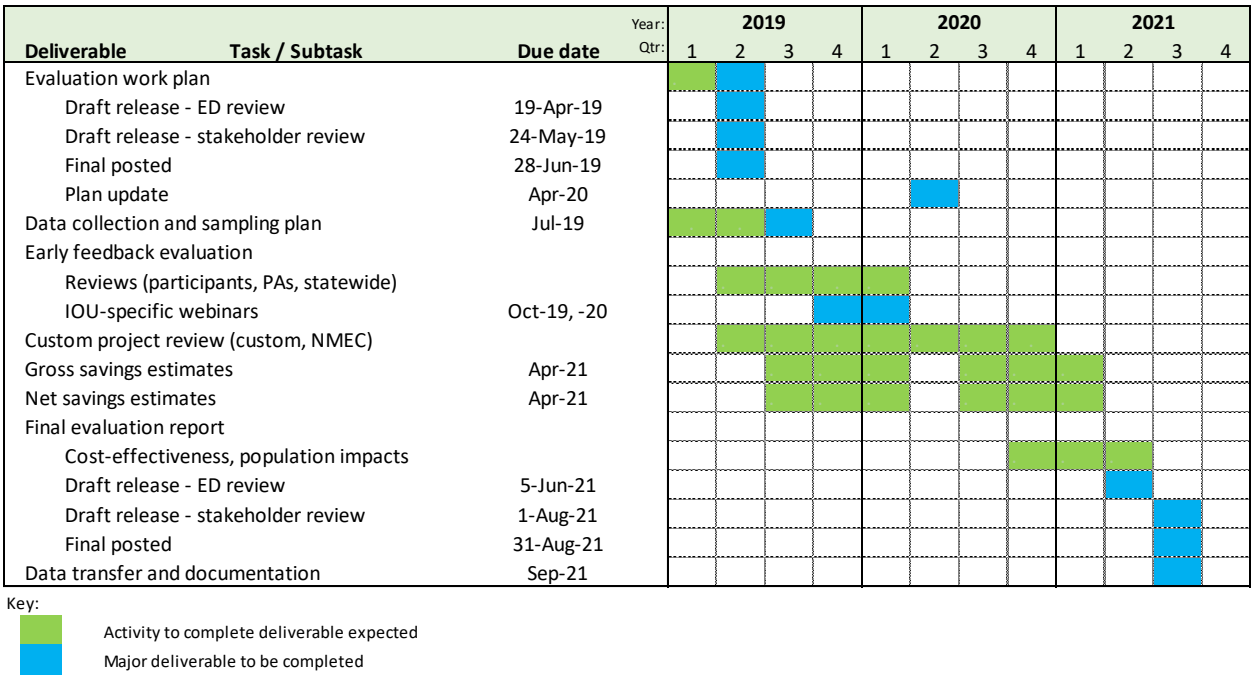


Figure 2: Schedule for SEM Evaluation Activities

1. Introduction and Background

In this section, we summarize our understanding of relevant CPUC policies and guidance, and the findings from evaluations of other SEM programs outside California.

1.1. Findings and Recommendations from Prior Evaluations

California SEM programs have not been evaluated as of May 2019. So, this section summarizes key findings from evaluations of similar programs conducted by Energy Trust of Oregon[3] and the Bonneville Power Administration (BPA).[4]

Energy Trust of Oregon published one evaluation for program years 2010-2013.[5] A second evaluation is nearing publication. The key recommendations from the published evaluation include:

- Program administrators should collect final versions of the model files and input data. As a result of incomplete data, version control issues between models and M&V reports and an inability to recreate model results, the evaluator was only able to evaluate 18% of sampled participants. This highlights the importance of collecting actual model files, not just data and reports.
- The SEM implementer should better train the participants in how to model nonroutine events. While the evaluation found that most participants continued to practice SEM and update the model, 80% of those interviewed also reported that their facilities underwent major changes after the conclusion of their SEM engagement and that they were unable to reflect those changes in the model.

BPA has published two evaluations of SEM. Key findings and recommendations from the most recent evaluation[6] include:

- The energy consumption of 22% of participants increased. While there are many potential explanations for the increases, the most likely explanation is inaccurate modeling. SEM models quantify savings that are a small portion of overall usage. This inevitably results in some uncertainty and inaccuracy. Since modeling inaccuracies can contribute to underestimating and overestimating savings, it is important to claim negative savings where they occur in order to offset savings overestimates at other facilities. Not claiming negative

[3] https://www.energytrust.org/industry-agriculture/industry-strategic-energy-management/

[4] https://www.bpa.gov/EE/Sectors/Industrial/Pages/Strategic-Energy-Management.aspx

[5] The Cadmus Group. (2019). Energy Trust Production Efficiency Strategic Energy Management Evaluation Final Report.

[6] SBW Consulting, Inc. & The Cadmus Group. (2017). Industrial Strategic Energy Management (SEM) Impact Evaluation Report.

savings means disregarding only those models with underestimated savings, which biases estimates of program savings.

- The presence of custom measures within SEM introduces uncertainty into the estimates of savings due to:
  - Deviations between expected and actual performance of custom projects
  - Accuracy of savings from deemed measures
  - Persistence of savings from custom measures
- The evaluation found no relationship between the rate of savings and the presence of Consortium for Energy Efficiency's (CEE) minimum elements.[7] CEE's minimum elements define the minimum conditions that a facility should have in place to effectively and continuously improve their energy performance.

1.2. Commission Guiding Policies, Protocols and Decisions

This section highlights some key policy statements, as well as our evaluation approach with respect to the policies.

1.2.1. Default Net-to-Gross Ratio of 1

Decision 16-08-019 calls for a default net-to-gross ratio (NTGR) of 1.0 to apply to SEM custom projects when program influence is evident. We will evaluate the effectiveness of PAs at documenting program influence, as well as the appropriateness of the 1.0 assumption. Methodologies for evaluating NTGR are described in section 7.2 of this workplan.

1.2.2. Baseline Defined as Existing Condition

Decision 16-08-019 states that for SEM programs in the industrial sector, which can include custom projects as well as BRO efforts, the baseline should be existing conditions. We will evaluate the appropriateness of this requirement by investigating the age and condition of replaced equipment, as well as the decision-making process for replacing it.

1.2.3. Review and Evaluation Process

Decision 16-08-019 states that the Commission does not intend for SEM projects to be exempt from custom project review or ex post evaluation. However, the decision clarifies that these activities may be modified based on the uniqueness of SEM projects. Section 6 describes our methods for conducting custom project reviews for a portion of the evaluation sample.

[7] Consortium for Energy Efficiency. *CEE Strategic Energy Management Minimum Elements.* https://library.cee1.org/system/files/library/11283/SEM_Minimum_Elements.pdf

2. SEM Program Brief

Before launching SEM programs in California, Sergio Dias Consulting developed a consistent statewide framework for SEM programs in California. This framework was informed by SEM program designs developed by Energy Trust of Oregon, Bonneville Power Administration (BPA), and the US Department of Energy (DOE). The deliverables from this design effort include the California Industrial SEM Design Guide[8] and the California Industrial SEM M&V Guide.[9] These guides describe best practices and a single statewide program design to each PA's SEM program. Notably, these guides have not been approved by CPUC.

The statewide design calls for a two-year customer engagement, followed by two months for data collection, analysis, and reporting. Savings are claimed after year one and again after year two. Each savings-claim year is referred to as a "reporting period," and each two-year engagement is referred to as an "achievement period" (see Figure 3). The design is structured to allow participants an additional 24-month achievement period, provided the programs are extended. Table 2 summarizes the statewide design activities and timeline. Each PA has the flexibility to vary from the standard design.

[8] Sergio Dias Consulting LLC. (2017) *California Industrial SEM Design Guide.*

[9] Sergio Dias Consulting LLC & Raven Energy Consulting. (2017) *California Industrial SEM M&V Guide*

Table 2: SEM Activity Timeline per Statewide Design Guide

SEM Month	Workshops	Site-Specific Activities	M&V Activities	Modeling Period
1		#1 Kick-off Meeting and Existing Project Review	#1 Develop Energy Data Collection Plan & Collect Baseline Data	Year 1 Performance Period
2	#1 Starting SEM			
3	#2 Energy Efficiency 101	#2 Energy Map 101 #3 Treasure Hunt 101	#2 Develop Hypothesis Model and Technical Review of model	
4				
5				
6	#3 Tracking: Performance 101			
7			#3 Collect Mid-Year Data and Opportunity Register	
8	#4 Employee Engagement 101	#4 Employee Engagement		
9			#4 Model Trial and Participant Review of Model Trial	
10		#5 Energy Management System Assessment	#5 Model Application and Finalization	
11	#5 Making It Stick			
12				
13			#6 Collect Year 1 Data and Opportunity Register	Year 2 Performance Period
14		#6 Completion Report #1	#7 Year 1 M&V Report and Technical Review	
15	#6 Energy Efficiency 201	#7 Energy Map 201 #8 Treasure Hunt 201		
16				
17				
18			#8 Year 2 Model Review & Update	
19	#7 Tracking Performance 201	#9 EMIS Planning/ Implementation		
20				
21				
22				
23		#10 Energy Management System Assessment #2		
24	#8 Celebrating Accomplishments		#9 Collect Year 2 Data and Opportunity Register	
25				
26		#11 Completion Report #2	#10 Year 2 M&V Report	

Figure 3 summarizes the baseline, achievement, and reporting period timelines. The preferred baseline for the first achievement period is the 12 months immediately prior to the start of the SEM engagement. If the 12-month baseline hinders the modeler's ability to develop a strong hypothesis model, the guide allows for an extended 24-month baseline period. The second reporting period uses the same baseline as the first. Savings reported for the second reporting period are incremental to those achieved in the first. Should the programs be extended into a second achievement period, the baseline is reset to be based on the 12 months prior to that achievement period. This ensures that all savings claimed are incremental and not cumulative.

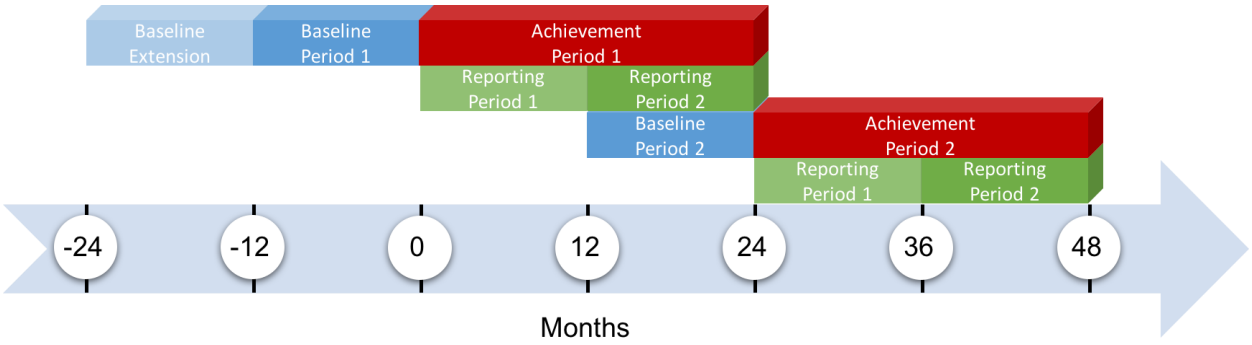


Figure 3: SEM Baseline and Achievement Periods per Statewide M&V Guide

Since the completion of the SEM design and M&V guides, the PAs have issued RFPs, selected implementers, and written implementation plans for their industrial SEM programs. The PAs have recruited 37 participants. The programs launched in 2018 and PAs plan to begin claiming savings in 2019. The PAs' programs are subtly different, but all generally follow the statewide design and M&V guides. Table 3 summarizes each PA's program. Southern California Edison (SCE) and Southern California Gas (SCG) have launched a single, integrated program targeting participants for both gas and electric savings. Also noteworthy is that Pacific Gas and Electric (PG&E) has launched two programs with different implementers.

Table 3: Overview of Industrial SEM Programs

Description	SDGE	PG&E Food	PG&E Mfg	SCE/SCG
Program Name	Industrial SEM	SEM-Food Processing	SEM for Manufacturing Facilities	Industrial SEM
Engagement Start Date	July 1, 2018	July 1, 2018	October 1, 2018	August 1, 2018
Market Sector	Industrial	Food Processing	Manufacturing	Industrial
Participants	7	12	10 (2 cohorts of 5)	8
Milestone Incentives*	$5,000 paid as earned.	$6,000 paid as earned.	$25,000 paid as earned	$5000 paid as earned

Performance Incentives	BRO: no incentive. Custom: traditional programs	BRO: $0.025/kWh; $0.20/therm at months 9 & 26. Custom: traditional programs	BRO: $0.03/kWh (Y1), $0.02/kWh (Y2); $0.40/therm (Y1), $0.20/therm (Y2) Custom: traditional programs	BRO: $0.02/kWh; $0.75/therm paid at months 12 and 24. Custom: traditional programs

* Milestone incentives are paid for accomplishing SEM practices during the engagement, not directly tied to energy savings.

Figure 4 presents a summary of projected two-year program budgets and net annual savings.

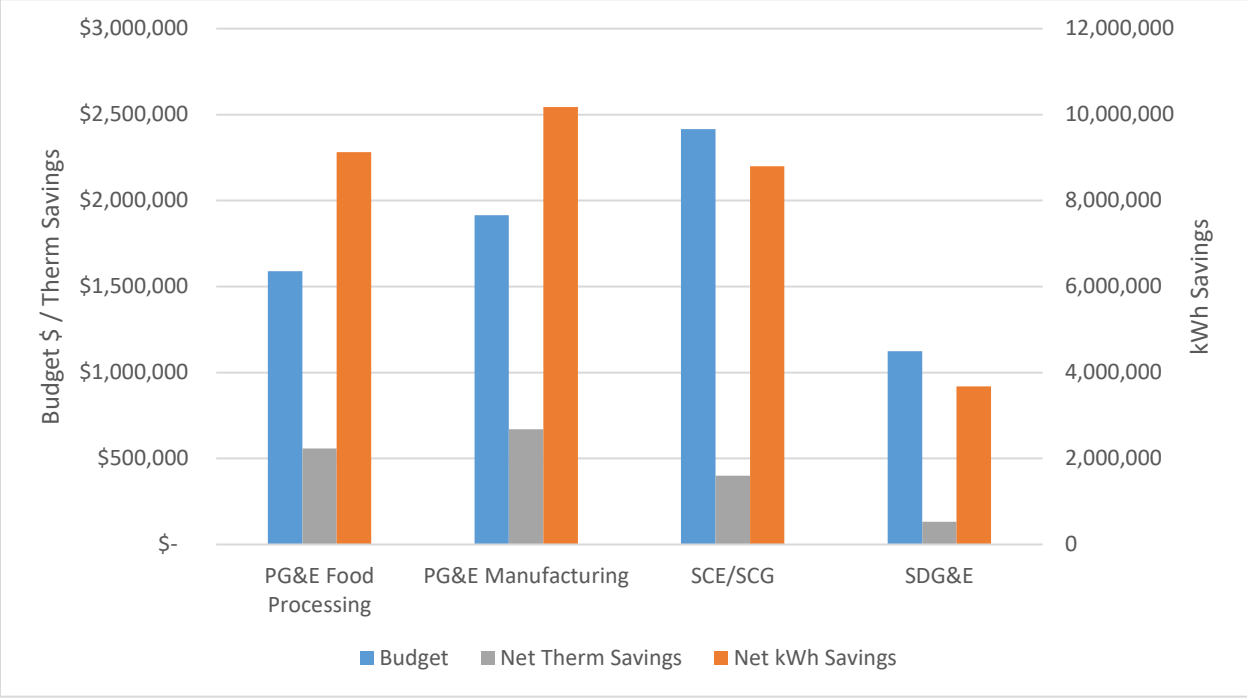


Figure 4: Projected 2-Year Budgets and Savings from PA Implementation Plans

3. Evaluation Objectives

SEM evaluation spans several Group D deliverables, each with unique objectives and approaches. While to some extent SEM is involved in a large number of deliverables, those with significant SEM scope are listed below in sequence of work progression:

- Early Feedback (Deliverable 13)
- Custom Project Review (Deliverable 17)
- Ex Post Gross Savings Evaluation (Deliverable 9)
- Ex Post Net Savings Evaluation (Deliverable 10)

3.1. Early Feedback Objectives (Deliverable 13)

For this deliverable, we will evaluate the effectiveness of all aspects of program design and delivery. We will identify areas for improvement and recommend specific actions. While all programs generally follow the framework identified in the statewide design guide, each of the four programs is unique in design. Additionally, three implementers are implementing four different programs. This gives us a unique opportunity to compare approaches and their associated outcomes in order to identify the most effective practices.

Key objectives of this research will incorporate real-time feedback on the following program components:

- Customer eligibility and document review
- Cohort and participant engagement
- Participant responsiveness and cooperation with implementers
- Workshop and activity quality
- SEM model accuracy and adherence to modeling guidelines
- Engagement with treasure hunts, success and effectiveness
- Applicability, effective technical communication, completeness and support with participant's opportunity registers
- Implementer's interim savings estimates, forecasts, and communications with PAs
- Support for program attribution
- Implementation progress assessment

3.2. Custom Project Review Objectives (Deliverable 17)

We will conduct custom project reviews of all SEM custom projects and of the final models at sites selected for early feedback. The primary objective of the custom project review is to review

the effectiveness of the PA's due diligence review as they vet key factors such as customer and project eligibility and accuracy of energy-savings estimates prior to customers incentives and implementers performance payments. Our custom project review will result in dispositions that summarize our review and feedback for the PAs. Over time, the intent is that the PAs will incorporate project-specific feedback from the dispositions more broadly. Additionally, our dispositions inform the Efficiency Savings and Performance Incentive (ESPI) Custom Project Performance Review component of the ESPI Performance Awards Incentives received by the PAs.

3.3. Objectives of Ex Post Gross Savings Evaluation (Deliverable 9)

The objective of ex post gross evaluation is to determine the realized gross energy savings for the SEM programs. We will estimate total savings and the portion of total savings associated with SEM capital and SEM BRO measures, accounting for the effects of non-routine events during the baseline or performance periods. The approach for estimating gross savings based on ex post evaluation will be informed by custom project review activities and findings.

3.4. Objectives of Ex Post Net Savings Evaluation (Deliverable 10)

Commission Decision 16-08-019 establishes a default NTGR of 1.0 for SEM custom projects when program influence is evident. This evaluation will explore the appropriateness of this NTGR and whether this default NTGR should be continued or modified. The objective of this effort will be to estimate an NTGR for SEM capital and SEM BRO measures.

4. Sample Design

We acknowledge that sampling for several activities from a small population is challenging. However, given the constraints on the evaluation budget, a sample-based approach will allow us to dive deeper into a small number of participants. We expect that this will yield more valuable results than a less rigorous evaluation of all participants.

We selected the evaluation sample as follows:

- We asked the PAs to estimate (in broad ranges) the net kWh and therm savings likely to be achieved for each of the participants. We converted these savings to MMBtu equivalent (site energy) and combined them into an estimate of total net energy savings expressed as MMBtu.
- We randomly separated the population of 37 participants into two groups of 18 and 19 participants: an early-feedback group and a no-early-feedback group. We then divided each of these groups into three strata based on total net savings, using the Delenius-Hodges method.[10]
- We selected a stratified random sample of 15 participants from the early-feedback group to receive early feedback. These 15 participants will also receive custom project review and ex post evaluation. We similarly selected a stratified random sample of 15 participants from the no-early-feedback group to receive ex post evaluation, for a total of 30 participants selected to receive ex-post evaluation. The sample was optimally allocated among the three strata in each group using the Neyman Method.[11]

Figure 5 depicts the population and sample design. To ensure that all participants receive consistent treatment from the PAs and their implementers, we recommend not disclosing the sample results to the PAs.

[10] Cochran, W. G. (1977), Sampling Techniques (Third Edition), Section 5A.7, John Wiley and Sons, New York.

[11] Cochran, W. G. (1977), Sampling Techniques (Third Edition), Section 5A.7, John Wiley and Sons, New York.

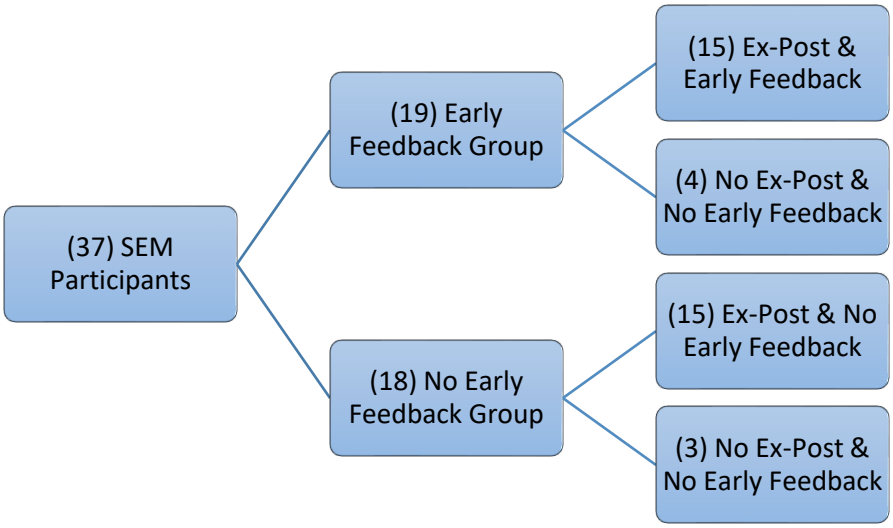


Figure 5: Sample Design Illustration

Table 4 summarizes our sample, including an estimate of its relative precision. Note the high level of confidence and precision despite the small population of 37 participants.

Table 4: Sample Summary

Early Feedback	Participants	Sample	Total Net MMBtu	Sample Estimate of Total Net MMBtu	Confidence Level	Relative Precision
No	18	15	116,348	119,095	90%	2.6%
Yes	19	15	141,565	139,505	90%	1.6%
All	37	30	257,913	258,600	90%	1.4%

This sample will be used several ways:

- Participants who receive early feedback will also be selected for custom project review.
- All custom projects at early-feedback sites are also selected for custom project review.
- All 30 sampled sites will receive ex post evaluation.
- All custom and deemed projects completed during baseline or reporting periods will be evaluated during ex post.
- We will evaluate net savings for all 30 sampled sites as part of the ex post evaluation.
- We will use the remaining seven sites not sampled for a pre-test to refine the NTG survey questions and evaluation approach.

We acknowledge that elements of our early feedback may cause PAs to improve their program design, quality control, and delivery for any participant. This could affect the no-early-feedback group, hindering our ability to measure the impact of our early feedback. However, a primary goal of early feedback is to improve the programs. As such, if our involvement increases

effectiveness of the programs and influences the no-early-feedback group, we view this as a success, not a bias. It is important to note that the objective of measuring the impact of early feedback is secondary to the original Group D SEM objectives, which were: to provide early feedback to help improve program delivery, and to estimate the ex post savings from the program.

If we see large differences between the two domains, this may indicate that our early feedback was effective and that either it was not disseminated to the no-early-feedback group, or the feedback was plant-specific and not generalizable (e.g., improving the operating sequence for a piece of equipment unique to a plant).

If there is virtually no difference between the two groups, we will not know whether this indicates no effect of early feedback or if it was effective but transferred to the no-early-feedback group. This is a downside to the design, but no other design would fix this problem, and quantifying the effect of early feedback is not the primary purpose of the study. To gather additional evidence on the effect of early feedback, we will also interview implementers to determine the extent to which they applied lessons learned through early feedback to the no-early-feedback group. Whether the differences are large, small, or non-existent, the two-domain sample design will achieve its primary purpose to reduce the chance of bias being introduced in our estimate of ex post savings.

To reduce the risk that the program will treat the seven unsampled participants differently, we plan to hide the identity of the untreated (no-early-feedback) sample from the PAs and implementers. This should be possible as our data request is for information about all 37 plants and we can use the seven unsampled plants as the group for a pre-test of our NTGR survey, which should make it difficult for anyone to determine which untreated plants are in the ex post savings sample. Certainly, there is some risk that the seven unsampled would be treated differently by the program, but we believe that our sampling plan offers the best hope for minimizing bias in the ex post savings estimate for the program.

5. Early Feedback Methods

This section describes our approach to early feedback.

5.1. Communication Protocol

It is important to communicate deliberately in order to avoid disrupting or delaying program implementation. This section lays out our communication protocol.

5.1.1. Communicating Findings and Recommendations

We will report findings and recommendations to the CPUC and PAs, but not to implementers or participants. This will help us to minimize confusion amongst participants, and to avoid interfering with the contractual relationships between the PAs and their implementers.

5.1.2. Communication with Participants

We request that the PAs notify participants selected for early feedback that they will be hearing from evaluators. We will notify PAs prior to contacting participants. Experience shows that this advance notification from PAs or implementers enhances participant responsiveness in the evaluation recruiting process.

5.1.3. Communication with Program Implementers

Our communication with program implementers will be limited. Data requests and primary communications will be through the PAs. We expect that the PAs will often relay requests to the program implementers, which may lead to some follow up communication between evaluation staff and implementation staff. Additionally, we will interview program implementers as part of the evaluation process.

5.2. Participant Level Early Feedback

Our activities with participants during early feedback will focus on confirming that key activities and documents meet their design intents and contribute toward progressing each site toward the objectives of the SEM programs.

5.2.1. Scoping Reports

Scoping reports enable the SEM program administrators and evaluators to understand the conditions that existed at each site before the SEM engagements. The California SEM Design Guide outlines ten sections for scoping reports. Our evaluation will confirm that each scoping report addresses each of the ten areas.

The California SEM Design Guide dictates that the SEM coach is responsible for developing the scoping reports. However, much of the information required in the scoping reports pertains to the status of the site before the SEM engagement. Therefore, the SEM coaches must get this information from site contacts or the utility account executive. We will pay attention to the amount of detail provided by the SEM coach to ensure that the conditions of the site before the SEM engagements are sufficiently documented.

If during our evaluation (after our initial review) we discover pertinent details that were not mentioned in the scoping reports, we will note any omissions and potential benefits if the information had been provided in the scoping report. This could include, but is not limited to, changes to production that may trigger non-routine events, other energy-efficiency projects being planned or implemented, and capital upgrades not considered energy projects but that affect energy use.

5.2.2. Opportunity Registers

Opportunity registers chronicle the opportunities identified in treasure hunts and by other means, and help participants track and prioritize the implementation of measures based on these opportunities. In addition, they provide evidence of the influence of the SEM program on savings.

During our review of the opportunity registers we will evaluate the following:

- Do the opportunity registers contain all the sections required by the statewide SEM Design Guide?
- Are the listed opportunities technically valid (will they save energy and impact the facility negatively)?
- Are the listed opportunities specific and actionable?
- Do the listed opportunities include other DSM opportunities such as demand response or combined heat and power?
- Do the identified opportunities include energy management system opportunities?
- Is the opportunity register designed to be user-friendly?
- Is the opportunity register being updated and used to create accountability by assigning responsibilities and due dates and tracking progress of implementing measures?
- Have persistence strategies been documented for each completed activity?
- Do opportunities span across multiple energy-using systems at the site?
- Because most SEM models for industrial facilities are production-normalized, improvements in process efficiency (increasing throughput) and direct-energy-reduction projects will both yield SEM savings. Therefore, we will also review the opportunity register to see if both process improvement and opportunities for direct energy reduction have been documented.

- Have all custom capital projects been noted on the opportunity register?

5.2.3. Hypothesis Model Review

We will review the hypothesis model for all early-feedback sites once in the first reporting period and again in the second reporting period.

Our review of the hypothesis model will evaluate the independent variables tested, model statistics, residual trends, sign and magnitude of coefficients, measurement boundary, documentation and rationale for excluded observations, and identification and treatment of non-routine events. If the model falls short of the M&V Guide requirements, we will explore alternative models, which may include additional or alternative independent variables, using an extended 24-month baseline period or adjustments for non-routine events. We have developed a model-review template to ensure an organized, consistent review among reviewers. This template documents inputs, outputs, regression equations, and other key information. Since implementers use a variety of regression software tools, the templates allow for repeatable results regardless of regression software. See Appendix Appendix B: for an overview of this template.

Our technical approach to reviewing models is:

- Review the data flow through the model to identify input sections, analytics sections, and output sections.
- Independent Variables
  - Review independent variable charts. Based on our understanding of the site, do the relationships seem reasonable? During the baseline or reporting period, are there any outlier data points greater than three standard deviations from the mean? If so, we will investigate how those outlier data points are handled in the energy model tool. If points are excluded, we will ensure that sufficient justification has been provided.
  - Are the independent variables reasonable for the type of facility? We want to ensure that the predicted-energy-use equation is not only statistically valid but also physically reasonable. For example, if a facility with large refrigeration loads uses evaporative condensers, does the model use wet-bulb instead of dry-bulb temperature as an independent variable? If not, and it is not discussed, we will likely try modeling with wet-bulb temperature to see if it improves model fit.
- Modeling Approach
  - Appropriate model resolution. If fine-resolution data has been aggregated into a weekly or a monthly model, why? This may be required to improve the model or may be limited by the granularity with which the site tracks independent variables such as production. However, models should be as granular as possible to maximize their ability to predict energy use and provide feedback on energy performance.

- Are separate models created for different physical aspects of the plant's operation or facility, such as weekend and weekday models or peak and off-peak season? Our review will consider whether the modeling approach and rationale is justified or if alternative approaches may result in a better model.
- If a monthly model is used, did the implementer use a weighted-regression approach?

- Baseline Energy Use
  - Review energy-use data for each affected gas and electric meter and for any model inputs that combine use for more than one meter. Does the energy-use profile align with our expectations of how the site uses energy? Does the baseline annual energy use align with our understanding of how much energy the site uses?
  - Review the scatter plots of baseline estimated and actual energy use to ensure that the data follows a linear trend. If concentrations of data are not collinear, we may consider splitting the model into two separate models with a differentiator such as weekend and weekday models, high- and low-production models, etc.
  - Review time-series plots of baseline estimated and actual energy use. Is the estimated energy use adequately capturing variations in energy use across the range of operating conditions? Do any included periods not reflect typical operation? Should those non-typical periods have been excluded? Are any periods excluded and if so, is there a defensible justification for doing so?
  - Review the baseline cumulative sum of differences (CUSUM) for significant trends. Have those trends been investigated and explained in a baseline report?
- Statistics
  - Confirm that all statistics that are required to be reported are being reported in the model. Compare those statistics with the recommended values established in the M&V Guide, which are summarized in Table 5.
- Instructions
  - Are instructions for how to populate the energy model provided to the participant? The instructions should be very clear to minimize incorrect data.

Table 5: Recommended Statistical Indicators from Statewide M&V guide

Statistical Tests	Statistical Test Threshold Values
Model R^2	> 0.75
F-test overall model p-value	< 0.10
At least one relevant variable p-value	< 0.10
All relevant variables p-value	< 0.20
Net Determination Bias	< 0.005%
Coefficient of Variation	< 20% for daily models
	< 10% for weekly models

Statistical Tests	Statistical Test Threshold Values
	< 5% for monthly models

5.2.4. Capital (Custom and Deemed) Project Evaluation

The SEM model review described in section 5.2.3 will determine total energy savings at the facility, though this may include capital projects funded separately through custom and deemed incentives. To isolate the SEM BRO savings, we must determine, to a high level of certainty, the savings resulting from any capital projects implemented during the baseline and reporting periods. Additionally, when capital projects have been completed, our evaluation will determine whether they were planned before SEM participation or initiated because of SEM participation.

We will evaluate all custom and deemed projects completed at SEM sites during the SEM baseline or reporting periods. The projects will be evaluated to determine actual realized gross savings so that they can be netted out of the model accurately. This will include visiting the site, collecting data, and, where necessary, accurately quantifying savings and M&V of custom and deemed projects.

5.2.5. Site Visits

We will visit sites at or near the end of the first and second reporting periods. This will allow us to:

- Gain familiarity with the facility and factors that affect energy use (independent variables)
- Interview facility staff
- Attempt to identify any process, product, or facility changes that may trigger the need for a non-routine adjustment
- Assess participants' energy-management infrastructure to determine if they have incorporated energy management into their business practices. This may provide some insight into the likely persistence of savings.
- Evaluate all custom or deemed projects completed during the baseline or reporting period. This will include collecting additional data to gain certainty around realized energy savings. Collecting data may include:
  - Validating nameplate data and performance specifications
  - Collecting trend data
  - Reviewing control setpoints and sequence of operations
  - Interviewing facility staff to identify key parameters such as standard operating procedures (SOPs) and schedules
  - Taking spot measurements of key values such as power, temperature, pressure, etc.

  - Logging key data values
- While interviewing facility staff, we will also seek to better understand their level of engagement and any feedback or critiques they have of the program and delivery. This will help inform our findings described in sections 5.3 and 5.4. We'll design a survey in advance to help ensure strong and consistent interview data.

5.2.6. Formulation of Participant-Level Recommendations

We will compile our findings from the activities described in section 5.2 to recommend ways to improve effectiveness in delivering participant-level activities.

5.3. PA-Level Early Feedback

This section describes our approach to reviewing PA-level delivery and documents. This includes aspects of delivery that are not site-specific, such as delivering workshop, curriculum, and program design per implementation plans.

5.3.1. Reporting and Workshop Delivery

We will review the workshop materials, meetings, and evidence of customer engagement. We will compare our personal observations with the results and outcomes documented in the SEM participant tracking reports and SEM workshop summary reports. The minimum elements that we expect in the SEM participant tracking reports include the following:

- General cohort information. This includes key dates for start of engagement, milestones, and reporting periods.
- General information for each facility. This includes company name, location, SEM coach, utility account executive, facility savings goals, annual consumption by fuel type, and key staff member names with SEM team assignments.
- A list of key activities with dates for each facility

The SEM workshop summary reports, at a minimum, will be expected to include:

- General workshop information (name, date, location)
- List of attendees including participants, utilities, implementers, and guests
- Summary of the workshop and material delivered
- Participant feedback including surveys of workshop attendees

5.3.2. Customer Engagement, Effectiveness of Curriculum, and Transformation

A key objective of SEM programs is to influence energy-management culture within a facility. We will attempt to understand the effectiveness of this influence by evaluating the engagement of participants. Aside from participation at workshops, we will compare the level of engagement across programs within the sample group. This will allow our evaluation to compare the levels of engagement and assess what is and is not working for various program features, curriculum, or other factors. Key indicators that we will consider to gauge customer engagement include:

- Frequency of contact between participants and implementers
- Frequency and/or presence of energy-team meetings
- Actions taken towards early implementation of opportunities
- Participation at workshops
- Internal communications amongst participant facility staff to share energy saving goals, results, and ideas for energy savings opportunities
- Meeting program requirements on time and with substance and quality
- Demonstrated support of SEM at facilities by executive management

Participant feedback after workshops will be a crucial means to evaluate the curriculum, delivery, and effectiveness of workshops.

5.3.3. Implementation Plan Review

We will review each PA's implementation plan for adherence to the statewide program design and M&V guide. While reviewing the implementation plans, we will note deviations from the statewide program design and consider the ramifications of the deviations. During our review we will also consider the PA's incentive structure, customer-eligibility criteria, calendar of events, required tools provided to the customers, and other factors that may impact the participants' experiences and outcomes.

This task will also include interviews of each program implementer to seek additional recommendations for program improvements.

After the first year of SEM engagement, we will consider whether the implementation plans accurately communicated the program design and if they were consistent with our observed implementation. We will provide specific feedback as needed.

5.3.4. Formulation of PA Level Recommendations

We will compile our findings from the activities described in section 5.3 to recommend ways to improve effectiveness in program-level aspects of design and delivery. Additionally, through our

participant-level review activities described in section 5.2, we will search for findings that are common among the participants, indicating the need for improving program design or implementation.

5.4. Statewide-Design-Level Early Feedback

This section describes our approach to reviewing the statewide design and M&V guides.

5.4.1. Program Design Guide

The foundation of our review will be to ensure that SEM has been structured to naturally progress through the high-level objectives defined in the Statewide SEM Design Guide, which include:

- Implementing energy-efficiency projects and saving energy
- Establishing the energy management system or business practices that will help a facility manage and continuously improve energy performance
- Normalizing, quantifying, and reporting facility-wide energy performance

To ensure that these high-level objectives are met in an effective and cost-effective manner, our review will include:

- Roles and Responsibilities
  - Are the divisions of labor designed so that the responsible individuals should already have the requisite knowledge and experience to efficiently complete their tasks?
- Timing and Content
  - Are there any missing activities related to the high-level objectives that might diminish participant experiences or undermine the ability to generate SEM savings?
  - Are there any superfluous activities related to the high-level objectives that might distract participants from key activities or add delivery costs with little benefit?
  - Is the proposed timing likely to unnecessarily hinder implementation or will there be other negative consequences from the proposed timeline?
  - Are activities sufficiently defined to differentiate SEM from other demand-side-management initiatives such as retro-commissioning? A key distinction is that successful SEM engagements will encourage participants to learn and implement various energy-management practices, whereas other demand-side-management programs rely more heavily on third-party implementers to identify and complete projects.
- M&V
  - Is enough data being requested and collected early enough to allow hypothesis models to be developed early?

- Does language in the program design guide contain enough M&V detail and contain appropriate references to the other more comprehensive M&V resources such as the industrial SEM M&V guide?

- Incentives
  - Are the milestone incentives motivating participants to work towards the intended outcomes?
  - Have any of the individual PA incentive structures proven successful? If so, we will explore recommendations to a modified incentive structure in the statewide design.
- Tools
  - Are the provided physical tools (e.g., items in the data-logging toolkit) expected to be useful in identifying or implementing energy projects or improving persistence?
  - Have the opportunity register, energy model, and energy map tool been sufficiently defined to ensure that they support the high-level objectives and to ensure participants across different program administrators and third-party implementers will receive a similar experience?
  - While the energy model is often addressed for its role in SEM M&V, it also plays a pivotal role in engaging and motivating participants. Therefore, we will also review language and guidance regarding how the energy-model tool should be designed and delivered from the perspective of employee motivation in order to promote engagement and persistence.
- Reports
  - Do the required reports document the process and results of the SEM engagements? Can reports be streamlined?
  - Is the intended audience clearly communicated for each report and is the information appropriate for the intended audience?
- General
  - Is there guidance on how program implementers should proceed when they encounter situations that are not explicitly detailed in the SEM Design Guide (for example customer inaction)?

5.4.2. M&V Guide

The California Industrial SEM M&V Guide (M&V Guide) draws heavily from existing documents including:

- Bonneville Power Administration's Monitoring Tracking and Reporting Guide
- Energy Trust of Oregon's Energy Intensity Modelling Guideline

- US Department of Energy's Superior Energy Performance Measurement and Verification Protocol

SEM M&V is founded in statistical analysis; however, few SEM M&V practitioners are statisticians. Therefore, the M&V Guide will be an invaluable tool to communicate expectations and establish best practices. In reviewing the M&V Guide we will consider the following:

- Content and Sequence
  - Is the M&V Guide comprehensive enough to lead an M&V practitioner through the process from requesting the initial data to documenting the results?
  - Is the order of information logical and aligned with the sequence of activities that are likely to be completed by the M&V professional?
- Terminology and Definitions
  - Are terms sufficiently defined to minimize confusion?
  - Are the definitions of terms aligned with other industry reference guides?
- Exceptions – An M&V guide should detail how to conduct SEM M&V under typical conditions and provide some direction on handling exceptional cases. Therefore, our review will consider how this M&V guide addresses typical cases but also how it covers exceptional cases to ensure that program administrators and third-party implementers have a guiding resource for unusual situations.
- Energy Model – SEM M&V occurs within each third-party implementer's NMEC models. Our review of the M&V guide will consider the guidance given on modeling. We will consider how the M&V guide leads users through developing and documenting the model for the purpose of final reporting. However, we will also consider how the M&V guide leads users through the SEM M&V process from the perspective of customer engagement. To this end, we will evaluate what data the M&V guide recommends practitioners display in their energy models and in reports.
- Model Selection – One particularly challenging task for SEM M&V practitioners is evaluating competing models. We will therefore review the M&V guide's recommendations for how M&V practitioners should evaluate competing models. Specifically, we will check to see if the M&V guide provides direction on weighing model fit versus simplicity, the robustness of various independent variable types, and the balancing of the need for customer understanding versus model complexity.
- Model Validity
  - We will review the statistical metrics used to test the validity of energy models and compare the statistical metrics to other guiding documents for SEM M&V.
  - SEM models require ongoing validation throughout the implementation of the SEM engagement to ensure that the results of the model align with the expected results based

on known changes or completed energy projects at the site. Our review will also evaluate the language in the M&V guide regarding how a model is tested and finalized.

- Reports
  - Are the required reports sufficient to document the M&V process and savings from the SEM engagements?
  - Is the intended audience clearly communicated for each report, and is the information appropriate for the intended audience?
  - Can the report be streamlined to reduce delivery costs and provide participants a more concise deliverable?

5.4.3. Formulation of Statewide Recommendations

We will compile our findings from the activities described in section 5.4 to recommend ways to improve effectiveness in statewide design and M&V. Additionally, through our participant-level and program-level review, we will search for findings that are common among many participants and PAs, indicating the need to improve statewide design.

6. Custom Project Evaluation Methods

This section describes our custom project review methods. We will subject the final models of all sites that receive early feedback to custom project review. We expect that, for early-feedback sites, PAs do not issue performance-based incentives until our custom project review of the final model is complete, and that incentives issued are consistent with our custom project review. Additionally, all custom projects at sites sampled for early feedback will undergo custom project review, following the traditional custom-project ex-ante-review process. Custom project reviews will yield dispositions that inform the ESPI process.

6.1. SEM Custom Projects Review

We will ascertain whether any sites selected for early feedback will be monitored for the presence of custom projects. All custom projects at early-feedback sites will undergo custom project review. The custom project review of SEM custom projects will follow the standard custom project review process and result in dispositions.

The purpose of the custom project review of an SEM custom project is to review the project eligibility, industry standard practice determination, measure application type, program influence, engineering analysis of energy savings, and PA due-diligence review process. This engineering analysis will review calculations used to estimate the custom-project savings, inputs to the tools, and M&V plans associated with the project. Our disposition may require M&V of the custom project if the energy analysis has sufficient risk or uncertainty.

6.2. SEM BRO Review

This section describes our process for custom project review of SEM BRO projects.

6.2.1. Timing of Custom Project BRO Review

SEM BRO projects are not suitable subjects for the typical custom project custom project review as there are no savings estimates prior to the implementation of the measures in the project. So we need to decide when in the process to conduct custom project review. The two logical options are at the completion of the hypothesis model or at the completion of the final model and M&V report. While the former option is better aligned with the timing of a typical custom project review, there are no savings to review at this point, which significantly reduces the value of the custom project review. Consequently, to allow for a more substantive custom project review, we have chosen to review the final model and M&V report after the conclusion of the performance period.

Our SEM BRO custom project review will happen before the PAs issue performance-based incentives at the end of the first and second reporting periods. We expect that performance-based incentives issued be consistent with our custom project review. Our custom project review

will be conducted within the 30 business days defined by SB1131. It is noteworthy that milestone incentives can be issued without a custom project review.

6.2.2. Approach to Custom Project BRO Model Review

This section describes our approach to custom project review of the final SEM models.

- Review final model. At the time of our custom project model review, we will have reviewed the hypothesis models through early feedback. As such, our first step in final model review will be to review our feedback of the hypothesis model and ensure that our recommendations have been incorporated into the final model.
- Evaluate the model-development documentation submitted by the implementer. This will include reviewing narratives in the M&V report and/or competing model summary tables that illustrate alternative model iterations. This will help us evaluate the modeling process and the testing of various modeling approaches to vet competing models and select the best model.
- Check for alignment of model results with known activities. We will use the opportunity register (and/or opportunity tracker) to align site activities with model results throughout the SEM engagement. Figure 6 shows a sample CUSUM graph from the M&V guide that shows the actions taken over time. We will seek to explain any inflection points or significant changes in slope of the CUSUM.

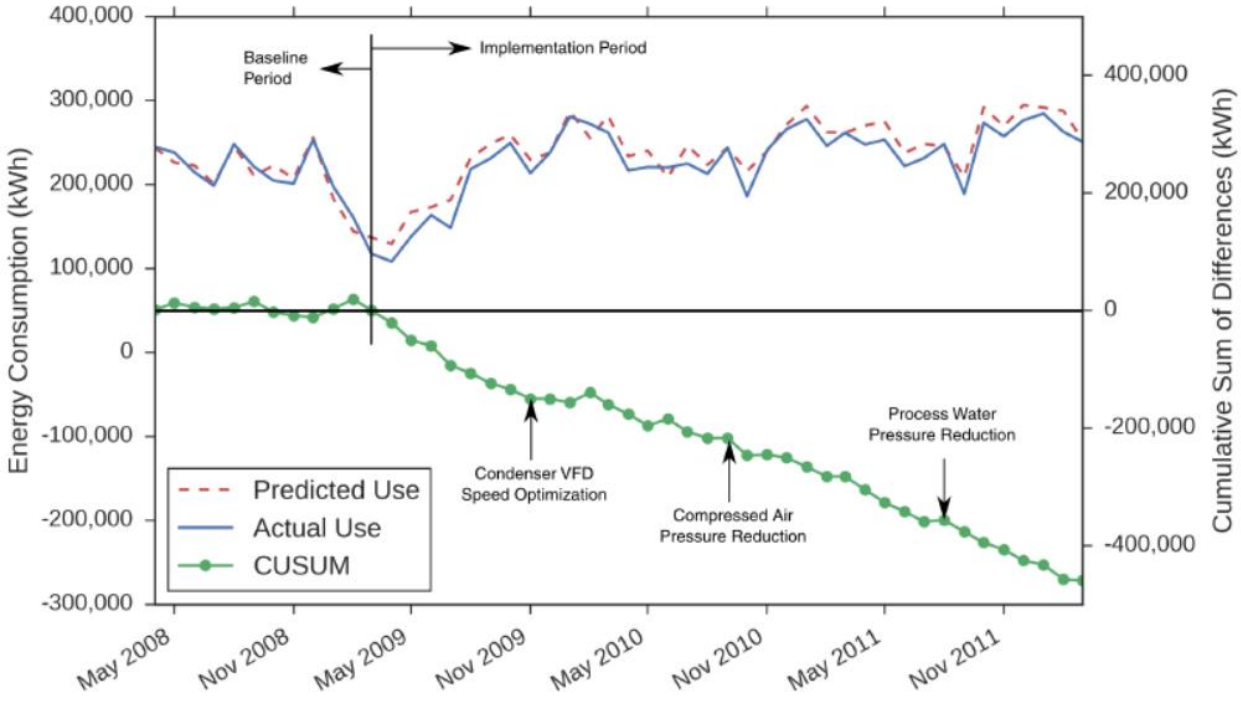


Figure 6: Sample CUSUM Graph with Actions Annotated

- Adjust for non-routine events including custom and deemed projects. We will confirm that all custom projects that were incentivized by the PA during the baseline and performance periods have been properly accounted for using the methods defined in the M&V guide. The savings estimates for custom improvement projects will include any modifications suggested during custom project review (which will be revised using M&V results) and as-built and operated input parameters.
- Review model-validation activities. We will also evaluate steps taken by the implementers to use independent variable data during the baseline and reporting period to validate their models. We will check if independent variable data stays within ±10% of the baseline range to prevent excessive model extrapolation. If energy-driver data extends beyond ±10% of the baseline range, we will review how those observations were handled in the M&V report .

7. Ex Post Evaluation Methods

This section summarizes our ex post gross and net evaluation methods. Additionally, this section describes our approach to analyzing cost effectiveness and our application of the results of sampled site-level evaluations to the population.

7.1. Gross Savings Estimation Methods

This section describes our methods for estimating gross energy savings. For the 15 early feedback sites, we will have just completed these activities through early feedback and custom project review. As such, we will estimate gross savings only on those 15 sites that were not selected for early feedback. Additionally, if our early feedback and custom project review resulted in revisions, we will review those revisions and update our findings as part of our ex post activities. Finally, while we have set the targeted level of confidence and precision at 90/10 for estimates of gross savings, we expect to achieve much better confidence and precision in the range of 90/1.4 to 2.6, as shown in Table 1. This will be important since we will lose some precision when we propagate the precisions for gross and NTGR estimates, and our final target is to estimate net life-cycle savings with 90% confidence and 10% relative precision.

7.1.1. Data Collection and Site Specific M&V Plan

We will begin data collection by requesting complete and updated participant files from the PAs for all SEM participants at the end of each of the two reporting periods. We will request that the PAs identify all deemed and custom projects completed at each site during the baseline and performance periods.

We will review these project files to gain an understanding of the facility, to identify significant actions taken, and to identify any custom or deemed projects or non-routine events.

Based on this file review, we will develop a site-specific evaluation plan. This plan will ensure a productive and efficient evaluation that minimizes the time required by facility staff, and will:

- Identify measures implemented
- Identify missing data or key deliverables
- Identify strategies for improving models through additional data collection and alternative modeling approaches, such a better understanding of the facility, its independent variables, and non-routine events
- Plan site visit activities to ensure that we meet all of our evaluation needs in a single visit
- Draft guidelines for interviewing facility staff – This will be coordinated with our NTG surveys to consolidate activities when possible and minimize survey time required by participants. Additionally, we will coordinate participant contacts, seeking to identify

appropriate decision makers (most relevant to NTG interviews) and technical staff (most relevant to gross interviews).

- Identify an M&V approach to any custom or deemed projects completed at the facility during the baseline or reporting periods

We will begin collecting data at the end of the first reporting period, so data collection will take place over a five-month period from August through December 2019, as each program concludes its first reporting period.

7.1.2. Communication Protocol

To avoid participant fatigue, we will seek to minimize contact with participants. We will first attempt to gather additional data from PAs and implementers before contacting participants. We will also ask the PAs to notify participants that they should expect to be contacted by us, which should pave the way for a more successful recruitment.

7.1.3. SEM BRO Model Reviews

Our ex post model review will follow the same process as our custom project review described in section 6.2.2. For sites that have not already been visited during early feedback, our ex post evaluation will include visiting the site, interviewing facility staff, collecting on-site data, and evaluating all custom and deemed projects completed during the baseline and reporting periods. Since this will have just been conducted on the 15 early feedback sites, we will use those findings for ex post results. If, however, our early feedback and custom project review recommendations result in model updates, we will include those updates in our ex post review.

7.1.4. Treatment of Custom, Deemed, and Non-Incentivized Projects in SEM Models

The treatment of custom, deemed, and non-incentivized projects that are not attributable to SEM in the models will be critical in determining the impact of SEM, net of other activities. For all sites sampled for ex post evaluation, we will evaluate all custom and deemed projects completed during the baseline and reporting periods. We will net out the evaluated gross savings of each project from the SEM model to estimate SEM-BRO savings. Additionally, we will review the scoping reports and interview facility staff to identify any projects implemented that were not a result of SEM and were not incentivized by the PA. For those projects, we will evaluate the resulting energy savings and net those out of the SEM models.

It is important to note that the savings netted out of the model may be different from those that are appropriate to claim through the custom or deemed project. For example, if an SEM participant replaces a chiller as an end-of-life replacement, we recommend that the program should claim savings incremental to Title 24. However, the savings that would show up at the meter (and SEM model) would not be incremental; it would be the full savings relative to the

old chiller. So, as we evaluate custom and deemed projects at SEM sites, we will take care to consider the baseline. Where the PA, or our evaluation, uses an end-of-life baseline, we will calculate an alternative baseline as if it were accelerated-replacement for the purposes of netting those savings out of the SEM model. When we encounter this need for two baselines we will document our methodology.

7.1.5. Estimating SEM Custom and SEM BRO Savings

We will quantify the savings of SEM custom and deemed measures using methods that comply with the International Performance Measurement and Verification Protocol (IPMVP) Options A or B M&V techniques.

We will review SEM BRO models as described in this section, and re-model as needed to improve model accuracy and account for any non-routine events. This option is consistent with IPMVP Option C.

Last, we will calculate the SEM BRO savings as the NMEC modeled savings adjusted for non-routine events and net of SEM custom and deemed savings.

Thus, SEM BRO savings will be derived as follows:

SEM BRO Savings = [*SEM NMEC Modeled Savings*] – [*SEM Custom Savings*] – [*Deemed Savings*]

When no custom or deemed projects have been completed, we will quantify all savings using only NMEC models. This will let us quantify and report uncertainty, some level of which will be present in all NMEC models. Statistical models allow us to quantify the uncertainty of savings, while custom-project M&V does not. Therefore, if custom or deemed projects have been implemented during the baseline or reporting periods, we will be unable to calculate and report uncertainty.

7.1.6. Treatment of Negative Savings

All NMEC models have a level of uncertainty that will inevitably either under- or over-report savings at a given site. Through refined modeling techniques we seek to minimize that uncertainty. Ideally, the model uncertainty will not be biased; that is, portfolio-wide claims will have reduced uncertainty as one site's overreported savings will be balanced by another site's underreported savings.

However, if sites with negative savings are reported as having zero savings rather than the actual negative savings, this would introduce a bias that ignores the underreported negative savings, while including the unintentionally overreported savings at other sites.

Before reporting negative savings, we will work with our modelers and facility staff to understand the cause of negative savings and account for it in the model, likely through a non-routine adjustment. However, despite our best efforts to explain negative savings, we expect that the best model available for some facilities will show negative savings. To minimize bias and

maintain strong portfolio-wide accounting, we recommend that our evaluation report negative savings when the model shows it.

7.2. Net Savings Evaluation

Since SEM programs are new to California, DEER does not include an NTGR for SEM. Given this lack of DEER value, Decision 16-08-019 calls for a default net-to-gross ratio (NTGR) of 1.0 to apply to SEM custom projects when program influence is evident. Our research will evaluate the validity of this assumption.

Given the size and complexity of the SEM projects, we will base our approach on the most rigorous set of NTGR methods, the standard-very large version of NTGR methods.[12] Because the SEM program promotes both capital measures and BRO actions, we will modify the wording of the questions to distinguish between these two categories of energy-reduction actions.

One important question is when to interview participants. Since it might require some time for recommended energy-efficiency measures to be approved and installed, or for many of the BROs to be routinized, we propose that interviews be conducted at the end of the first reporting period for each participant site. At that time, participants should be able to recall the details of their decisions and tell us in their own words regarding the development of various capital and BRO actions. We will closely coordinate all NTG interviews with the engineers conducting the ex post gross-impact evaluation to eliminate redundancy in questions and minimize the evaluation time required by site staff.

Because there is typically no single indicator of program influence in complex projects, we will rely on a preponderance-of-evidence approach based on multiple metrics. Using a rules-based approach, we will combine this data into an internally consistent, coherent story about program attribution. Finally, we will develop best-practice protocols for estimating program influence that other evaluators can use going forward.

Though we have yet to determine exactly how we will combine this data, the following data will be included:

1. **Program Files**. Program files provided by IOUs can contain information that is relevant when analyzing free-ridership. For example, letters written by the utility's customer representatives can indicate what the customer had planned to do in the absence of the program and explain the customer's motivation for taking the energy-efficiency actions.
2. **Decision-Maker Surveys**. When a site is recruited, we will determine who was involved in the decision-making process that led to taking the energy-efficiency actions under the program. We will ask these decision makers to complete a survey that includes questions

[12] The Nonresidential Net-To-Gross Ratio Working Group (Dr. Ridge, Lead). (2012). *Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers*. Prepared for the Energy Division of the CPUC.

about the role the implementer played in helping the participant to identify and implement energy-savings opportunities, what proportion of the items on the opportunity register were generated by the implementer and by the participant, and in what combination (not to assign influence based on whose idea it was, but rather to contribute to the program description). The survey will include highly structured questions and responses concerning the probability that the customer would have identified and taken the same actions at the same time in the absence of the SEM program. Survey questions will be designed and tested to ensure they do not include leading questions. First, we will ask participants about the timing of their program awareness relative to their identification of and decision to take the energy-efficiency actions. Next, we will ask them to rate the relative importance of program and non-program influences in their decision making. Third, we will ask them to rate the significance of various factors and events that may have led to their decisions to take the actions at the time that they did. These include, among others:

- Information from the treasure hunt and opportunity register
- The age or condition of the equipment
- Information from a feasibility study or facility audit
- The availability of an incentive or endorsement through the program
- A recommendation from an auditor or consulting engineer
- Previous experience with a similar program
- Information from a program-sponsored training course or marketing materials provided by the program
- A suggestion from IOU program staff, a program implementer, or a utility account representative
- A standard business practice
- An internal business procedure or policy
- Stated concerns about global warming or the environment
- A stated desire to achieve energy independence

The survey will contain a core set of questions and several supplemental questions. For example, if a corporate policy was a primary consideration in their decision, we will ask them a series of questions about the specific policy that led to their taking the energy-efficiency actions. If they indicate that the actions were a standard practice, supplemental questions will address the origin and evolution of that standard practice within their organization. These questions are intended to provide a deeper understanding of the decision-making process and the relative influence of the program versus these internal policies and procedures. Responses to these questions also help gauge consistency when investigating conflicting answers regarding the relative influence of the program compared to the non-program elements. These sites may also receive additional scrutiny into various

aspects of their decision to modify the operation of their facility based on industry-specific issues and other information sources. All this data will be used to construct an internally consistent "story" that supports the calculated NTGR.

3. **Utility and Program Staff Interviews**. We will also interview utility staff and program staff, including account representatives. These interviews are designed to gather information on the historical background of the customer's decision to change the operation and maintenance of their facility or purchase energy-efficient equipment, and the role of the utility and program staff in those decisions.
4. **Implementers**. We will conduct phone interviews with the implementers to determine how this utility program fits into their general business model. For example, would the implementers be targeting the same customers and making the same recommendations even without the SEM incentives? Or does their business model depend on utility subsidies and/or incentives?
5. **Other information**. For these standard-very large-project NTGR sites, we will perform secondary research of other pertinent data sources. For example, this could include reviewing secondary sources such as the U.S. Energy Information Administration (https://www.eia.gov/todayinenergy/detail.php?id=33392). In addition, the standard-very large NTGR analysis might involve interviews with other employees at the participant's firm, sometimes in other states, to provide further insights regarding standard practices within each company regarding capital and BRO actions.

7.2.1. Sampling Actions Within Site

Historically, the self-report method of the standard-very large NTGR has been used for programs in which multiple strategies were used to minimize the number of measures the respondent is asked about, while still achieving a reasonable level of confidence and precision at the site level.[13] When participants installed more than three measures, evaluators drew a random sample of three of the total measures installed, or they deliberately selected those projects with the largest savings so that the NTGR addressed the program influence on the largest portion of the savings, even though the results were not necessarily generalizable to the population of measures at a given site. In other cases, the measures were sufficiently homogeneous (e.g., LEDs) and therefore could be grouped for the purpose of assessing attribution; thus effectively keeping the number of measures manageable. Sometimes, measures with very small savings were dropped from the sample frame since it would not have been cost-effective to estimate the influence of the program on the decision to install them. In still other cases, even though the measures were not homogeneous, the influence of the program on the decision to install all three was essentially the same, and the NTGR battery only had to be asked once. These strategies kept the sampling fraction sufficiently large so that adequate levels of

[13] As examples, see: Itron. SBW Consulting et al. (2010). *Major Commercial Contract Group: Volume 1: Final Impact Evaluation Report: 2006-2008 Program Years*; Itron. (2017*). 2015 Custom Impact Evaluation: Industrial, Agricultural, and Large Commercial*.

confidence and precision at the site level were achieved mostly without overburdening the participant. However, the SEM program differs from the typical nonresidential program in that it attempts to achieve savings in the industrial sector with highly heterogeneous measures and thus requires a different sampling approach.

The most important difference from custom programs is that the actions that fall into the SEM category are many and varied, including BRO measures. There are so many actions and so much variation that choosing the actions to ask NTGR questions about is complex. Asking about each one would be much too burdensome on the interviewee. In our experience, we can ask questions about three or four actions. However, any four actions, chosen randomly, might not be representative of the overall influence of the program. We considered asking about *groups* of actions, such as all BRO measures and all capital measures, but this would likely pose difficult choices for the interviewee about how to characterize the influences on the entire group. We needed to build a process for categorizing and selecting actions based on aspects of the actions that were most likely to have been influenced by the program. For example, capital measures are likely to be subject to different decision-making participants and processes than O&M measures, and thus the program is likely to have a different point of influence. This makes the capital-versus-BRO distinction an important one for sampling. At the same time, it would be a mistake not to represent the interventions expected to yield the majority of savings.

The rest of this section describes the categories and processes we will use to select actions listed in the opportunity register as the basis for questions in NTGR interviews.

Before using the opportunity register for sampling, we will group the actions in several ways. First, we will filter the items in the opportunity register for those that have been completed or begun by the end of the first-year performance period, and only consider those for interviews. This means that some actions taken at the end of that period will be the subject of NTGR questioning but will not be reflected in the gross-savings estimate for the first performance period. These gross savings will be reflected in the second-year performance period.

Second, we will consider all the strategies mentioned earlier for reducing the population of measures from which a sample must be drawn at a given site. In addition, all measures that clearly pertain to the same piece of equipment or process will be grouped. In some cases, actions listed in the opportunity register are distinct from others on the list, while other items in the register could be called micro-actions taken on the same piece of equipment or process. We would not want to treat each micro-action on its own; they should be considered as a unified group that reduces the energy use of the equipment or process.

Third, it will be important to represent both BRO and capital measures, since that distinction is likely to differentiate decision processes and influences. Thus, we will stratify actions or interventions into those two categories and sample within them.

Fourth, within the two categories of BRO and capital, we will further stratify actions into high- and low-savings actions. For each site, we plan to choose four interventions, divided across BRO versus capital and high versus low-savings potential. How they are allocated across the

four groups for the collection of the 30 sampled sites will depend on the distribution of actions within and across the groups.

This sample design attempts to balance three objectives:

1. achieve a representative sample,
2. minimize non-response and measurement error by not overburdening the respondents, and
3. stay within the evaluation budget.

The achieved level of confidence and precision for the estimated NTGR will be calculated in a way that is consistent with our two-stage cluster sample design[14] (i.e., sampling 30 of the 37 sites and sampling four actions within the sampled sites). For the first stage, we assumed a coefficient of variation (CV) of 0.35 which was the CV observed in the evaluation of the 2017 Custom program. For the second stage, we assumed an even lower CV due to the expected low variation for each site within the capital and BRO strata. Given these two assumptions, the expected level of achieved confidence and precision for the NTGR is approximately 90/8. This, in combination with the expected 90/1.4 level of confidence and precision for the estimate of gross savings, should, using the propagation of errors, result in an achieved level of confidence and precision well within the 90/10 level for *net program-level life-cycle savings*.

Finally, two remaining unknowns will be resolved before data collection begins. First, while we do not currently know, at the site level, the number of sampled projects as a fraction of the population of projects, or the savings for the sampled sites as a fraction of total savings, it is possible that these fractions could be rather small. As the performance period ends for each of the 30 sites, we will be able to calculate these fractions within each site and assess the representativeness of our site-level sample of four. For those sites with small fractions and, as a result, large confidence intervals, we will consider offering an incentive to the respondent to increase the number of actions covered in the survey.

Second, we are somewhat uncertain whether respondents will be able to think clearly about the influence of the program for each of the two sampled capital measures and each of the two sampled BROs. We will explore their ability to do so as part of the two pretest interviews discussed later in section 8.2.4. We might discover that they cannot assign a unique level of program influence to each of the four, but that *are* able to think separately about program influence for the group of capital measures and the group of BROs. If that occurs, we will revise the questionnaire to ask about program influence for each of two groups: capital measures and BROs.

[14] Valiant, Richard, Jill A. Dever, and Frauke Kreuter. (2013). Practical Tools for Designing and Weighting Survey Samples. New York: Springer.

7.2.2. Program Theory and Logic Model

To further guide our thinking about what data to collect and to help identify causal mechanisms for this complex program, we will initially rely on a program-theory-and-logic model.[15] At a high level, logic models describe inputs and activities and how they combine to produce expected outputs which, in turn, may produce expected short-term, mid-term, and long-term outcomes. Each pathway or linkage in the logic model describes a hypothesized cause-and-effect relationship. The key elements of any logic model are the inputs, activities, outputs, and outcomes. *Inputs* can include human, financial, organizational, community, or systems resources in any combination. These inputs serve as the catalyst for program *activities*, which reflect the processes, events, technologies, and other devices that are intentional in the program. The direct results of program activities are the *outputs,* which can include the production or availability of different types of program assistance. *Outcomes* are about change, and indicate what ideally will occur because of the program activities and outputs. Common outcomes include specific changes in awareness, knowledge, skills, behavior (including capital purchase behavior), and energy savings. Outcomes are often categorized by time to indicate which ones are expected in the short, intermediate, and long term.

Each program implementer provided a logic model as part of their plan. Appendix C shows a sample logic model. We will review each model and might, *in collaboration with the implementer*, modify it. Then we will focus on why we expect this set of SEM strategies to change individual and organizational behavior in a way that will lead to the desired outcomes, including sustainable reductions in energy use. Since the implementers said little about the underlying individual and organizational theories that were implicit in their logic models, we will conduct a brief review of the organizational changes to identify the theory underlying each key linkage (a hypothesized cause and effect relationship) in the logic model, which can be tested using data collected by the SBW team. We will incorporate the results of this analysis into the estimation of the NTGR discussed in section 7.2.6.

7.2.3. Participant Spillover

As noted above, SEM can be defined as a holistic approach to managing energy use to continuously improve energy performance by achieving persistent energy and cost savings over the long term. It does this by focusing on changing business practices from senior management through shop-floor staff to affect organizational culture, reduce energy waste, and improve energy intensity. Given this, the goal of SEM is to identify all conceivable equipment changes and BROs that could cost-effectively reduce energy use. As a result, spillover is expected to be minimal since few energy-saving actions would remain to be taken by any participants.

[15] Rogers, Patricia J., Timothy A. Hacsi, Anthony Petrosino, and Tracy A. Huebner (Eds.) (2000). *Program theory in evaluation: Challenges and opportunities*. San Francisco, CA: Jossey-Bass Publishers; Frechtling, Joy A. (2007). *Logic Modelling Methods in Program Evaluation*. San Francisco, CA: Jossey-Bass

7.2.4. Questionnaire Development

Before finalizing the NTGR methods and questionnaire, we will draw on lessons learned from the most recent evaluations of custom programs in California, as well as SEM programs in other regions.[16] Then we will develop a draft questionnaire to reflect the unique characteristics of the industrial SEM program. We will then conduct pretest interviews[17] with two of the seven participants who were not included in the evaluation sample.[18] The primary objective of these pretest interviews is to minimize measurement error by improving the clarity of the survey questions, particularly for counterfactual questions, so that they are correctly interpreted by those survey respondents who take the final version of the survey. To our knowledge, such a comprehensive investigation of the nonresidential survey instrument has never been done, and we expect it to yield some important improvements. We will then administer the draft final questionnaire that emerges from these pretest interviews to the remaining four non-fully-evaluated participants to ensure that the survey length is not onerous, that the question length and order are correct, that the wording is clear, and that skip patterns and consistency checks are functioning correctly.

7.2.5. Data Collection

We will begin collecting data at the end of the first reporting period. Interviewing participants at this time is essential since participant recall is one of the biggest threats to validity and reliability of the NTGR estimate. The longer one waits to interview a participant the less likely they are to recall the important details and motivations of their decisions and actions taken as a result of the program. While we expect to interview all 30 participants, we will use several strategies to improve response rates. We will work closely with utility account representatives to gain the cooperation of the participants and contact each prospective respondent up to five times at different times of the day until we receive a firm refusal. We will also send multiple reminder emails to each prospective respondent at different times of the day and week. To identify non-response bias, we will compare survey respondents to the population based on data which is available in all of the project files such as business type, number of projects, and project savings.

---

[16] For example: Itron, Opinion Dynamics, and Michaels Energy. (2016). *ComEd and Nicor Gas Strategic Energy Management (SEM) Evaluation Report*. Prepared for the Commonwealth Edison Company; NEEP. (2017). *Evaluation Measurement & Verification (EM&V) Best Practices & Recommendations for Industrial Strategic Energy Management Programs*; DNV-GL. (2016). *Impact Evaluation of Commercial Strategic Energy Management*. Prepared for the Energy Trust of Oregon; Stewart, James. 2017. *Chapter 24: Strategic Energy Management (SEM) Evaluation Protocol*. Golden, CO; National Renewable Energy Laboratory. NREL/SR-7A40-68316. http://www.nrel.gov/docs/fy17osti/68316.pdf; PWP and Evergreen Economics. (2017). *Current Methods in Free Ridership and Spillover Policy and Estimation*. Prepared for the Energy Trust of Oregon.

[17] See Willis, Gordon B. (2015). *Analysis of the Cognitive Interview in Questionnaire Design*. New York: Oxford University Press; Biemer, Paul P., Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman. (2004). *Measurement Error in Surveys*. Hoboken, NJ: John Wiley & Sons.

[18] Dr. Ridge was recently involved in conducting cognitive interviews for a NTGR questionnaire for the Commercial Lighting Program implemented by Xcel Energy.

Collecting data at the end of the first reporting period means that data collection will take place over a five-month period from August through December 2019, as each program concludes its first reporting period.

We will conduct the NTGR surveys via telephone interviews. Trained professionals with experience that is commensurate with the interview requirements will perform these interviews. Due to the complex nature of standard–very large NTGR projects and the related decision-making processes associated with them, these interviews will be conducted by a tandem team consisting of a senior interviewer and an experienced engineer. More than likely, these will involve interviews of multiple people involved in the project, including the primary decision maker, utility account executives, program staff, and other decision influencers.

Note that collecting data for NTGR estimates will be coordinated between two teams: the team responsible for estimating the gross savings and the team responsible for estimating the NTGR.

7.2.6. Core NTGR

We will base the core NTGR on the responses to the closed-ended questions in the participant survey. We will use the average of three scores described below. Each of these scores will be chosen to represent the highest response or the average of several responses given to one or more questions regarding participant decisions to take energy efficiency actions. The scores are as follows:

- **Program attribution index 1 (PAI–1) score** that reflects the influence of the most important program elements related to the customer's identification of opportunities and their decisions to take the energy efficiency actions at that time.
- **Program attribution index 2 (PAI–2) score** that captures the perceived influence of the program (including help with identifying opportunities, rebates, recommendations, training, or other program interventions) relative to non-program factors in the decision to take the energy-efficiency actions. This score is calculated by asking respondents to assign importance values to both the program and the most important non-program influences so that the two total 10. The program-influence score is adjusted (i.e., divided by 2) if respondents say they had already made decisions to take the actions before they learned about the program.
- **Program attribution index 3 (PAI–3) score** that captures the likelihood of participants taking the actions at this time and in the future if the program had not been available (the counterfactual). We will rely on existing CPUC guidelines to support an analysis of preponderance of evidence for early retirement claims.[19]

The self-reported core NTGR is the average of the PAI-1, PAI-2, and PAI-3 scores. The core NTGRs will be weighted by the ex post gross savings.

[19] CPUC and SCE, (2014). Early Retirement Using Preponderance of Evidence (Version 1.0 Final)

To ensure that measurement of program influence is internally consistent we will use Cronbach's alpha[20] to assess and report the reliability of NTGR estimates. Cronbach's alpha is a widely accepted measurement of the internal consistency and reliability of a multivariate measurement and is designed to measure an underlying construct (e.g. program influence). The alpha represents one type of reliability of an overall set of measurements (e.g., PAI_1, PAI_2, and PAI_3). Generally, alpha should be 0.70 or above. While Cronbach's alpha increases as the number of items in the scale increases, we usually have only three or four items distilled from multiple questions. So even with consistency checks, inconsistencies invariably emerge and lead to low inter-item correlations. As a result, an alpha lower than 0.70 might be acceptable. For large energy savers, we will consider additional quantitative and qualitative data that is not reflected in the alpha. When at least two raters are involved in independently estimating the NTGR we will also calculate inter-rater reliability.

We will also conduct sensitivity analyses (e.g., changing weights, changing the questions used in estimating the NTGR, changing the probabilities assigned to different response categories, etc.) to assess the stability and possible bias of the estimated NTGR.

7.2.7. Final NTGR

The calculation of the core NTGR is straightforward and will be based on the answers to the closed-ended questions. However, because the approach used for standard NTGR — very large analyses relies on more information from multiple sources, it requires greater effort to determine whether it is appropriate to modify the core NTGR to estimate the final NTGR. The Self-Report Guidelines point out that a case study[21] is one method of assessing both quantitative and qualitative data when estimating an NTGR. A case study is an organized presentation of all the data collected about a particular customer site with respect to the relevant aspects of the customer's decision to take energy-efficiency actions. When multiple interviews have been conducted and we have collected program documentation and both quantitative and qualitative data, we will integrate all this information into an internally consistent and coherent case study that supports the final NTGR.

Depending on the situation, we may consider the following data sources to supplement the information collected during the decision-maker interviews.

- Account representative interview
- Energy coach interview
- Utility program manager/staff interview
- Utility technical contractor interview
- Evaluation engineer interview

[20] Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

[21] Yin, Robert K. (1994). *Case study research: Design and methods*. Newbury Park, California: SAGE Publications.

- Gross impact site plan/analysis review
- Corporate green/environmental policy review (if mentioned as important)
- Corporate standard practice review (if mentioned as important)
- Industry standard practice review (if mentioned as important)

As explained in the Self-Report NTGR Guidelines, when complementing the quantitative analysis of free-ridership with additional quantitative and qualitative data from multiple respondents and sources, one must keep in mind some basic considerations. For example, additional data may reveal important influences on the customer's decision to take the actions, such as interviews with third parties who were involved in the decision to take energy-efficiency actions.

Sometimes, the data clearly points in the same direction, while at other times the data will be more ambiguous. In all cases, to maximize reliability it is essential that multiple people analyze the data. Each person must analyze the data separately and then their respective results can be compared and discussed. Using this approach, important insights can emerge from the different ways that two analysts look at the same data. Ultimately, differences must be resolved, and a case made for a particular final NTGR. Careful training of analysts in the systematic use of rules is essential to ensure inter-rater reliability.[22]

Once the individual analysts have completed their reviews, they will discuss their respective findings and the rationale for their recommended changes to the core NTGR. Key points of their rationales will be summarized and then presented in greater detail in a workpaper. This will help to ensure that an independent reviewer can understand and judge the data and the logic underlying each final NTGR estimate. Equally importantly, the CPUC will have all the essential data to enable their staff to replicate the results and, if necessary, to derive their own estimates.

7.2.8. Subsequent Program Years

The SEM Program aims to minimize energy use through conservation and efficiency. Equally importantly, it also aims to make these efforts sustainable. Thus, the focus of the first-year impact evaluation will be on estimating the program's influence on participant decisions and actions taken to reduce energy use. In the second year, we will continue to our NTG research as well as focus on estimating the influence of the program on sustaining these actions. In particular, we will seek to better understand the extent to which energy management systems has been embedded within participants corporate culture. Using the logic model and organizational change literature[23] as guides, we will develop sustainability metrics and measure

[22] Inter-rater reliability is the extent to which two or more individuals (coders or raters) agree. Inter-rater reliability addresses the consistency of the implementation of a rating system.

[23] Burke. W. Warner. (2018). *Organizational Change: Theory and Practice*. Los Angeles: SAGE; Wolfe, A.K., E.L. Malone, J. Heerwagen and J. Dion. (2014). *Behavioral Change and Building Performance: Strategies for Significant, Persistent, and Measurable Institutional Change*. Prepared for the Pacific Northwest Laboratory.

them during the fourth quarter of 2020. Data sources will include interviews with implementer staff, such as the energy coach, executive sponsor, energy champion, and members of the energy team. We will also review program documentation as part of this effort We will collect energy management assessments from both the onset of participation as well as at the conclusion of the second reporting period. This will provide us with insight into the progress that has been made in instilling a culture of energy management. As done during the evaluation of the first program year, we will interview key decision makers at the end of the second-year reporting period for each project. Data collection will take place over a five-month period from August through December 2020, depending on the date that the second reporting period ends.

7.3. Cost Effectiveness Analysis

This section describes our approach to analyzing cost effectiveness. Because costs are tracked at a program level rather than the participant level, our final reported cost effectiveness will be at the program level.

Uncertainty in the load shape of energy savings and demand savings will present a challenge when assessing cost effectiveness. Since SEM models are correlated to production and because production is tracked on a daily or weekly basis, we will not be able to use hourly load shapes or production-normalized demand savings. However, based on the mix of measures implemented, we will estimate approximate savings load shapes to analyze cost effectiveness.

Likewise, we will assign a savings-weighted effective useful life (EUL) for the BRO savings, which may include a mix of BRO and capital measures. While the CPUC has assigned a five-year measure life to BRO measures, EUL estimation is further complicated by the fact that SEM participants are welcome to implement capital projects and receive SEM BRO incentives rather than participate in the PA's custom capital programs. When this occurs, we will estimate the capital portion of the savings and assign the DEER-defined EUL. We will also estimate non-capital BRO savings and assign a five-year EUL. Last, we will calculate a savings-weighted measure life for the SEM BRO savings, which may include some capital measures.

7.4. Program Impact Estimates

We will estimate ex-post program impacts based on the gross and net savings estimates for the sample of 30 projects.

7.4.1. Gross Savings

We will estimate the gross life-cycle savings (kWh and therms) for the combined effect of all SEM-initiated measures. Separately we will do the same for SEM-capital and SEM-BRO projects. We will compute savings for each performance period and for the lifetime impact of the measure implemented during each performance period. We will estimate these program-

level gross life-cycle savings and the achieved precision for each using the stratified mean estimation method from Cochran.

For the stratified mean estimation method, the steps are as follows:

1. Calculate the mean savings

$$\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$$

where:

W_h = $\frac{N_h}{N}$ which is the stratum weight

N_h = population of stratum h

N = population of group

$\bar{y}_h$ = the mean of y for stratum h

$\bar{y}_{st}$ = the mean resulting from a stratified random sample (*st* for *stratified*).

2. Calculate the group savings (defined by whether a participant was part of the early feedback group or part of the no-early feedback group)

$Y = N^* \bar{y}_{st}$

The realization rate is the ratio of the group ex post savings to the group custom project savings

$RR = Y/X$

where X = the sum of the custom project savings in the group.

3. Calculate the variance of the mean, s^2

$$s^2(\bar{y}_{st}) = \sum_{h=1}^{L} \frac{W_h^2 \, s_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_h^2}{N}$$

where s_h^2 is the stratum variance and n_h is the stratum sample size.

Note that the second term in the equation represents the finite population correction.

The achieved relative precision at 90% confidence is

$RP = 1.645^* s(\bar{y}_{st}) / \bar{y}_{st}$

7.4.2. Net Savings

We will weight the NTGR for each sampled project by its ex post life-cycle gross savings and then sum them to arrive at the estimated NTGR for the entire sample of 30 projects. We will apply this sample-based estimate of the NTGR to the total ex post gross performance period and

lifetime savings for all 37 projects, as described above. We will calculate the achieved level of confidence and precision for the estimated NTGR in a way that is consistent with our two-stage cluster design (i.e., sampling 30 of the 37 sites and sampling four actions within the sampled sites). For each group and within each stratum, the standard error of the NTGR for the two-stage cluster design will be calculated as follows.

$$\widehat{SE}(\bar{\bar{x}}_{clu}) = \left(\frac{1}{\bar{N}\sqrt{mf_2}}\right)\left[\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m-1}\right]^{1/2}\left(\frac{N-n}{N}\right)^{1/2}$$

where:

$\bar{x} = \frac{x}{m}$	= average level of characteristic X per cluster in the sample
$x = \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} x_i$	= sample total for characteristic X
$x_i = \sum_{j=1}^{\bar{n}} x_{ij} =$	= sample total of characteristic X for ith sample cluster
f	= overall sampling fraction
$f_1 =$	= 1st-stage sampling fraction
$f_2 =$	= 2nd-stage sampling fraction
n	$= m\bar{n}$
$\bar{n} =$	= number of listing units sampled from each cluster
M	= number of clusters in population
m	= number of clusters in sample
$N = M\bar{N}$	= total number of listing units in population
$\bar{N} = N/M$	= average number of listing units per cluster in the population
x'_{clu}	$= x/f$

We will calculate the achieved level of confidence and precision for the total ex post gross performance period and lifetime savings for all 37 projects as follows. First, for each group (early-feedback versus no-early-feedback) we will calculate the two parameters, the ex post gross savings (GS) and the NTGR, the 90% relative precision. Next, we will calculate the 90% relative precision of the net savings for each group using the equation below in a way that takes into account the propagation of errors involved in multiplying these two parameters.[24]

$$RP \text{ Net Savings} = \sqrt{rp(GS)^2 + rp(NTGR)^2}$$

[24] TecMarket Works, Megdal & Associates, Architectural Energy Corporation, RLW Analytics, Resource Insight, B & B Resources, Ken Keating and Associates, Ed Vine and Associates, American Council for an Energy Efficient Economy, Ralph Prahl and Associates, and Innovologie. (2004). *The California Evaluation Framework.* Prepared for the California Public Utilities Commission and the Project Advisory Group; Taylor, John R. (1997). *An Introduction to Error Analysis: The Studies of Uncertainties in Physical Measurements*. Sausalito, CA: University Science Books.

where:

$rp(GS)$ = the 90% relative precision of the ex post gross savings

$rp(NTGR)$ = the 90% relative precision of the ex post NTGR

The error bound of the net savings for each group will then be calculated as follows:

$EB = NS \times RP$ Net Savings

where:

EB = the 90% error bound of net savings for each group

NS = the net ex post savings for each group

Once we calculate the error bound of net savings for each of the groups, we can calculate the error band for the portfolio across both groups using the following equation:

$$EB_{SEM\ Program} = \sqrt{(EB\ Early\ Feedback)^2 + (EB\ No\ Early\ Feedback)^2 \ldots .}$$

The above error bound equation is based on the following three assumptions:

1. There are no interactions between the two groups.
2. Each of the two groups has been evaluated independently.
3. Each evaluation has provided an unbiased estimate of the actual savings of the corresponding group.

The result is a consequence of (a) the fact that the standard deviation of a sum of statistically independent random variables (e.g., the estimated savings of each group) is the square root of the sum of the squares of the standard deviations of each of the random variables, and (b) the error bound being defined as 1.645 times the standard deviation.

Finally, the relative precision of the SEM program, defined as a combination of the two groups, will be calculated using the following equation:

$$RP_{SEM\ Program} = \frac{EB_{SEM\ Program}}{NS_{SEM\ Program}}$$

where:

$EB_{SEM\ Program}$ = the error band of the SEM program.

$NS_{SEM\ Program}$ = the net impacts (kWh, kW or therms) of the SEM program.

Note that results will also report electric and gas savings separately for the same combinations.

8. Data Management and Quality Control

We will develop an Excel workbook to standardize how we document all findings and corrections from the review of each project. The workbook will have a series of locked tabs that contain data-entry ranges and validate each value entered. Instructions for entering data will appear in adjacent cells. Our analysts will add additional tabs to the workbook to document the models and inputs used to calculate savings.

Our senior engineering staff will:

- Develop the workbook
- Train our project-review staff
- Provide quality control
- Resolve problems

In their quality-control role, our senior engineering staff will review each of the complete project-evaluation workbooks and interact as necessary with our analysts to ensure that all projects are treated consistently.

Our data-processing team will compile data for the project workbooks and associate that data with claim data. They will apply additional quality-control tests to the data and refer any issues that need resolution to the assigned analyst.

9. Task Plan

The task plan describes the sequence of the work, logistics related to completing the work, and the deliverables. Its structure corresponds to the budget and schedule portions of the work plan.

9.1. Workplan

This document represents our SEM workplan. Subsequent to this initial workplan, we will submit updated workplans in 2020 and 2021 to continuously refine and improve our approach, and to respond to changes as programs evolve.

9.2. SEM Data Request

Two data requests have already been issued. We have determined that the Custom Measure and Project Archive (CMPA) will serve as the platform to issue data requests and receive data.

9.2.1. Review Obtained Files for Sampled Sites

Before any work begins, we will review the data files obtained for sampled sites for quality. It is important to ensure that the data sets are complete before beginning the evaluation. If any gaps, clarifications or issues arise we will contact the appropriate CPUC representative for guidance in resolving these issues with the IOUs.

9.3. SEM Evaluation Sampling

We have completed our sample design and run the sample. As discussed in section 4, we established samples from the population of 37 SEM participants across four SEM programs. To complete this task, we aggregated all SEM participants in a spreadsheet and assigned a random number between 1 and 37 to each participant. We then arranged sites numbered 1-19 into the "early feedback" sample and sites 20-37 into the "no early feedback" sample.

Within each of the sample groups we used a stratified random sample approach to sample 15 participants from the early-feedback group to receive early feedback. We similarly sampled 15 participants from the no-early-feedback group. We categorized the three strata as small, medium, and large savers, and based the assignments to each category on the range of IOU-projected energy savings.

To prevent bias, we will not disclose sample status to the IOUs or implementers.

9.4. Early Feedback

We will conduct early feedback consistent with the approach described in section 5 of this workplan.

Deliverable – Our deliverable for early feedback is to conduct working sessions with each PA individually. In these sessions, we will provide site- and program-level feedback.

After the PA sessions, we will conduct a session with the CPUC to review trends in site- and program-level feedback. We will summarize this feedback in the form of recommended revisions to the statewide design and M&V guides.

9.4.1. Participant-Level Early Feedback

To confirm participant eligibility, we will conduct a detailed review of scoping reports, site data, map details, and other publicly available information. Beyond eligibility of a site to participate in SEM, we plan to understand how the sites are metered; what sources of energy they consume; if they self-generate; and what their physical characteristics are. In addition, we will conduct site visits, interview facility staff, and review the opportunity registers, hypothesis models, and other deliverables. Additionally, we will evaluate all custom and deemed projects that are completed at the site during the baseline and reporting periods.

9.4.2. PA-Level Early Feedback

The tasks associated with our PA review include assessing the workshops, curriculum, and program reporting. We will work to better understand the communication between the implementers and the PAs. Additionally, we will work to assess the level of customer engagement and provide constructive feedback to improve program design. Our early feedback to PAs will include implementer interviews. Additionally, our participant-level review activities will identify common findings that apply to multiple participants, indicating the need for PA-level improvements.

9.4.3. Statewide Design & M&V Guide Early Feedback

We will review and provide feedback on the Statewide SEM Design Guide and M&V Guide. Additionally, through our participant-level and program-level review activities, we will identify common findings shared across participants and PAs, indicating the need for statewide design improvements.

9.5. Custom Project Review

We will conduct our custom project review consistent with the approach proposed in section 6 of this workplan. As previously mentioned, we are segregating the custom project review work between SEM custom and SEM BRO.

Deliverable – Each custom project review will result in a disposition. Dispositions for custom projects will follow the CPUC's custom project review procedure.

9.5.1. SEM Custom Projects Review

It is our understanding that the PAs have agreed to flag all SEM projects as SEM custom or SEM NMEC in column 15 of the bimonthly custom project review project list. We support this designation and will review each custom project review project list for both SEM NMEC and SEM custom projects. We will select all SEM projects for custom project review at SEM participants sampled for early feedback. Additionally, we recommend establishing a protocol for PAs to flag deemed projects at SEM sites. Perhaps a SEM deemed designation.

The custom project review of SEM custom projects will follow a similar approach and technical rigor as our other custom project review work for non-SEM sites (Deliverable 17). The SEM custom project custom project review will be treated as any other custom project review of custom project and will result in dispositions.

All engineering calculations, calculator tools, observations and project specifications will be evaluated for completeness. If there are any uncertainties or gaps in the analysis, we will work with the implementers, PAs, and sites to seek resolution. Further definition is provided in section 6.1.

9.5.2. SEM BRO Savings Review

We will review SEM BRO savings through a review of all final SEM models at sites selected for early feedback and custom project review as described in section 6.2 of this workplan.

Deliverable – The custom project review of SEM BRO models will result in dispositions. We expect that the PAs will not issue performance-based payments prior to our custom project review. We commit to following the custom project custom project review timing defined by SB 1131.

9.6. Ex Post Evaluation

Our ex post evaluation will be consistent with the approach described in section 7 of this workplan. This work will include site visits (where not recently conducted for early feedback) to gain a deeper understanding of the facility, variables that affect energy use, possible non-routine events, etc. Additionally, we will interview facility staff and evaluate all custom and deemed projects completed during the baseline or reporting periods.

Deliverable – We will provide a report of our activities, findings, and recommendations. It will provide sampled project and program-level gross realization rates and NTG ratios, as well as gross and net evaluated annual and life-cycle energy savings.

9.6.1. Gross Savings Evaluation

We will evaluate gross savings on all 30 SEM sites sampled for ex post evaluation as described in section 7.1 of this workplan. Key tasks include:

- Review SEM model and M&V report
- Identify non-routine events, including custom and deemed projects
- Visiting site and interviewing facility staff
- Conduct ex post evaluation of all custom and deemed projects completed during the baseline or reporting periods
- Quantify ex post evaluated energy savings

9.6.2. Net Savings Evaluation

We will conduct a net-to-gross evaluation to determine net savings, as described in section 7.2 of this workplan. Key tasks include:

- Sampling
- Program Theory and Logic Model
- Questionnaire development
- Data collection
- Core NTGR
- Final NTGR

9.6.3. Cost-Effectiveness Analysis

We will analyze cost effectiveness as described in section 7.3 of this workplan. This will use the existing CPUC cost-effectiveness calculator as well as our ex post results, EUL and savings load shapes to quantify cost effectiveness. We will estimate of EUL and savings load shape through a bottom-up review of the significant energy-savings actions taken in the opportunity register.

9.6.4. Program Impact Estimates

Once we have analyzed all sampled sites, we will estimate program-level gross and net savings as described in section 7.4.

10. Budgets and Schedule

This section presents our budget and schedule for evaluating SEM programs.

10.1. Evaluation Budget

The SEM evaluation is part of the larger budget for all early feedback and ex post evaluation (Group D contract deliverables 9, 10, 13, and 17). The other portion of this larger budget is for the evaluation of large commercial, industrial and agricultural (LCIA) projects. Table 6 presents our budget in total and separately for LCIA and SEM activities discussed in this workplan. We have not included the budget for custom project review of SEM custom or BRO projects in this table.

Table 6: SEM Evaluation Budget

	Budget by category ($)		
Deliverables	**LCIA**	**SEM**	**Early feedback and ex post evaluation (Total of LCIA and SEM)**
Workplans and updates	168,000	168,000	336,000
Data collection and sampling plans	129,000	43,000	172,000
SEM early feedback	-	495,000	495,000
Gross savings estimates	3,632,000	353,000	3,985,000
Net savings estimates	640,000	60,000	700,000
Final reports	343,000	171,000	514,000
Data documentation	133,000	67,000	200,000
Total	**5,045,000**	**1,357,000**	**6,402,000**

10.2. Timeline and Milestones

Figure 7 presents our timeline and milestones.

Deliverable	Task / Subtask	Due date	Year:	2019				2020				2021			
			Qtr:	1	2	3	4	1	2	3	4	1	2	3	4
Evaluation work plan															
	Draft release - ED review	19-Apr-19													
	Draft release - stakeholder review	24-May-19													
	Final posted	28-Jun-19													
	Plan update	Apr-20													
Data collection and sampling plan		Jul-19													
Early feedback evaluation															
	Reviews (participants, PAs, statewide)														
	IOU-specific webinars	Oct-19, -20													
Custom project review (custom, NMEC)															
Gross savings estimates		Apr-21													
Net savings estimates		Apr-21													
Final evaluation report															
	Cost-effectiveness, population impacts														
	Draft release - ED review	5-Jun-21													
	Draft release - stakeholder review	1-Aug-21													
	Final posted	31-Aug-21													
Data transfer and documentation		Sep-21													

Key:

Activity to complete deliverable expected

Major deliverable to be completed

Figure 7: Timeline and Milestones

11. Management and Staffing Plan

Figure 8 shows an organizational chart of our management and staffing structure.

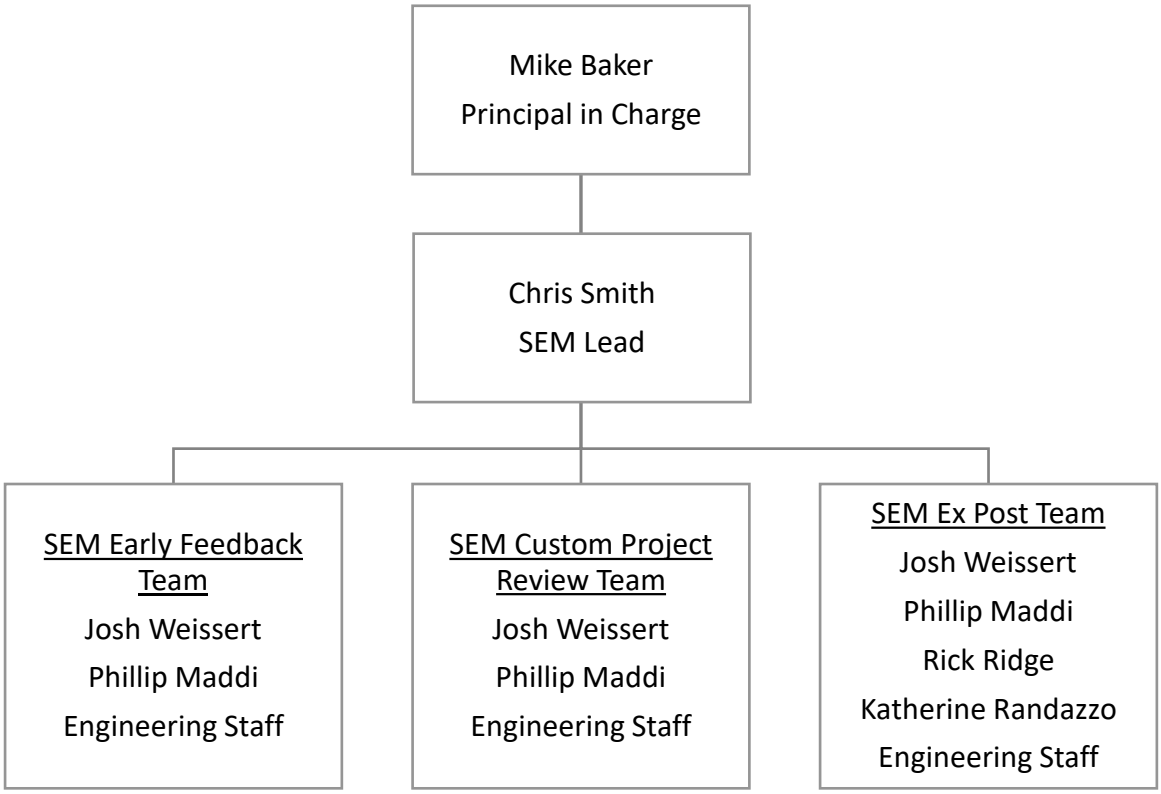


Figure 8: SEM Evaluation Organizational Chart

Appendices

Appendix A: Data Requests

The following data requests have been fulfilled by PAs.

A.1 Issued 2-11-2019

1. Data Request 1: Any draft or final materials used in the marketing or delivery of the SEM program, including but not limited to program applications, eligibility screening tools, real time review guide, or other documents that describe program procedures and operations, beyond what is available in the published program implementation plans.
2. Data Request 2: For all sites, site-specific documentation including, but not limited to - scoping report, opportunity register, hypothesis model, primary data used in developing the hypothesis model, data collection plans, and description of non-routine events which occurred in the baseline period. In addition, provide two lists for each site:
   a. List 1: Non-SEM Savings projects: Pre-existing projects identified and planned prior to SEM engagement and implemented during the SEM engagement, whether receiving incentives or not.
   b. List 2: SEM Incented Project Savings projects: Incented projects (i.e. custom capital projects) identified, planned, and implemented during the SEM engagement receiving incentives at or near the incentive rate for another program (i.e. "capital project" incentive rate).

A.2 Issued 3-5-2019

Group D – Custom EM&V

Selecting and Processing SEM 2018 Sample for Custom Project Review and Early Feedback

Summary

The CPUC has selected SBW Consulting, Inc. (SBW) as the Group D – Large Commercial, Industrial and Agricultural evaluation, measurement and verification and custom project custom projects review contractor. They are now assisting CPUC staff in conducting both custom project review and Ex Post evaluation of the Strategic Energy Management (SEM) program. The PA's have previously notified CPUC staff that they have engaged work with 37 SEM participant sites, which constitute the entire population of the participants in this program. CPUC staff has directed SBW to select samples of these participants, to complete custom project review and Ex Post evaluation.

As described below (Procedure) SBW will prepare (with the assistance of the PA's) a listing of the 37 participant sites, including a "rough guess of savings." Each of the 37 sites will be randomly assigned to one of two groups. A randomly selected sample of 15 sites from one group

will receive early feedback and Ex Post evaluation. The sample of 15 selected from the second group will receive only Ex Post evaluation.

The sample of 15 sites selected for early feedback will also be subject to custom project review. The CPUC designates that all SEM custom projects (motivated by the SEM program) developed for any of these sites are sampled for custom project review and must be submitted to CMPA in accordance with the custom project review procedure. CMPA folders will be created for each of these participant sites and all documentation required for custom project review should be placed in those folders for any SEM custom projects that are proposed and become ready for review during the two-year SEM performance period.

The CPUC staff anticipates that more than one SEM custom project may be proposed for any one of these sites and is establishing site-level folders so that all such projects associated with a site can be collected in a single place.

Procedure

PG&E, SDG&E and SCE (Filing jointly with SCG) will submit a list of their SEM participants to PA's Non-project Files directory in the CMPA using the attached template. The CMPA upload folder will be named as follows: [Program Administrator]19SEM. The attached template requires the following information:

PA. PG&E, SDG&E or SCE/SCG.

Participant ID. ID that uniquely identifies each participant in the program.

Site Name. Name of the site as it appears in the Opportunity Register.

Site Type. A brief description of the business conducted at the site, e.g., Food Processing

Customer Name. Name of the organization that has agreed to participate in the program.

Street Address. Street address for the facility where the SEM measures will be implemented.

City. City where the facility is located.

Net kWh Savings. Rough guess for the net kWh savings for the 2-year performance period for this participant. Select the most appropriate range from the drop-down list:

< 250,000

250,000 to 500,000

500,000 to 1,000,000

1,000,000 to 2,000,000

2,000,000

Net Therms Savings. Rough guess for the net Therms savings for the 2-year performance period for this participant. Select the most appropriate range from the drop-down list:

<10,000

10,000-25,000

25,000-50,000

50,000-100,000

>100,000

SBW will combine the lists provided into a statewide list of SEM participants. SBW will use the kWh/Therms savings rough guesses to assign each participant to one of three savings strata (small, medium and large). Next, SBW will randomly assign each of the participants to one of two domains: early feedback and no-early feedback. SBW will then select a stratified random sample of 15 participants from each of the two domains. ED will review and approve this sample.

CPUC staff will notify PA's of which participants have been selected for early feedback and that all SEM custom projects associated with these participants have been selected for custom project review.

SBW will create a CMPA folder for each of the 15 early feedback participants. These folders will be named: [Program Administrator]19SEM[Participant ID]

As SEM custom projects, for early feedback participant sites, become ready for review, PA's will upload documentation needed to support custom project review to the corresponding site folder. The collection of files for any one project should be placed in a ZIP file prior to upload. The ZIP file name should be: [Program Administrator]19SEM[Participant ID][Application Number].zip.

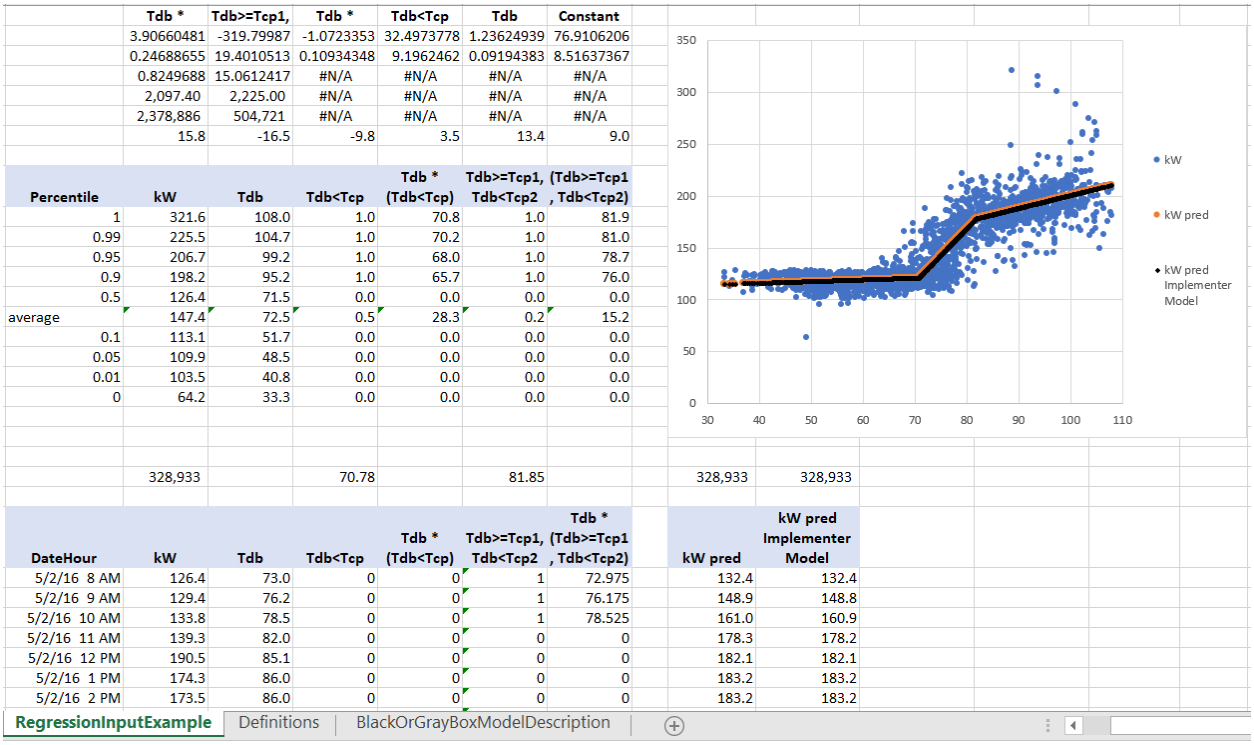
Appendix B: Sample Model Review Tool

A challenge in reviewing NMEC models is conducting standardized reviews of a wide range of models using various Excel-based tools and specialized regressions software (such as ECAM, JMP, R, Stata, or SAS). To facilitate a uniform and organized review of a range of models and approaches, we have drafted an Excel-based model-review template. While we will work outside of this file, we will standardize the results, inputs, and statistical indicators in the template.

The Excel-based template includes three tabs. For sites with multiple models, each model will require an additional tab. The tabs are defined as follows:

- RegressionInput_Name – This tab includes the independent variables used to derive the regression equations, and the statistical results of the models. Each independent variable used in the regression model is listed along with relevant raw data, sorted by timestamp. Any parameters necessary to fully define the independent variables (e.g. indicator variables, inflection points, base temperature for HDD/CDD) are defined immediately above the raw data.
- Definitions – This tab is used to define each variable used in the model.
- BlackOrGrayBoxModelDescription – On this tab, the evaluator can summarize the model review if it will not easily fit into a linear regression.

Figure 9 shows a sample of the model review template.



	Tdb *	Tdb>=Tcp1,	Tdb *	Tdb<Tcp	Tdb	Constant
	3.90660481	-319.79987	-1.0723353	32.4973778	1.23624939	76.9106206
	0.24688655	19.4010513	0.10934348	9.1962462	0.09194383	8.51637367
	0.8249688	15.0612417	#N/A	#N/A	#N/A	#N/A
	2,097.40	2,225.00	#N/A	#N/A	#N/A	#N/A
	2,378,886	504,721	#N/A	#N/A	#N/A	#N/A
	15.8	-16.5	-9.8	3.5	13.4	9.0

Percentile	kW	Tdb	Tdb<Tcp	Tdb * (Tdb<Tcp)	Tdb>=Tcp1, Tdb<Tcp2	(Tdb>=Tcp1 , Tdb<Tcp2)
1	321.6	108.0	1.0	70.8	1.0	81.9
0.99	225.5	104.7	1.0	70.2	1.0	81.0
0.95	206.7	99.2	1.0	68.0	1.0	78.7
0.9	198.2	95.2	1.0	65.7	1.0	76.0
0.5	126.4	71.5	0.0	0.0	0.0	0.0
average	147.4	72.5	0.5	28.3	0.2	15.2
0.1	113.1	51.7	0.0	0.0	0.0	0.0
0.05	109.9	48.5	0.0	0.0	0.0	0.0
0.01	103.5	40.8	0.0	0.0	0.0	0.0
0	64.2	33.3	0.0	0.0	0.0	0.0
	328,933		70.78		81.85	

DateHour	kW	Tdb	Tdb<Tcp	Tdb * (Tdb<Tcp)	Tdb>=Tcp1, Tdb<Tcp2	Tdb * (Tdb>=Tcp1 , Tdb<Tcp2)	kW pred	kW pred Implementer Model
5/2/16 8 AM	126.4	73.0	0	0	1	72.975	132.4	132.4
5/2/16 9 AM	129.4	76.2	0	0	1	76.175	148.9	148.8
5/2/16 10 AM	133.8	78.5	0	0	1	78.525	161.0	160.9
5/2/16 11 AM	139.3	82.0	0	0	0	0	178.3	178.2
5/2/16 12 PM	190.5	85.1	0	0	0	0	182.1	182.1
5/2/16 1 PM	174.3	86.0	0	0	0	0	183.2	183.2
5/2/16 2 PM	173.5	86.0	0	0	0	0	183.2	183.2

Figure 9: Model Review Template Sample

Appendix C: Sample Logic Model