GROUP D

Evaluation, Measurement, and Verification of Program Year 2023 Strategic Energy Management Projects

California Public Utilities Commission
CALMAC ID:
Date: April 25, 2025

Table of contents

List of figures

List of tables

Glossary of key terms and acronyms

As-built conditions – Refers to any site-specific or measure-specific parameters that could influence the energy savings, including quantities, sizes, load profiles, sequences of operation, setpoints, etc., as found and verified by the evaluators during the data collection phase.

Baseline period – The baseline period is the 12 or 24 months leading up to the energy efficiency intervention or retrofit.

Bottom-up savings calculation – A savings methodology that calculates the SEM energy savings utilizing measure-level calculations. This methodology uses measure-specific formulas, inputs, and assumptions, etc., to calculate the measure-specific savings. The overall site savings are then calculated by aggregating the energy savings of all implemented measures.

BRO measures – Refers to implemented or planned SEM measures that are behavioral, retrocommissioning, and operations.

CCT – Refers the Custom Core Template, which is an Excel-based tool utilized by the SEM evaluation team to report site-specific data collection efforts, review of participant documentation and methods, and documenting SEM evaluation team's methods and findings.

California Database for Energy Efficiency Resources (DEER) – Refers to the Database for Energy Efficient Resources. This database contains information on energy efficient technologies and measures. DEER provides estimates of the energy-savings potential for these technologies in residential and non-residential applications. DEER is used by California Energy Efficiency (EE) Program Administrators (PAs), private sector implementers, and the EE industry across the country to develop and design energy efficiency programs.[1]

California Energy Data and Reporting System (CEDARS) – Refers to the database that securely manages California Energy Efficiency Program data reported to the Commission by Investor-Owned Utilities, Regional Energy Networks (RENs), and certain Community Choice Aggregators (CCAs).[2]

Custom measure and project archive (CMPA) – Refers to the CPUC regulatory supervision website (Energy Division Non-DEER Resources) which is the archive of custom measures and projects utilized by California investor-owned utilities (IOUs) and reviewed by Energy Division staff. Every project supports secure uploading and browsing of files.

Custom project review (CPR) – Refers to the process of selecting custom projects, submitted biweekly by the PAs, for review of all forecasted savings parameters and documents of selected projects.

Effective useful life (EUL) – An estimate of the median number of years that the measures installed under the program are still in place and operable.

Forecasted gross savings – Engineering-based savings estimate derived before installation.

Gross realization rate (GRR) – Refers to the ratio of achieved (evaluated) energy savings to forecasted energy savings; as a multiplier on Unit Energy Savings, the GRR considers the likelihood that not all CPUC approved projects undertaken by IOUs will produce gross savings.

---

[1] Public utilities commission of California, Resolution E-5152, August 5, 2021. http://www.deeresources.com/files/DEER2023/Resolution%20E-5152%20DEER2023%20Complete.pdf

[2] California Energy Data and Reporting System (CEDARS), "Welcome to CEDARS," cedars.sound-data.com, https://cedars.sound-data.com/

Gross savings – Gross savings count the energy savings from installed energy efficiency measures irrespective of whether those savings are from free riders, i.e., those customers who would have installed the measure(s) even without the financial incentives offered under the program.

Initial claimed savings – For SEM projects, the savings claimed in CEDARS following project implementation.

Lifecycle savings – Refers to the savings associated with the lifetime of an efficiency measure undertaken by a program participant. Equipment replaced early in its useful life might receive reduced savings for a portion of its lifetime.

Measure – Specific customer action that reduces or otherwise modifies energy end-use patterns. A product whose installation and operation at a customer's premises reduces the customer's on-site energy use, compared to what would have happened otherwise.

Measure application type (MAT) – Refers to the installation basis for each claim. There are seven approved measure application types: Add-on Equipment, Accelerated Replacement, BRO-Behavioral, BRO-Operational, BRO-Retro-commissioning (RCx), New Construction, and Normal Replacement.

Mixed-analysis sites – Refers to sites that used both top-down and bottom-up analysis methods in some way. Typically, this would involve a bottom-up analysis approach in one year of the two-years cycle and a top-down modeling approach in other year. In some cases, this was done when estimating savings for both fuels where one fuel would use one approach and the other fuel would use the other to estimate savings.

Net savings – The savings realized when free-ridership is accounted for. Savings are calculated by multiplying the gross evaluated savings by the net-to-gross ratio.

Net-to-gross ratio (NTGR) – A ratio or percentage of net program savings divided by evaluated gross or total impacts. NTGRs are used to estimate and describe the free ridership that may be occurring within energy efficiency programs.

Non-routine adjustment (NRA) – Non-routine adjustments are used to account for the effects of non-routine events, where the changes affected by the NRE are not suitable to the baseline or reporting period adjustment models. Non-routine adjustments occur separately from the routine adjustments made using independent variables in the adjustment model. Non-routine adjustments are developed using methods including but not limited to engineering analysis, sub-metering, or other analyses using the metered energy use data.

Non-routine event (NRE) – A non-routine event is an externally driven (i.e., not related to the energy efficiency intervention) significant change affecting energy use in the baseline or the reporting period and therefore must be accounted for in savings estimations. Typical NREs include changes in facility size, changes in facility activity not affected by the energy efficiency measures (such as addition or removal of a data center) or other modifications to the facility or its operation that alter energy consumption patterns and are unrelated to the program intervention.

Normalized savings – Savings calculated as the difference between the weather normalized baseline and performance period statistical models.

Program administrator (PA) – An entity tasked with the functions of portfolio management of energy efficiency programs and program choice (i.e., Marin Clean Energy (MCE),[3] Pacific Gas & Electric (PG&E), Southern California Edison (SCE), Southern California Gas (SCG), San Diego Gas & Electric (SDG&E)).

[3] MCE is a not-for-profit public agency that provides electricity service to more than 1 million residents and businesses in 37 member communities across four Bay Area counties: Contra Costa, Marin, Napa, and Solano.

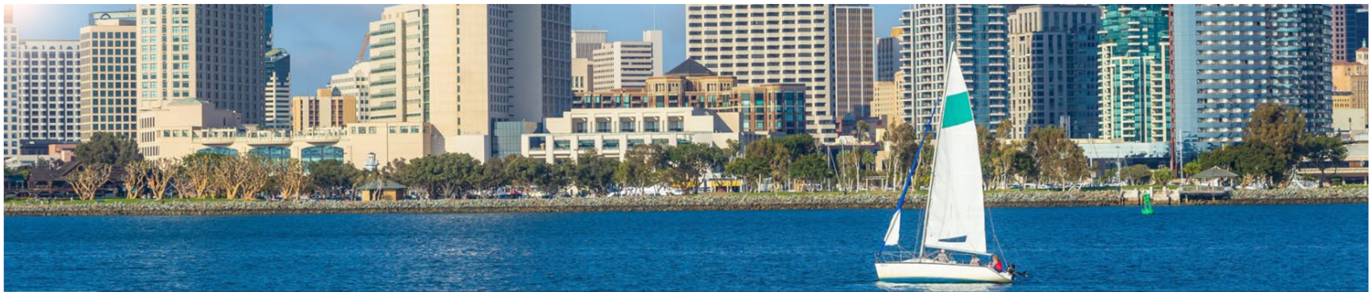
Peak demand – Refers to the average demand impact, for installed or implemented measures, as would be applied to the electric grid. CPUC Resolution E-4952 approved the Database for Energy-Efficient Resources (DEER) for 2020. Additionally, this resolution revised the DEER Peak Period definition from 2:00 p.m. to 5:00 p.m. to 4:00 p.m. to 9:00 p.m. effective January 1, 2020. In accordance with the CPUC memo issued on 03/21/19, operationalizing the 2020 DEER Peak Period change, effective January 1, 2020, per CPUC Res E-4952 for custom projects shall follow the Statewide Custom Project Guidance Document, Version 1.4.

Relative precision – A ratio of the error bound divided by the value of the measurement itself. This provides an error on a relative basis that is frequently used to show uncertainty as a fraction of a quantity. In this report, all relative precisions are provided at the 90% confidence interval, which means that in repeated sampling 90 times out of 100 the true value will fall within the lower and upper bounds of the estimate.

Top-down savings calculation – A savings methodology that calculates the SEM energy savings using facility models on the site level. This methodology uses a billing analysis utilizing multivariable regressions of utility meter data along with the relevant independent variables (such as levels of production or weather conditions) between the baseline period and the reporting period.

True-up savings – The savings claimed in CEDARS following the end of the performance period. This value is expected to be the difference between initial claimed savings and the normalized savings.

1 EXECUTIVE SUMMARY



1.1 INTRODUCTION

This interim report presents the key findings of the Strategic Energy Management (SEM) program impact evaluation conducted by DNV on behalf of the California Public Utilities Commission (CPUC) for program year (PY) 2023. The SEM program helps customers identify low- and no-cost operations and maintenance improvements to reduce energy consumption. This report marks the first phase of a comprehensive two-year evaluation covering both PY 2023 and 2024. The results of PY2023 will be combined with those from PY2024 to produce results for application to the overall program savings claims.[4] The focus of this study is to monitor program performance and provide recommendations to improve it through the evaluation of a partial sample of SEM projects. The PY2023 sample includes SEM projects Marin Clean Energy (MCE),[5] Southern California Edison (SCE), and Southern California Gas (SCG), while the PY2024 sample is expected to include projects from Pacific Gas & Electric (PG&E) and San Diego Gas & Electric (SDG&E). The SEM participating sites for PG&E and SDG&E are grouped by each Program Administrator to complete their cycles every other year, with PY2024 being the next scheduled completion year. Therefore, projects from these utilities will be included in the PY2024 sample.

Goal and objectives

The overall goal of this interim study is to evaluate energy and demand savings for SEM projects implemented in PY2023 and provide recommendations for improving program performance through the evaluation of a partial sample of SEM projects.

The primary objectives of this interim study are to:

1. Develop first-year and lifecycle evaluated gross[6] savings for the SEM program in program year PY2023.
2. Develop first-year and lifecycle evaluated net[7] savings for the SEM program in program year PY2023.
3. Identify the reasons for deviations between forecasted and evaluated savings.
4. Develop meaningful, actionable recommendations to improve program performance in delivering energy efficiency savings through 2024, with the goal of informing the overall results of the full two-year impact study, which will be completed next year.

---

[4] Due to the unique structure of the SEM program, energy savings projects are implemented over a two-year cycle, with site participation considered complete only after both reporting periods. Some Program Administrators (PAs) group cohorts together, completing site participation every two years. As a result, statewide results that include projects from these PAs can only be generated every other year.

[5] MCE is a not-for-profit public agency that MCE provides electricity service to more than one million residents and businesses in 37 member communities across four Bay Area counties: Contra Costa, Marin, Napa, and Solano.

[6] **Gross savings** count the energy savings from installed energy efficiency measures irrespective of whether those savings are from free riders, i.e., customers who would have installed the measure(s) even without the financial incentives offered under the program.

[7] **Net savings** are the savings realized when free-ridership is accounted for. Savings are calculated by multiplying the gross savings by the net-to-gross ratio.

Gross and net results from this study will be combined with results from a subsequent PY2024 study to achieve a high level of precision when fully combined.

Background

SEM is a unique program under California's statewide energy efficiency (EE) portfolio. Its purpose is to promote holistic, long-term energy savings in facilities through ongoing engagement, continuous education, and measurement of performance. The program, originally focused on industrial customers, enrolled the first cohort of participants in 2018, but is now open to some non-industrial sectors. SEM has a separate statewide Design Guide[8] and an M&V Guide,[9] which all SEM program implementers are required to use for designing and implementing the program. The preferred method of estimating energy savings for the SEM program is a top-down modeling approach, which calculates the SEM energy savings using facility models on the site level based on energy billing data. The alternative approach to calculate SEM savings is bottom-up, which calculates the SEM energy savings using measure-level calculations.

This is the third impact evaluation of the SEM program done on behalf of the CPUC. The first study was completed on the 2018 and 2019 program years (the PY2018/PY2019 Study).[10] The second study was completed on the 2021 and 2022 program years (the PY2021/PY2022 Study).[11]

1.2 METHODOLOGY OVERVIEW

The DNV team estimated the accuracy of gross and net savings the PAs claimed for SEM projects installed in PY 2023. Our gross and net savings calculation methods are summarized below. This evaluation study adheres to the International Performance Measurement and Verification Protocol (IPMVP) [12] and the California Evaluation Protocol.[13] Figure 1-1 shows the overall evaluation process.

---

[8] "The California Industrial SEM Design Guide" provides the program requirements for qualifying as a SEM Program. The guide includes the sequence and curriculum for the program participants and is delivered by program implementers.

[9] "The California Industrial SEM M&V Guide" establishes an M&V process to which industrial facilities as part of the SEM program must adhere for program engagement. The framework defines the protocols to determine a participant's energy baseline, track energy performance throughout the engagement, document energy savings, and validate the used M&V methods. M&V in SEM typically relies on a consumption-based energy model or measure-level engineering calculations.

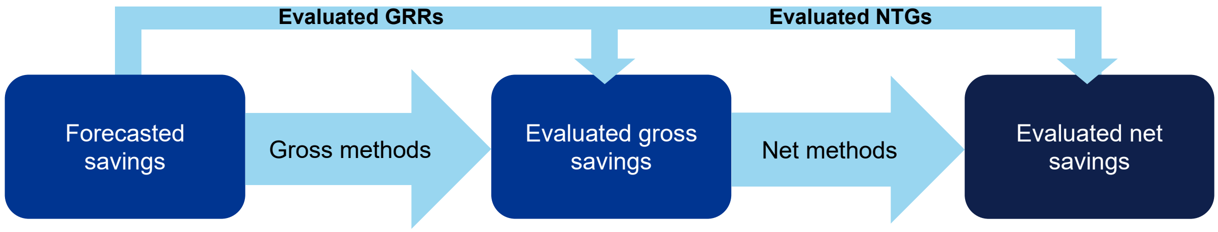
[10] SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

[11] DNV, *Group D Strategic Energy Management (SEM) 2021-2022 Impact Evaluation*, Final Report, March 5, 2024. https://pda.energydataweb.com/api/view/3970/CPUC%20Group%20D%20SEM%20Impact%20Report%20Final.pdf

[12] IPMVP is a protocol that facilitates a common approach to measuring and verifying energy efficiency investments. IPMVP incorporates M&V best practices in a non-prescriptive framework that allows it to be applied flexibly based on a measure's application and the information available.

[13] The California Evaluation Protocol (CEP) is a set of guidelines and procedures developed by the California Public Utilities Commission (CPUC) for conducting evaluations of energy efficiency programs.

Figure 1-1. SEM gross and net savings methods PY2023



Gross methods

The DNV team determined the appropriate evaluation approach for each sampled site based on the project documentation review and the collected data and information from the site contacts. We presented all site-specific M&V plans and evaluation findings in the Custom Core Template (CCT) tool. CCT is a macro-based Excel workbook that documented PA savings claims along with DNV's assumptions and methods used to estimate savings. It also reported evaluated savings, data collected from sites, analysis methods used, reported gross realization rates, and, finally, documented the reasons for discrepancies between claimed savings and evaluated savings. Data collection consisted mostly of participant interviews to determine which measures were installed and operating, photographs to verify installed measures, consumption data to estimate savings, and in some cases, trend data or performance logs to confirm operation.

Net methods

The methodological approach for the present evaluation is informed by Decision D16-08-019[14], which states that a well-designed SEM program's holistic and long-term approach encourages the implementation of behavioral, retrocommissioning, and operational (BRO) measures as well as custom and capital measures. The decision concludes that, when program influence is demonstrated, capital measures may apply the SEM default net-to-gross ratio (NGTR), which has been 1.00. This evaluation cycle employed two evaluation methods:

- **Self-reported attribution (SRA) for capital measures**. Under the SEM framework, capital measures may apply a SEM NTGR "when program influence is evident." The SRA is well suited to the assessment of capital measures that typically entail installation of equipment and a structured internal decision-making process that can be investigated through interviews of decision makers. The influence of the capital measures was determined using survey instruments and methods approved in the PY2021/2022 evaluation with improvements.
- **Theory-driven attribution (TDA) for BRO measures**. SEM seeks to change organizational behavior to produce persistent ongoing savings through a series of prescribed activities. The TDA approach is well suited to assessing whether the activities are delivered, and the organization is producing the intended outcomes. The methods and survey instruments were developed in the 2018–2019 TDA evaluation, with further adaptions in this evaluation.

These 2023 results will be combined in the 2024 evaluation cycle to develop a combined 2023–2024 NTGR for future application.

---

[14] Public Utilities Commission of the State of California, Order Instituting Rulemaking Concerning Energy Efficiency Rolling Portfolios, Policies, Programs, Evaluation and Related Issues, August 18, 2016, http://ccag.ca.gov/wp-content/uploads/2016/10/6.5-Attachment-ALJ-Decision-16-08-019-081816.pdf

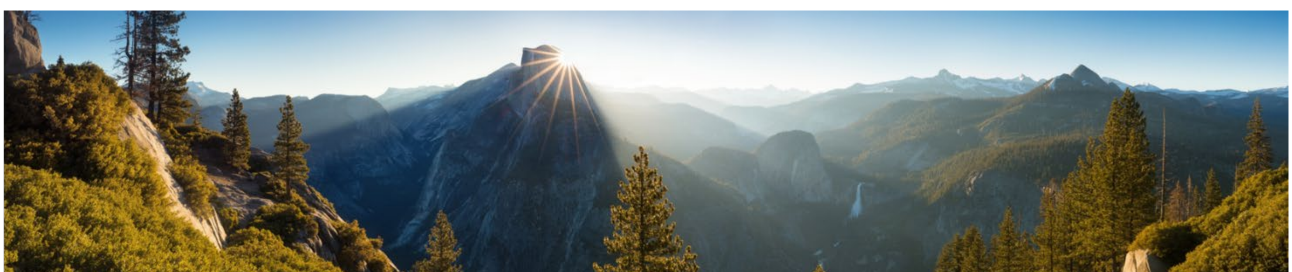
DNV

1.3 EVALUATED PROGRAM SAVINGS CLAIMS

This SEM impact evaluation focused on customers who completed a two-year cycle in 2023. We verified the aggregated two years of savings for each site we reviewed. Table 1-1 provides a summary of this SEM population. There were 42 unique customers in this SEM participant pool who met these criteria. Of these 42 customers, 31 completed and reported savings for electricity-saving measures, and 12 completed gas-saving measures. Table 1-1 shows the number of participants and the first-year and lifecycle savings that were claimed for this group of customers.

Table 1-1. SEM PY2023 evaluation population summary

Savings parameter	Number of participants	Reported savings	
		MWh	MW
Electric first-year savings	31	29,890	3.7
Electric lifecycle savings	31	149,448	18.4
		Therms (1,000)	
Natural gas first-year savings	12		986
Natural gas lifecycle savings	12		4,928
		MMBtu	
Total energy first-year savings	42		200,538
Total energy lifecycle savings	42		1,002,690

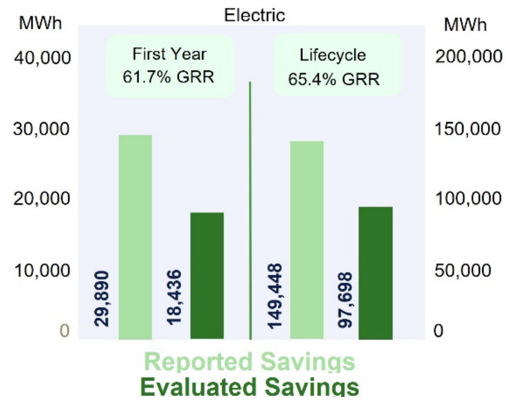


1.4 RESULTS

Gross savings results

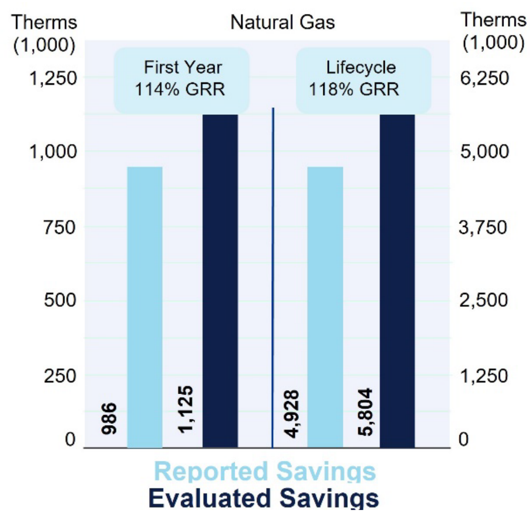
This section presents the overall electric and natural gas savings and gross realization rates (GRRs) for the evaluated PAs. All relative precisions in the tables that follow are calculated at the 90% confidence interval.

Key electric gross findings



- As shown in the chart to the right, the SEM participants who completed a two-year cycle in 2023 **achieved an aggregated first-year gross electric savings of 18,436 MWh** with a statewide[15] GRR of 62%. A key driver of the electric first-year and lifecycle realization rates was model adjustments made to enhance the statistical significance of the models used to estimate forecasted savings and ensure they accurately reflect typical facility operations. Additionally, annualization discrepancies occurred, where participants calculated energy savings by prorating savings from a limited period to annual savings, rather than using all available valid data points recorded throughout the reporting period. The model optimization and savings annualization issues mentioned above contributed to the decrease in the electric GRR. We anticipate that addressing these issues will improve the combined results for PY2023 and PY2024.
- PY 2023 SEM program participants **achieved an aggregated lifecycle gross electric savings of 97,698 MWh** with a statewide GRR of 65%. The GRR for lifecycle gross electric savings is slightly higher than for first-year gross electric savings, as most capital projects reviewed have an effective useful life (EUL) exceeding five years. Examples include LED lighting and lighting controls installation.

Key gas gross findings



- As shown in the chart to the right, the SEM customers who completed a two-year cycle in 2023 **achieved an aggregated first-year gross natural gas savings of 1,125 thousand therms** with a statewide GRR of 114%. Key drivers of the natural gas first-year and lifecycle realization rates are adjustments to savings calculation methodologies, inputs, and parameters based on an in-depth engineering review and as-built data collected from facility personnel. The natural gas GRR was driven high primarily by one site, where the post-installation operational load was significantly lower than forecasted due to the installed SEM measures. We anticipate that the impact of this site on the overall results will be reduced when PY2023 results are combined with those of PY2024.
- PY 2023 SEM program participants **achieved an aggregated lifecycle gross natural gas savings of 5,804 thousand therms** with a statewide GRR of 118%. The GRR for lifecycle gross natural gas savings is slightly higher than for first-year gross natural gas savings, as many capital projects we reviewed have an effective useful life (EUL) exceeding five years. Examples include new boilers and furnaces installations.

These GRRs will be integrated with the results from the upcoming PY2024 study to ensure a high level of precision when both program years are combined.

[15] This evaluation result is based on a partial sample that includes MCE, SCE, and SCG, while the PY2024 sample is expected to include projects from PG&E and SDG&E. The GRRs reported in this report will be combined with the results from the upcoming PY2024 study to ensure a high level of precision.

Table 1-2 presents the electric first-year and lifecycle evaluated energy and demand gross savings and relative precisions at 90% confidence interval. Overall, the precisions were lower than anticipated due to the significant variability in site-specific results. We will account for this variability in the sample design for PY2024 by increasing the sample size and adjusting the stratification approach to achieve the MMBtu statewide target of a 90/10 confidence/precision level across both program years.

Table 1-2. Electric first-year and lifecycle evaluated gross energy savings by PA

PA	First-year				Lifecycle			
	Forecasted savings	Evaluated savings	GRR	Relative precision	Forecasted savings	Evaluated savings	GRR	Relative precision
Energy (MWh)								
MCE	471	473	100.3%	0.0%	2,357	2,622	111.2%	0.0%
SCE	29,418	17,963	61.1%	47.8%	147,091	95,076	64.6%	43.6%
Statewide	**29,890**	**18,436**	**61.7%**	**46.6%**	**149,448**	**97,698**	**65.4%**	**42.4%**
Demand (MW)								
MCE	0.1	0.1	100.3%	0.0%	0.4	0.4	111.2%	0.0%
SCE	3.6	2.0	55.2%	56.8%	18.0	10.6	58.6%	51.1%
Statewide	**3.7**	**2.1**	**56.1%**	**54.7%**	**18.4**	**11.0**	**59.7%**	**49.1%**

A discussion of the drivers of each PA's realization rates is provided below:

MCE: The evaluated electric savings for MCE were higher than forecasted due to rounding discrepancies that impacted the electric savings for one site.

SCE: The evaluated electric savings for SCE were smaller than forecasted savings, primarily due to model adjustments. We calculated evaluated savings after adjusting models to enhance statistical significance, better reflect typical operations, and improve savings accuracy. Additionally, SCE's forecasted savings overstated post-installation operating loads based on as-built data collected from facilities personnel.

Table 1-3 presents the gas first-year and lifecycle evaluated therm gross savings and precisions statewide, by PA. Similar to the electric savings summarized above, the precisions were lower than anticipated due to the significant variability in site-specific results. We will account for this variability in the sample design for PY2024 by increasing the sample size and adjusting the stratification approach to achieve the MMBtu statewide target of a 90/10 confidence/precision level across both program years.

Table 1-3. Natural gas first-year and lifecycle evaluated gross energy savings by PA

PA	First-year				Lifecycle			
	Forecasted savings	Evaluated savings	GRR	Relative precision	Forecasted savings	Evaluated savings	GRR	Relative precision
Energy (therms/1,000)								
MCE	164	74	45.1%	9.5%	821	383	46.6%	42.5%
SCG	821	1,051	127.9%	46.3%	4,106	5,422	132.0%	45.4%
Statewide	**986**	**1,125**	**114.1%**	**42.5%**	**4,928**	**5,804**	**117.8%**	**42.5%**

A discussion of the drivers of each PA's realization rates is provided below:

MCE: The evaluated gas savings for MCE were smaller than forecasted savings, primarily because MCE used unsubstantiated savings factors in its bottom-up savings calculations rather than measure- and site-specific parameters that were used in the evaluated savings calculations.

SCG: The evaluated gas savings for SCG were higher than forecasted savings, primarily because we found the post-installation operation loads to be lower than forecasted, based on as-built data collected from facilities personnel.

Net savings results

The following sections present the 2023 results of the net savings analysis, starting with electric energy and demand savings. Both first-year and lifecycle savings are provided for each PA and combined.

Table 1-4 presents the first-year and lifecycle forecasted savings, evaluated gross and net savings, NTGR, and relative precision at the PA and combined levels for gas savings.

Table 1-4. PY2023 electric net savings and NTGR

PA	First year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Energy (MWh)								
MCE[1]	473	469	99.2%	0.8%	2,622	2,592	98.8%	1.2%
SCE	17,963	17,820	99.2%	0.8%	95,076	93,966	98.8%	1.2%
Statewide	**18,436**	**18,289**	**99.2%**	**0.8%**	**97,698**	**96,558**	**98.8%**	**1.5%**
Demand (MW)								
MCE[1]	0.1	0.1	99.6%	0.6%	0.4	0.4	99.1%	1.3%
SCE	2.0	2.0	99.6%	0.6%	10.6	10.5	99.1%	1.3%
Statewide	**2.1**	**2.1**	**99.6%**	**0.6%**	**11.0**	**10.9**	**99.1%**	**1.3%**

1 No MCE sites with electric savings were recruited; therefore, statewide results were applied.

Table 1-5 presents the first-year and lifecycle forecasted savings, evaluated gross and net savings, NTGR, and relative precision at the PA and statewide levels for gas savings. These results are based on SCE results only, since MCE did not report electric savings in this cycle. Both the TDA and capital attribution research support a NTGR close to 1.0.

Table 1-5. PY2023 natural gas net savings and NTGR

PA	First year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Gas (therms)								
MCE	74	51	68.9%	18.5%	383	204	53.2%	48.5%
SCG	1,051	1,042	99.2%	0.9%	5,422	5,354	98.8%	1.5%
Statewide	**1,125**	**1,093**	**96.9%**	**1.3%**	**5,804**	**5,508**	**95.2%**	**2.5%**

Both the TDA and capital attribution research support the current NTGR of 1.0. However, both the TDA and capital attribution outcomes were poor for MCE which is apparent in the natural gas NTGR.

SRA capital measure NTGR results

We found that 19% of program claimed savings were derived from capital measures. The results show that the SEM program had a substantial influence on the installation of capital measures, with a statewide average capital NTGR value of 0.92 on a MMBtu basis, which holistically captures both electric and natural savings for each site using the SEM algorithm. This indicates that SEM's immersive nature is successfully leading customers to install more capital projects than they would have without program participation. We do note that the MCE SRA capital NTGR averaged 0.45, however MCE accounted for a small portion of overall savings, thus the impact on the SEM program NTGR was mitigated.

TDA site-specific results

Table 1-6 summarizes the NTG site-specific indicators by project. The TDA results include the Delivery score (how well the implementer delivered the SEM program to the site), the Engagement score (how engaged the site was in the SEM activities), a Combined TDA score (average of the delivery and engagement results), and a Final TDA score reflecting whether the preponderance of evidence supports that the program more likely impacted the outcome or that it did not. The table also includes the capital NTGR. The summary indicators are unweighted averages by PA.

Table 1-6. Attribution indicators by project

Site ID	Vendor	Delivery score	Engagement score	Combined TDA score	Final TDA NTGR score	Capital project NTG
SCE_01	Cascade	0.98	0.95	0.97	1	0.64
SCE_02	Cascade	0.94	0.98	0.96	1	0.93
SCE_03	Cascade	0.99	0.83	0.91	1	No Cap
SCE_04	Cascade	0.94	0.88	0.91	1	1.00
SCE_05	Cascade	0.92	0.74	0.83	1	No Cap
SCE_06	Cascade	0.88	0.75	0.82	1	0.75
SCE_07	Cascade	0.79	0.76	0.78	1	No Cap
SCE_SCG_01[1]	Cascade	0.90	0.93	0.91	1	No Cap
SCE Summary		**0.92**	**0.85**	**0.89**	**1.00**	**0.83**
SCG_01	Cascade	0.84	0.88	0.86	1	0.90
SCG_02	Cascade	0.91	0.79	0.85	1	1.00
SCG_03	Cascade	0.86	0.83	0.84	1	0.77
SCE_SCG_01[1]	Cascade	0.90	0.93	0.91	1	No Cap
SCG Summary		**0.88**	**0.86**	**0.87**	**1.00**	**0.89**
MCE_01[2]	CLEAResult	0.65	0.74	0.70	1	0.17
MCE_02[2]	CLEAResult	0.60	0.71	0.66	1	0.72
MCE Summary		**0.63**	**0.73**	**0.68**	**1.00**	**0.45**

1 No gross evaluated savings due to missing data.
2 Commercial site

The indicators are consistent with the final program NTGR results. Two of the PA programs (both delivered by Cascade) have high NTG indicators and high NTGR results, shown in Table 1-4 and Table 1-5. The MCE program, delivered by CLEAResult, received lower TDA indicator values for the two gas-only sites, and a lower gas NTGR. Interview results (summarized below) provide additional nuances to differences in delivery that support this finding. While the preponderance of evidence framework established in the PY2018/PY2019 SEM Impact Evaluation Study[16] supports the application of a 1.0 MCE's program compliance, as measured by the TDA data, is on the edge of a "failing grade." The PY2018/PY2019 framework used a preponderance of evidence standard (a 50% threshold) to establish whether the program was more likely than not to have impacted measure adoption.

In contrast to the generally laudatory comments from SCE/SCG participants, the MCE participants noted:

[16] SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, Section 3.4.2.9.1, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

- Expectations with the gas SEM program were unclear to one MCE participant, who reported disappointment with the types of opportunities identified in the Treasure Hunt. The Energy Champion expected a more detailed review of their gas-using equipment but said the implementation staff did not have the necessary expertise to provide valuable insight.
- The other MCE site was a school and there were very few opportunities for gas savings in the classrooms outside of the HVAC system. The maintenance staff was already aware that the HVAC units were over 30 years old and had plans to replace them even without the assistance of the SEM program.
- Both MCE participants were also disappointed with the granularity of the gas data that was only available at monthly intervals and did not allow for a detailed model of energy use.

The forthcoming SEM Activity Report will provide a more comprehensive overview of program implementation practices.

1.5 CONCLUSIONS & RECOMMENDATIONS

DNV developed the following key conclusions and recommendations based on all reported impact evaluation activities. These recommendations aim to inform PAs of opportunities to enhance savings estimation practices and support future evaluation efforts as the program continues to evolve and expand. Overall, the PY2023 SEM participants achieved aggregated first-year gross savings of 16,780 MWh in electricity and 1,125 thousand therms in natural gas which is lower than the achieved savings in PY2021/2022.

SEM gross savings analysis methodology findings

We observed a decline in the use of the top-down modeling approach, with fewer sites adopting it and a decrease in the percentage of savings from top-down models for electric and gas compared to PY2021/2022. The SEM M&V Guide designates top-down modeling as the preferred methodology for calculating SEM savings. While bottom-up calculations are allowed and justified in certain situations, they may obscure zero or negative savings—particularly at sites beyond the first reporting cycle, where SEM-related savings could regress. As a final point, the SEM program's special status, warranting its existing baseline and SEM NTGR, is based on Decision 16-08-019. This decision describes SEM as founded on a "whole-facility approach that uses NMEC and a dynamic baseline model to determine savings from all program activities". However, since most reported projects and savings do not employ NMEC methods (top-down modeling), this raises concerns about whether the current delivery conforms to the Decision.

We found that one of the primary reasons for the increased use of bottom-up calculations was failed top-down models, primarily due to insufficient statistical significance or lack of reliable data. While implementers review all failed models, the PAs may only review and approve the conclusions, rather than conducting their own independent review of the failed models. We found that attempted and abandoned models were not consistently provided to evaluators in the project packages for PY2021/2022 and PY2023. Therefore, it remains unclear whether unreliable data or models are provided to PAs and thoroughly reviewed by them. Additionally, some reasons for switching from top-down to bottom-up modeling were found to be non-site-specific and unjustified, which contradicts the intent of the program design and guidance provided in the SEM M&V Guides.

When savings from either savings calculation approach appear disproportionate to the measures installed, it should prompt further investigation, and consideration should be given to using the alternative approach (either top-down or bottom-up). Some SEM participants in PY2023 have already adopted this as a best practice. Furthermore, identifying actionable items, including necessary resources and timelines, to address barriers preventing top-down modeling can expedite the resolution process. This approach has also been adopted by some SEM participants in PY2023 as a best practice.

Recommendations

- Implementers should provide justifications for using bottom-up calculations that are site-specific, reasonable (in alignment with the SEM M&V guide examples), including any failed top-down models or unreliable data. PAs must conduct thorough reviews of root causes noted in bottom-up justifications before approving. This will minimize the unnecessary use of bottom-up calculations and ensures compliance with the SEM M&V guide.
- For sites using bottom-up calculations to claim savings beyond the first reporting cycle, implementers should collect and provide evidence and documentation of savings persistence from the previous cycle and include it in the project files package. Following this recommendation will allow validation of the savings persistence of SEM measures.

Savings annualization discrepancies

Savings annualization refers to prorating the savings calculated within a short period of time to represent a full year of savings. Annualization is often used when SEM projects were installed late in the year, and consequently, the full annual impact of those savings would not appear in the billing analysis. In some cases, annualization may be used when certain periods are considered unrepresentative of typical facility operations and are therefore excluded from the modeling analysis consideration.

The current version of the SEM M&V Guide limits the use of annualization to only when the model is being retired or a customer will not be participating in the SEM program after the current reporting period, with PA authorization.[17] We acknowledge that program participants who used annualization followed the guidelines of earlier versions of the SEM M&V Guide where annualization was permitted. However, the annualization approach often overlooks seasonality in the typical annual operation for facilities, which results in inaccurate savings estimation.

Overall, the use of annualization based on insufficient or unrepresentative periods of facility operation in PY2023 resulted in a 6% decrease in total forecasted savings, compared to an 8% decrease in PY2021/2022, reflecting significant improvement from PY2021/2022. This improvement is primarily due to sites using longer periods for annualization in PY2023.

[17] Sergio Dias Consulting. "California Industrial SEM M&V Guide, Version 3.02." Section 1.4. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

Recommendations

- Implementers should follow the SEM M&V guide, which recommends limiting annualization to only when the model is being retired or a customer will not be participating in the SEM program after the current reporting period, with PA authorization. Therefore, annualized savings should be rejected when annualization is being used outside of these two reasons as they are likely to produce inaccurate annual savings, such as seasonally impacted savings, or measures with that fluctuate over time, such as shutdown-type measures. Following this recommendation will result in more accurate savings estimations.
- Implementers should continue to investigate further when savings from either top-down modeling or bottom-up calculations are inaccurate or disproportionate to the measures installed and consider using the alternative savings calculation approach (either top-down or bottom-up) if the disparity cannot be explained or resolved. Following this recommendation will ensure that the claimed savings accurately reflect the performance of installed measures, leading to a more precise representation of the SEM program's impact.
- If SEM participants decide to use bottom-up calculations for the second reporting period after using the top-down modeling approach in the first, implementers should consider recalculating the savings from the first reporting period using the bottom-up calculations and truing up the savings at the end of the cycle. Following this recommendation will minimize the use of annualization based on a limited time period, reducing the risk of inaccurate savings estimate.

Top-down model discrepancies

We made top-down model adjustments that resulted in a 13% decrease in total forecasted savings, compared to a 4% decrease in PY2021/2022. We reviewed all top-down models used by SEM participants to calculate savings for projects implemented in PY2023. Several models required adjustments to improve statistical significance, better reflect typical operations, and enhance the accuracy of savings calculations. These adjustments included removing variables that showed no impact on energy consumption, adjusting changepoints of weather variables to better reflect impact of weather conditions on facilities' energy consumption, and excluding non-routine events (NREs).

Recommendations

- Implementers should continue to follow the SEM M&V guidelines on creating top-down models and assessing their validity. Following this recommendation will result in more accurate savings reporting and regulatory compliance.
- Implementers should avoid using hard-coded values in savings calculations. The use of hard-coded values prevents the participants, PA reviewers, and evaluators from tracking the sources of the used values and complicates the process of updating and validating model results. Following this recommendation will result in enhanced vetting, validation, and accuracy of the parameters and inputs used in model calculations.

Bottom-up calculation discrepancies

We reviewed all submitted bottom-up calculations for sampled projects as described in Section 3.2.5.2. We adjusted savings calculation methodologies, inputs, and parameters, based on in-depth engineering reviews and as-found data collected from facility personnel, following CPUC evaluation protocols. Overall, these adjustments to the bottom-up calculations led to a 3% increase in forecasted savings once correction were made.

Recommendations

- Implementers should continue using measure- and site-specific parameters with documented references and substantiation for all inputs, to the extent feasible. This will result in more accurate savings estimations.
- Implementers should continue normalizing baseline production and occupancy profiles based on as-built operations to result in calculated savings that reflect only installed measures and improvements.
- Implementers should include any trend or metered data used for forecasted savings estimation in project files which will result in more accurate savings impact analysis results.
- Implementers should collect invoices, photographs, and any available documentation to substantiate assumptions and parameters used in forecasted savings estimations. Following this recommendation will result in more accurate savings estimations.

Program reporting and tracking data

The unique structure of the SEM program allows a single participant to submit multiple claims over several years —typically within a two-year cycle. However, there have been instances where the cycle extends to three years, such as when a site pauses SEM participation for a year due to unexpected circumstances like turnovers or temporary shifts in priorities and resources. Therefore, as the program expands, it is essential for the CPUC staff to continue collaborating with PAs to enhance program claims reporting and tracking.

Recommendations

- PAs should continue working in collaboration with CPUC staff to enhance SEM program reporting and data tracking. This continuous collaboration will ensure that as the program expands, the data tracking systems are developed to effectively monitor and support this growth, allowing for more accurate tracking of the program's expansion and overall impact.
- The CPUC staff should continue to review, refine, and implement improvement opportunities for program tracking identified in prior studies and statewide discussions, which will result in better informed program and policy outcomes.
- The CPUC staff and PAs should prioritize exploring the addition of an “analysis methodology” field when reporting claims in CEDARS to indicate whether each claim used top-down modeling or bottom-up calculations. Following this recommendation will improve tracking of top-down vs. bottom-up methodologies, allowing for better monitoring of program trends and enhanced stratification of sampling for evaluation.

SEM NTGR methods and results

NTG methods

The SEM SRA instrument developed in the PY2021/2022 evaluation was optimized through revisions to reduce the number of program and non-program influence factors. The 2018–2019 TDA instruments were adapted for the PY2023 evaluation referencing the approved SEM Design Guide rather than an evaluator-developed Program Theory and Logic Model (PTLM).

SEM, like no other current program, seeks to change organizational behavior in a manner that produces persistent and ongoing savings through a series of prescriptive activities. TDA is an essential method for assessing whether these broad objectives have been met. However, TDA does not directly produce a NTGR. In addition, the nuanced picture produced by TDA of the site delivery and engagement is boiled down to a one or a zero. While this method was used to meet a 50% threshold of whether the program influenced the participants actions in order to earn a NTG of 1, if a PA’s program-level TDA score fails to meet the preponderance of evidence criteria, there is no lower fallback NTGR.

Finally, the Expansion Study[18] tightly linked the success of the program to following the SEM Guides. The Combined TDA Score is a measure of how well the program is complying to the letter and spirit of the Guide. Another perspective is that a project should meet a minimum standard of compliance to be considered a SEM delivery warranting the SEM NTGR. Standards for compliance or competency typically meet a higher bar than 50% which is an accepted threshold for preponderance of evidence but may be too low a threshold for compliance.

Recommendations

- We recommend using the SRA and TDA survey instruments developed and refined in this impact evaluation for the 2024 evaluation cycle, which will result in better alignment of program tracking objectives with current CPUC data tracking policies.
- We recommend the development of a non-capital, BRO-oriented SRA NTGR battery to provide additional perspectives on program influence. Following this recommendation will provide an alternate source of NTGR should a PA fail to demonstrate a high level of influence on program outcomes.
- We recommend reconsidering the algorithm threshold for converting a TDA score to a NTGR score for applications after the 2024 evaluation. Following this recommendation will result in more accurate NTGR considerations for the unique characteristics of SEM.

NTGR results

The two MCE projects (of four) missed delivery and engagement metrics and scored notably lower than the other PAs on the site-specific metrics and the gas NTGR. Notably, the worst-performing site is an MCE commercial site.

Recommendations

- We recommend that MCE consider these findings in conjunction with the forthcoming SEM Activity Report (targeting a July 2025 publication) to review the delivery with the current vendor to re-enforce the SEM Design guide. Following this recommendation will result in enhanced performance calculations for MCE commercial sites in terms of participant engagement.

[18] DNV, *Group D Strategic Energy Management Expansion Study*, Final Report, May 23, 2024. https://www.calmac.org/publications/CPUC_SEM_Expansion_Study_Final_Report_240731.pdf

2 INTRODUCTION

This report presents the key findings of the Strategic Energy Management (SEM) program impact evaluation conducted by DNV and Guidehouse (together, the DNV team) on behalf of the California Public Utilities Commission (CPUC) for program year (PY) 2023. This evaluation includes SEM projects that completed a full two-year cycle in program year 2023 (PY2023) and will account for the savings claims from the first year of the cycle reported in 2022 or earlier. This study is the first half of a two-year evaluation cycle. The results from this study will be combined with the evaluation of SEM projects completing a full two-year cycle in program year 2024 to produce an update to program gross realization rates and net-to-gross ratios.

The overall purpose of this study is **to evaluate energy and demand savings for SEM projects implemented in PY2023.** This impact evaluation quantifies the evaluated gross and net first-year and lifecycle electric and gas energy savings and peak demand reduction. The study also presents recommendations aimed at improving program delivery, the quality of documentation and savings estimation practices, and the submission of program savings claims. This evaluation effort is guided by the SEM final work plan dated September 30, 2024.[19]

2.1 Background

SEM is a unique program under the Energy Efficiency Portfolio. It has a separate statewide Design Guide[20] and M&V Guide[21] that all SEM program implementers must use. As such, the program evaluation requires adherence to the CA evaluation framework and the documented approaches approved for SEM program implementation. A top-down[22] modeling approach is the preferred methodology to calculate SEM savings. The M&V guide requires participants to provide justification if bottom-up calculations[23] are used instead of a top-down modeling approach. All SEM energy savings calculations are expected to leverage the existing conditions recorded during the baseline period. The DNV team has also aligned with the California Public Utilities Commission (CPUC) directive, as decided during the PY2021/2022 Study,[24] to review and evaluate SEM projects in a manner consistent with other custom programs in an advisory capacity, despite SEM's unique design approach.

The SEM program has primarily been available to industrial sector facilities, with a few exceptions in certain jurisdictions. However, feedback from PAs and other stakeholders expressed interest in expanding SEM to include non-industrial market sectors such as commercial, agricultural, education, and public. Therefore, the CPUC commissioned the "Group D Strategic Energy Management Expansion Study."[25] The purpose of the study was to determine:

1. The characteristics of a SEM program that achieves high net-to-gross ratio (NTGR) outcomes
2. The characteristics of successful SEM participants
3. Whether non-industrial participants could be successful

---

[19] DNV, *Evaluation, Measurement, & Verification of Program Year 2023 Strategic Energy Management Projects Work Plan*, September 30, 2024. https://pda.energydataweb.com/#!/documents/4068/view

[20] The *California Industrial SEM Design Guide* provides the program requirements for qualifying as a SEM Program. The guide includes the sequence and curriculum for the program participants and delivered by program implementers.

[21] The California Industrial SEM M&V Guide establishes an M&V process which industrial facilities as part of the SEM program must adhere to for program engagement. The framework defines the protocols to determine a participant's energy baseline, track energy performance throughout the engagement, document energy savings, and qualify the methods. M&V in SEM typically relies on a consumption-based energy model or measure-level engineering calculations

[22] **Top-down savings calculation** – A savings methodology that calculates the SEM energy savings using facility models on the site level. This methodology uses a billing analysis utilizing multivariable regressions of utility meter data along with the relevant independent variables (such as levels of production or weather conditions) between the baseline period and the reporting period.

[23] **Bottom-up savings calculation** – A savings methodology that calculates the SEM energy savings utilizing measure-level calculations. This methodology uses measure-specific formulas, inputs, and assumptions, etc., to calculate the measure-specific savings. The overall site savings are then calculated by aggregating the energy savings of all implemented measures

[24] DNV, *Group D Strategic Energy Management (SEM) 2021-2022 Impact Evaluation*, Final Report, March 5, 2024. https://pda.energydataweb.com/api/view/3970/CPUC%20Group%20D%20SEM%20Impact%20Report%20Final.pdf

[25] DNV, *Group D Strategic Energy Management Expansion Study*, Final Report, May 23, 2024. https://www.calmac.org/publications/CPUC_SEM_Expansion_Study_Final_Report_240731.pdf

DNV

4. Whether adjustments to the California SEM program would be required to accommodate non-industrial participants.

The expansion study concluded that the current CA SEM program, as described in the California SEM Guides, fosters a high level of participant engagement through a prescribed delivery patterned on other successful programs with high NTGR in other jurisdictions. The study also found that successful SEM programs in other jurisdictions include non-industrial participants using a similar program of engagement. The study recommended that non-industrial participants be included in the SEM program, with the same EUL and NTGR as the industrial sector, both following the SEM Guides.

This PY2023 round of evaluation is the third round of savings impact evaluation for the SEM program done for the CPUC. Table 2-1 summarizes the previous two rounds of SEM evaluation.

Table 2-1. Summary of previous SEM evaluations

Study title	Evaluation scope	Enrollment year and cycles
2018-2019 Industrial Strategic Energy Management (SEM) Impact Evaluation[26] (2018/2019 Study)	Evaluating the accuracy of gross and net savings the PAs claimed for SEM projects installed in PYs 2019 and 2020.	Enrollment year 2018: Cycle 1 (PY2019 & PY2020)
Strategic Energy Management (SEM) 2021-2022 Impact Evaluation[27] (PY2021/2022 Study)	Evaluating the accuracy of gross and net savings the PAs claimed for SEM projects installed in PYs 2021 and 2022. Net savings research focused on capital measures.	Enrollment year 2018: Cycle 2 (PY2021 & PY2022) Enrollment year 2019: Cycle 1 (PY2020 & PY2021) Enrollment year 2020: Cycle 1 (PY2021 & PY2022)

2.2 Evaluation objectives

The primary objectives of this study are to:

1. Develop first-year and lifecycle evaluated gross savings for the SEM program in PY2023.
2. Develop first-year and lifecycle evaluated net savings for the SEM program in PY2023.
3. Identify the reasons for deviations between forecasted and evaluated savings.
4. Develop meaningful, actionable recommendations to improve program performance in delivering energy efficiency savings.

The key research questions for this impact evaluation are as follows:

- What are the SEM PY2023 annual gross kWh, peak kW, and therm savings?
- What are the PY2023 first-year and lifecycle gross kWh, peak kW, and therm savings by sampling domain (e.g., analysis approach, sector, PA)?

[26] SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

[27] DNV, *Group D Strategic Energy Management (SEM) 2021-2022 Impact Evaluation*, Final Report, March 5, 2024. https://pda.energydataweb.com/api/view/3970/CPUC%20Group%20D%20SEM%20Impact%20Report%20Final.pdf

- What are the evaluated gross realization rates (GRR)? What factors are driving gross realization rates, and as necessary, how can realization rates be improved? What is the corresponding GRR by sampling domain?
- What is the corresponding net-to-gross ratio (NTGR) by sampling domain? Determine the factors that characterize free-ridership and support the SEM NTG related to the SEM program design, and as required, provide recommendations on how the NTGR allocation might be improved with this program design in mind.
- What factors contributed to the difference in energy impacts between forecasted and evaluated savings?
- What assumptions or assumed parameter values should be adjusted based on evaluation findings and how?
- What gaps are there, if any, in the planned evaluation, measurement, and verification (EM&V) activities for SEM programs, including adherence to the SEM design and M&V guides? What emerging evaluation issues should be addressed going forward? What are the recommended changes to the SEM M&V guide and SEM program design?
- What are the remaining or new primary challenges, lessons learned, and potential best practices for key program components and related research questions?
- What are the actionable recommendations to address gaps and improve programs and projects in the future?
- What are the actionable recommendations to support and improve future evaluation activities?

2.3 CPUC policies and guidance

In designing and implementing this evaluation, the DNV team considered the following guidance documents and CPUC policies that were in effect at the time of project approval:

- The California Industrial SEM Design Guide
- The California Industrial SEM M&V Guide v2.01
- The California Industrial SEM M&V Guide V3.02
- Energy Intensity Model Guidelines v2.02
- ASHRAE Guideline 14-2014
- CPUC Energy Efficiency Policy and Procedures Manual Version 6
- PA-specific program policy and procedures manuals
- Energy Efficiency Industry Standard Practice (ISP) Guidance v. 3.1
- Fuel Substitution Technical Guidance for Energy Efficiency V2.0
- CPUC resolution E-4867 approving the DEER updates for 2020
- CPUC resolution E-4952 revising DEER update for 2020
- CPUC resolution E-4818 affecting assignment of project baselines
- CPUC D.19-08-009 Fuel Substitution Decision

3 METHODOLOGY

This section presents the methods the DNV team used to fulfill the evaluation objectives listed in Section 2.2, including the planned sample design, achieved sample sizes, gross savings, measurements and verification (M&V) activities, net savings approach, and final expansion procedures.

The DNV team reviewed 14 gross sample and unique customer sites and 13 net sample points. We assessed the provided project files for those data points, conducted phone interviews to verify project specifics, reviewed billing data and model parameters, and collected site-specific trend data and photographs, when applicable. The net evaluation combined separate capital and non-capital data collection methods to determine overall program-level NTG results. Capital measures used a self-reported attribution (SRA) approach to estimate a capital savings NTGR. The non-capital measures used a theory-driven attribution (TDA) in-depth interview approach to determine a non-capital savings NTGR. Both gross and net evaluation results are presented in Section 4 of this report

3.1 Sample design

The SEM evaluation will span two years, PY2023 and PY2024. Thus, sampling occurs in two waves, one wave for each year. The first wave sample was drawn using final PY2023 claims data. The second wave sample, to be conducted in 2025, will be drawn using the final PY2024 claims data. The two-wave sampling approach will allow the DNV team to achieve the precision target of ±10% at the 90% confidence interval over two years.

This report focuses on Wave 1 SEM projects that completed a full two-year cycle in PY2023, which are made up of savings claims from both PY2022 (Year 1) and PY2023 (Year 2). To achieve the two-year precision target of 90/10, each of the two evaluation waves will need to achieve a relative precision of ±14% at the 90% confidence interval.

SEM is unique in that its delivery is designed to engage customers over a six-year period, structured into three two-year cycles. Participants are grouped into “cohorts” defined by their first year of engagement. Each of the three cycles includes distinct planning and implementation activities over a two-year period. Often, projects completed in Year 1 are further developed or built upon in Year 2. This study will evaluate all claimed savings for the nominal Cycle 1 and Cycle 2 participants, which are shown in Table 3-1. This approach aligns with the program design and activity plan. Projects that are nominally eligible for the PY2023 impact evaluation are completed Cohort 2/Cycle 2 and Cohort 4/Cycle 1 projects. However, some observed variations in program delivery and participant pace have resulted in some cycles moving between calendar years.

Table 3-1. Cohort and cycles with evaluation reporting periods

Enrollment cohort	2019	2020	2021	2022	2023	2024	2025
1 – 2018	Cycle 1		Cycle 2		Cycle 3		
2 – 2019		Cycle 1		Cycle 2		Cycle 3	
3 – 2020			Cycle 1		Cycle 2		Cycle 3
4 – 2021				Cycle 1		Cycle 2	
5 – 2022					Cycle 1		Cycle 2
6 – 2023						Cycle 1	

The design of the SEM program introduces some complexities in tracking and related project documentation, as each unique customer may claim savings for a single engagement, annually, across six years. Evaluation of one cycle requires consideration of the previous cycle. In addition, sampling requires knowledge of the cycle and cycle-year, parameters that are not currently captured in California Energy Data and Reporting System (CEDARS) data, but these will be updated and become available starting in PY2025.

3.1.1 Gross and net savings sample design overview

Table 3-2 presents a summary of the SEM sample design approach for this study.

Table 3-2. SEM sample design assumptions and approach

Parameter	Description
Population	Tracking data set for the program year, aggregated at the cycle level for each participant. PY2023 wave: 2022 Year 1 claims + 2023 Year 2 claims
Explicit sampling strata	PA, Size of savings (MMBtu)
Gross sample allocation	15 projects, allocated for best overall precision while targeting 90/14 results by fuel type and 90/14 overall (MMBtu).
NTGR sample allocation	Separate sample allocation, starting by attempting NTGR surveys for all projects in the gross impact sample.
Target parameters	GRR, NTGR
Analysis domains	PA, Fuel (electric or gas)
Error ratios	Assumed value of 0.4.
Projected precision at 90% confidence (based on current error ratio assumptions)	Gross MMBtu savings by energy unit (electric): ±14% Gross MMBtu savings by energy unit (gas): ±14% NTGR by electric fuel type: ±14% NTGR by gas fuel type: ±14%
Savings size stratification	Custom – up to 3 levels based on savings, depending on the number of samples in the cell.
Contingency and back-up sample	Gross impact and NTGR sample: A prioritized list of backup projects was created for this study to support any valid dropped sample points or refusals.

The DNV team finalized the SEM population after performing extensive data cleaning to identify Year 1 and Year 2 savings claims across multiple program years. Each participant's Year 1 and Year 2 savings were aggregated to represent a completed two-year cycle. The DNV team used forecasted savings calculated by removing the default GRRs that the system had applied in calculating the savings reported in the ED tracking data. Table 3-3 provides a summary of the SEM program population determined for this study after aggregating each participants' Year 1 and Year 2 savings. Note that some participants included both electric and gas savings measures, while most had measures that impacted only one of the fuels. Neither PG&E nor SDG&E had any SEM participants who completed a two-year cycle in PY2023, so they are not represented in this sample. However, they are expected to be represented in the second phase of the study.[28]

Table 3-3. SEM PY2023 population summary

Evaluation cycle	# unique participants	# electric	# gas	FY MWh savings	FY MW reduction	FY Mtherms savings	FY MMBtu savings
MCE	4	2	3	471	0.1	164	18,033
SCE	31	31	N/A	29,418	3.6	N/A	100,379
SCG	9	N/A	9	N/A	N/A	821	82,126
Total	**44**	**33**	**12**	**29,890**	**3.7**	**986**	**200,538**

[28] The SEM participating sites for PG&E and SDG&E are grouped by each Program Administrator to complete their cycles every other year, with PY2024 being the next scheduled completion year. Therefore, projects from these utilities will be included in the PY2024 sample.

3.1.2 Gross sample completions and response rates

Table 3-4 presents the population counts, sample design quotas, and final achieved sample for key analysis dimensions, including PA and fuel, for the gross sample. The sample design quotas were designed to meet the annual precision targets of ±17% relative precision for both electric and gas first-year energy savings. Achieving the precision targets of ±17% relative precision for each program year (PY2023 and PY2024) will result in meeting the combined cycle precision target of ±10% relative precision for both PY2023 and PY2024. Overall, 80% of electric projects and 120% of gas projects in the primary sample design were recruited. This occurred because the sampling was based on MMBtu; in some cases where electric projects could not be recruited, the replacements -selected based on random priority- happened to be natural gas projects. While all gas projects were initially recruited, one site was unable to provide the required data late in the evaluation, preventing its completion. We plan to implement oversampling in the PY2024 sample design, based on the population size, to mitigate such incidents.

Table 3-4. PY2023 gross sample response rate by fuel and PA

PA	Electric				Natural gas			
	Population (N)	Sample design quota	Final sample (n)	% complete	Population (N)	Sample design quota	Final sample (n)	% complete
MCE	2	0	1	N/A	3	1	2	200%
SCE	32	10	7	70%	N/A	N/A	N/A	N/A
SCG	N/A	N/A	N/A	N/A	9	4	4	100%
Total	**34**	**10**	**8**	**80%**	**12**	**5**	**6**	**120%**

The final evaluation will include results from program years 2023 and 2024. The results from the 2023 analysis will guide any adjustments needed from the original sample design based on the 2023 program cycle evaluation. Table 3-5 presents the results of the gross data collection. A total of 14 sites were evaluated across the three PAs. For the interim analysis, a non-stratified analysis approach was used due to the small sample sizes for each PA. When the final analysis is run after the program year 2024 data is collected, the two years will be combined into a single analysis and the results will be post-stratified. At this point, DNV will review the optimal stratification based on the final data collected.

Table 3-5. PY2023 gross sample post-stratification

PA	Stratum	Strata maximum	Accounts population	Tracking savings MMBtu	Accounts sample	Weight
MCE	1	13,113	4	18,033	3	1.33
SCE	2	15,324	29	100,379	7	4.14
SCG	3	17,550	9	82,126	4	2.25
Total			**42**	**200,538**	**14**	

3.1.3 Net sample completions and response rates

The NTG sample for this study targeted all the sampled sites for gross impact. The planned research entailed interviews with customer and vendor staff associated with the selected projects. Table 3-6 summarizes the dispositions by their role in the program. The first three roles are defined in the SEM Design Guide. The purpose of the interviews was to collect data to support the capital SRA and TDA research and included an additional battery supporting a separate process study in order tounderstand their design and effectiveness and ensure consistent compliance with SEM principles.

Table 3-6. Disposition of interviews for SRA and TDA methods

Interviewee type	Interview purpose	Method	2023 target	Completed
Energy Champion. Responsible for the success of the SEM program at the site. Coordinates with the SEM Coach and internally.	Feedback from Energy Champions confirmed the evaluated capital projects, and informed the NTG algorithms, the TDA delivery and engagement scores, and the process findings.	TDA and SRA	15	13[1]
Executive Sponsor. The highest-level manager available at the site (typically the facility manager), who ensures resources for success.	Executive Sponsor feedback informed the TDA delivery and engagement scores as well as the process findings for improving the SEM program	TDA and SRA	5	5[2]
Energy Team. A cross-functional team (i.e., management, production, procurement, maintenance, HR) that meets regularly to manage and business practices and activities.	Energy Team, although in secondary roles, provides feedback to inform portions of the TDA and to support process findings.	TDA only	20	1[3]
Energy Coach. The lead staff member from the implementation team in charge of interacting with the Energy Champion at each site and driving SEM engagement (not SEM Guide defined).	Energy Coach feedback informed the TDA delivery and engagement scores as well as the process findings on how well the program is working in California.	TDA only	10	8 coaches, covering14[4] sites

1. Two Energy Champions provided contact information for another individual at the site but only one answered our interview questions. One Energy Champion declined the interview, and one did not respond to multiple email inquiries, even when copying the CPUC, PA, and Implementer contact personnel.
2. The team completed 6 interviews with Executive Sponsors but one covered a site where the Energy Champion did not participate in our research. The team will use feedback from the sixth Executive Sponsor in the process analysis but not the TDA scoring.
3. The team offered a $100 gift card to Energy Team members who responded to the 10-minute online survey and also asked participating Energy Champions to encourage team members to provide feedback. Energy Team member contacts were not included in the project file and plant staff turnover complicated recruiting, thus only one survey was completed.
4. The team successfully interviewed the Energy Coaches for all sites which also provided Energy Champion feedback; therefore, were able to understand the delivery and engagement scores from both the customer and the implementer.

The net analysis will use the same methodology as the gross analysis to combine the two years of evaluation data into a single data set for analysis. Table 3-7 presents the net analysis post-stratification table for the PY2023 results. Twelve sites were evaluated across the three PAs; two gross sites could not be recruited for the net analysis. For the interim net analysis, the same non-stratified analysis approach that was used in the gross analysis was used due to the small sample sizes for each PA. When the final analysis is run after the program year 2024 data is collected, the two years will be combined into a single analysis and the results will be post-stratified. At this point, DNV will review the optimal stratification based on the final data collected.

Table 3-7. Program year 2024 net sample post-stratification for NTGR

PA	Stratum	Strata maximum	Accounts population	Tracking savings MMBtu	2023 target	Completed	Sites with capital	Weight
MCE	1	13,113	4	18,033	3	2	2	2.00
SCE	2	15,324	29	100,379	7	7	4	4.14
SCG	3	17,550	9	82,126	4	3	3	3.00
Total			**42**	**200,538**	**14**	**12**	**9**	

Table 3-7 shows the number of sites with capital measures. The SRA NTG surveys were only conducted with sites that installed capital measures. Non-capital measure savings were given the NTGR of 1.0, which is the result of the 2023 TDA analysis as presented later in this report. For sites with capital measures, SRA and TDA surveys were attempted with both the Energy Champion and the Executive Sponsor at each sampled site, as well as reviewed program documentation as described below.

3.2 Gross savings methods

This section describes the DNV team's approach to evaluating gross savings.

3.2.1 Methods overview

The team determined the appropriate evaluation approach for each site based on a review of project documentation and data collected from the site contacts after the initial discussion. The team presented all site-specific M&V plans and evaluation findings in the custom core template (CCT) tool. The following subsections provide more details on this approach.

3.2.2 Custom core template (CCT)

The team used the Excel-based CCT—also used by other evaluation teams, such as CIAC—to organize and communicate evaluation information for each sample project selected. The CCT served as the final site-specific deliverable for evaluated savings and was the primary reference for the engineering team to create M&V plans and document data used in impact estimates.

The CCT stored claim information from the tracking database, organized M&V activities, savings calculation methods, supplemental data, energy model references, and realization rate determination in a standardized format. This consistency ensured adherence to CPUC guidelines and supported the systematic application of best practices for pre-implementation review and evaluation.

Upon completion of evaluation activities for each PA's sampled sites, we presented the site-specific results and findings to the relevant PA. We gave each PA the opportunity to select approximately 30% of their evaluated sites for a focused review of the evaluation methodology and findings. We addressed all comments and questions received, making adjustments as needed before finalizing the results. This review process took approximately two to four weeks, depending on the number of sites included in each PA's sample.

3.2.3 Project documentation review

The DNV team conducted a comprehensive review of the project files for the 15 sampled sites. For each site, we evaluated the SEM participants' calculation methods, assumptions, inputs, project documentation, and savings claims to ensure their appropriateness and adherence to the SEM M&V Guide in effect at the time. We relied on the following documents as the foundation for this evaluation study:

1. **Opportunity register:** Used to list all measures targeted by SEM participants, including completion statuses, installation dates, and impacted systems. We observed that while some opportunity registers reported savings and cost estimates for certain measures, this was not done consistently.
2. **Calculation files:** SEM participants used various methods to estimate savings for each reporting period, including site-level top-down models, measure-level bottom-up calculations, and demand savings calculators. We reviewed the savings reported in these calculators and the program-claimed savings to identify any tracking errors. Additionally, the team assessed the calculation approach, inputs, variables, parameters, and results for each site to ensure their appropriateness and adherence to the SEM M&V Guide in effect at the time.
3. **Completion report/reporting performance period report (RPPR):** Used to summarize sites' SEM activities for each reporting period. We reviewed the reported savings and list of installed measures to ensure alignment with the provided calculation files and opportunity register. However, this report was more frequently absent from the project packages

provided by the PAs in this evaluation round than in the previous (PY2021/2022) round, posing significant challenges for evaluators in obtaining the necessary background information on the SEM projects.

4. **M&V report or energy savings report:** Used to provide participant notes on the data and inputs used in the SEM savings calculation models, including, but not limited to, non-routine events (NREs), annualization considerations, variable range validity, and other relevant data observations. This document is typically provided for projects completed before PY2022, which is why evaluators encountered fewer of them in the project files for this round of evaluation.
5. **No model memo/Notification of Bottom-up (NOBU):** Used to provide the participant's rationale for using bottom-up calculations instead of a model. We reviewed the rationale provided by each participant that used bottom-up calculations to determine their validity.
6. **Technical review:** Used to record the findings of any peer reviews conducted to validate and support the savings model's inputs and findings, ensuring that the model parameters meet statistical significance requirements.
7. **Utility bills:** Used to verify the participation of SEM participants in the public purpose program (PPP).

The DNV team reviewed additional project documentation when more details were needed to supplement the documents listed above. We requested any missing files directly from the PAs and program implementers when necessary.

3.2.4 Recruitment and data collection

The DNV team used the same recruitment protocol that was developed and accepted by PAs during the PY2021/2022 Study. Prior to the start of the recruitment and data collection process, we reached out to each PA to establish a PA-approved communication protocol. We shared a proposed recruitment cover letter with each PA and allowed them the opportunity to comment and make recommendations. We also shared the list of sampled sites for each PA and their facility contacts, as provided in the project documentation. The PAs and program implementers supported the team's recruitment efforts in several ways, including:

- Answering participants' inquiries about the SEM evaluation process and requirements
- Making introductory calls connecting the DNV team and sampled sites
- Providing updated contacts in cases of personnel changes or turnovers
- Providing context and more information about facilities in cases of changes in their SEM participation or ownership changes
- Supporting the DNV team's data requests from participants, when requested

The DNV team started the recruitment process for each sampled site upon the completion of a site-specific M&V plan. We used the "Measure List" tab in the CCT to import the list of projects noted as completed in the Opportunity Register for each reporting period of the cycle under evaluation. For each completed measure, we planned to answer the following questions:

1. Was the measure installed as described in the Opportunity Register and project documentation? If not, why was the measure not installed?
2. Is the measure still in operation and delivering the expected level of energy savings? If not, when did the measure stop realizing savings, and why are savings not align with expectations? What are the reasons for the measure discontinuance?
3. Does the measure impact a single IOU meter? This information will be used to cross-check against meter data included in the models and project documentation.
4. Is the measure capital?

In addition to the participants' answers to the questions above, the DNV team collected any additional measure-specific information or customer feedback on their program participation.

For bottom-up and mixed[29] sites, we aimed to collect additional measure-specific data as needed. This included verification of operation parameters, trend data, equipment nameplates, photographs of equipment and setpoints, and facility operation and shutdown schedules.

For top-down sites, we aimed to collect more information on any observations or questions developed during the initial model review. This included NREs, explanations for any unexplained energy consumption spikes or drops, shutdowns, data points removal or adjustment, negative or zero savings claims, and any capital measures removed from the model savings.

Recruitment efforts started with an introductory email sent to prospective participants. The team attempted to reach the participants by phone at different times of day and different days of the week to maximize contact success. We used each M&V plan to guide site contact interviews to collect updated parameters for the savings calculations. The sample contained projects with multiple measures installed.

Recruitment efforts started in mid-October 2024 and concluded in early January 2025. These efforts engaged one lead recruiter and one active recruiter. Most of the sites recruited were within the expected number of attempts, ranging between one and five outreaches, with an initial email that included a description of the site. When recruitment efforts were unsuccessful due to departure of the listed primary contact or customer non-responsiveness, our team enlisted assistance from the 3P implementors to obtain updated contact information or facilitate contact. Cases were escalated to the PAs and CPUC as necessary.

We were able to fulfill our target of 15 sites recruited. Of these sites, 11 were part of the primary sample and 4 were back-up sites. After three unsuccessful attempts at outreach, our team enlisted assistance from the 3P implementors and began recruitment on a back-up site concurrently to ensure fulfillment of the target sample within the established timeframe.

The PAs and CPUC helped fulfill data requests after initial interviews when customers became non-responsive. After three to four unsuccessful attempts to obtain requested data, the PAs conducted outreach and leveraged their existing relationships to procure the information needed to complete the analysis. Four requests escalated to PA and CPUC involvement, and we were able to obtain data for three sites within the established timeframe. The team had difficulty obtaining data from the remaining site (SCG_7638.P5) despite assistance from SCG and the CPUC, which led to the inability to calculate site-specific results for. Hence, it was dropped from the overall PY2023 study results calculations. We will aim to collect the necessary to incorporate this site into the final combined results for PY2023/24.

3.2.5 Site analysis methodologies

This subsection addresses the site-specific analysis methodologies used to evaluate the savings forecasted by program participants.

The DNV team categorized the savings calculation methodologies that program participants used to calculate forecasted savings (for both electric and natural gas) into three analysis methods, as summarized in Table 3-8. Descriptions of each method are provided after the table.

[29] Mixed sites are sites that used both top-down and bottom-up analysis methods in some way. Typically, this would involve a bottom-up analysis approach in Year 1 of a cycle and a top-down modeling approach in Year 2 of that cycle. In some cases, this was done when estimating savings for both fuels where one fuel would use one approach and the other fuel would use the other to estimate savings.

Table 3-8. Breakdown of savings calculation methodologies

Parameter	Top-down	Bottom-up	Mixed analysis (top-down & bottom-up)
Reviewed projects, n=14	4	6	4
Percentage by count	29%	43%	29%
Percentage of electric savings	1%	40%	59%
Percentage of gas savings	25%	46%	29%
Percentage of overall savings (MMBtu)	13%	43%	44%

Top-down: This methodology calculates SEM energy savings by applying facility models at the site level. This process involves conducting a billing analysis using multivariable regressions of utility meter data, incorporating site-specific relevant independent variables (e.g., weather conditions or production levels) to compare the baseline and reporting periods. As outlined in the SEM guide, this method is recognized as the preferred approach for calculating SEM energy savings.

Bottom-up: This methodology calculates energy savings at the measure level. It applies measure-specific formulas, inputs, and assumptions to determine savings for each individual measure. The total site savings are then derived by aggregating the energy savings from all installed measures.

Mixed analysis: This methodology combines both top-down and bottom-up approaches to calculate energy savings. Typically, different calculation methodologies are applied to calculate savings for each fuel type or reporting period.

For each sampled site, the DNV team applied the same savings calculation methodology used by the participant (top-down, bottom-up, or mixed). The subsections below outline the specific tasks completed by the DNV team for each savings calculation methodology.

3.2.5.1 Top-down models

The DNV team followed the evaluation process outlined below for sites that calculated forecasted savings using the top-down methodology:

- Reviewed the statistical significance of top-down model parameters to ensure they fell within the required range.
- Verified that the provided utility billing data corresponded to the baseline and reporting periods as specified in the project documentation.
- Ensured the selected relevant variables adhered to the SEM M&V guidelines.
- Identified any model adjustments, such as the removal or modification of data points. In these cases, the DNV team determined whether to follow the participant's approach or adjust based on documentation or additional feedback from the site interview.
- Reviewed annualization methods, reasoning, and periods, when applicable, to ensure compliance with SEM M&V guidelines.
- Verified that the top-down model appropriately accounted for seasonality based on the type and schedule of site operations.
- Ensured the model accounted for any non-SEM projects completed during the baseline and/or reporting periods.
- Conducted measure-specific data collection for all measures marked as "Completed" in the provided opportunity registers, as described in Section 3.2.4. Note that top-down models calculate energy savings at the site level.

3.2.5.2 Bottom-up calculations

The DNV team followed the evaluation process outlined below for sites that calculated their forecasted savings using the bottom-up methodology:

- Conducted measure-specific data collection, as described in Section 3.2.4, for all measures marked "Completed" in the provided opportunity registers.
- Sampled the two highest savings measures for an in-depth review. If any completed measures were categorized as capital, we selected the highest savings capital measure as one of the two sampled.
- Performed a detailed review of the participant's engineering approach for the two sampled measures, including their methodology, formulas, assumptions, and inputs. If the participant's calculations were deemed appropriate, the DNV team applied the same approach with updated inputs based on data collected from site personnel. Otherwise, we adjusted the analysis as needed to enhance the accuracy of the estimated savings.
- Adjusted the overall site savings by removing the savings from measures that were not installed. For measures that were installed but later removed within the SEM program's five-year EUL, the DNV team prorated the savings.

3.2.5.3 Demand savings calculation

Program participants claimed demand savings for all electric energy savings. Consistent with SEM projects from PY2021/2022, participants continued to estimate demand savings for projects completed in PY2023 using two different demand calculators, as outlined below:

1. **The SEM Demand Calculator:** This calculator uses publicly posted load profiles documented by PAs to determine the summer peak hours (320 hours for all PAs) and percentage of kWh on-peak (4.07%, 4.25%, 4.11% for PGE, SCE, and SDG&E, respectively). This calculator was used by SCE.
2. **The SEM-NMEC Demand Savings Calculator:** This calculator uses the load shapes of the facility to calculate its demand savings. The calculator determines the summer peak hours by PA (742 hours for SDGE, and 786 hours for both PGE and SCE). This calculator determines the appropriate kWh summer on-peak percentage based on the facility's sector (commercial, industrial, or agricultural) and its type of operation (refrigeration, HVAC, lighting, etc.). This calculator was used by MCE.

According to the PY2018/19 Study, *"demand savings calculation help provide a savings metric for facility-level projects that incorporate different savings types from different resources (e.g., gas and electric, energy efficiency, demand response, and distributed generation)"*.[30] The DNV team notes that both demand calculators use the overall electric energy savings per reporting period to calculate demand savings. This approach does not account for the varying application types of installed measures and their potentially different impacts on overall demand, which could result in an inaccurate estimation of demand savings. However, we acknowledge that the SEM M&V guide, in effect at the time, did not provide program participants with additional specific guidelines for calculating demand savings. We also recognize that the reported demand savings were not considered in the determination of program performance-based incentives. Hence, we concluded that the calculators were used appropriately by participants, as per the program's instructions. As originally intended, the evaluation team will consider optimizing the Demand Savings Calculators in future activity reports.

Accordingly, we calculated the evaluated demand savings using the same calculator the participants used and updated only the electric energy savings input to use the evaluated savings instead of forecast claimed savings.

---

[30] SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

3.3 Individual measure savings

Program savings are estimated at the facility level; however, savings must be allocated to individual measures to support the NTGR calculations and to inform other analysis such as the allocation of savings between top-down and bottom-up methods and EUL implications.

The DNV team reviewed the list of measures provided in the opportunity registers for each sampled site. The registers were consistent in noting the measure classification (whether BRO or capital). We then conducted interviews with program participants, as described in Section 3.2.4, to verify the list of completed measures and confirm their classifications.

Following data collection, we allocated the estimated savings for each implemented measure. For sites using bottom-up calculations, we used the measure-level calculated savings. For sites using the top-down approach, we performed the following tasks to estimate savings for each measure:

- The opportunity registers generally included a classification indicating whether the savings impacts of each measure were considered high, medium, or low. We assigned a savings grade of 1, 2, or 3 for low, medium, and high impacts, respectively.
- The opportunity registers also typically noted the type of fuel impacted by each measure. In cases where this information was missing, we used engineering judgment to determine whether the installed measure impacted electricity, gas, or both.
- We then calculated weighted savings for each measure, based on the overall site's forecasted savings per fuel type and the savings grade assigned to each measure.

3.4 EUL methods

After estimating the forecasted savings for each measure, we focused on updating the EUL for each. For non-capital projects, we maintained the SEM EUL of five years. For capital measures, the team reviewed the Remote Ex Ante Database Interface (READI) to determine the appropriate EUL. If a capital measure was unique and the EUL could not be identified, the pre-existing SEM EUL of five years was applied.

3.5 Net savings methods

The methodological approach for the present evaluation is informed by Decision D16-08-019, which states that a well-designed SEM program's holistic and long-term approach encourages the implementation of BRO measures as well as custom and capital measures. The decision concludes that capital measures, when program influence is demonstrated, may apply the SEM default NGTR, which has been 1.0. In this evaluation cycle, two evaluation methods for net savings were employed:

- **Self-report attribution (SRA) for capital measures**. Under the SEM framework, capital measures may apply SEM NTGR of 1.0 "when program influence is evident." The SRA is well suited to the assessment of capital measures that typically entail installation of equipment and a structured internal decision-making process which can be investigated through interviews of decision makers. The influence of the capital measures was determined using survey instruments and methods approved in the PY2021/2022 evaluation with the improvements shared with stakeholders in a SEM Programs - Bi-Monthly ED/PA Update Meeting y.
- **Theory-driven attribution (TDA)**. The theory-driven approach was used to evaluate the overall design and delivery of the SEM program and non-capital measures NTGR, since it is well suited to the organization transformation intent of the SEM program. The methods and survey instruments developed in the 2018/2019 SEM Study were adapted with further improvements and optimization.

3.5.1 We also note that the SEM Activity Report (targeting a July 2025 publication) will leverage the NTGR interviews and related data collection. SEM programs have recently experienced growth, an expansion into the non-industrial sector, and the addition of third-party program implementers administering these programs, The purpose of the Activity Report research is to understand the current state of the program design and effectiveness to ensure consistent compliance with SEM principles in the face of these changes. SRA capital project-level NTGRs

The survey instrument and algorithms for scoring the survey responses follow California's standard Nonresidential NTG framework, comply with the California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals and the CPUC's Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches, and were further refined after an internal research process.

The specific survey instruments and algorithms used in this study were developed in the PY2021/2022 SEM Impact Evaluation Study. These were adapted from the Commercial, Industrial, and Agriculture Custom (CIAC) NTG survey battery.[31] However, the length of the instrument made it onerous for participants. The lists of program and non-program influences were too long and overlapping, and thus unlikely to reflect the responder's actual ranking of factors. The DNV team streamlined this battery of questions as part of the Group D NTG Methodological Update. These proposed revisions were reviewed and approved by CPUC Staff and shared with PAs during a statewide ED/PA Update Meeting.

The protocol for identifying capital measures was defined in the PY2021/2022 Study. The protocol starts with a review of the Opportunity Register to inventory measures classified as a capital measure as noted in Section 3.4.1. The capital classification is then confirmed by customer staff in the NTG battery. For projects with more than one capital project, separate NTGRs were calculated for each capital component (up to two) and then combined, proportional to the savings of each measure, into a composite capital NTGR

3.5.2 TDA methods

This section first describes the background of the methodology followed by the implementation approach.

3.5.2.1 TDA background

The 2018/2019 Study[32] established TDA as the most appropriate methodology for determining whether the SEM programs were exhibiting program influence broadly. As noted in the report, theory-driven attribution (TDA) research was designed with an "aim to demonstrate a reasonable association **… between the SEM program activities and the impacts that occurred** … [and] with a reasonable degree of confidence." The theory-driven methodology reviewed the Program Theory and Logic Model (PTLM)—built by the evaluator using input from each participating PA—to identify key linkages between SEM program activities and the program expected outcomes and gathered data from program actors to verify the strengths and validity of these linkages. The study concluded that there was "strong support, via a preponderance of the evidence, that customers' decisions to implement energy-efficiency improvements in Industrial SEM aligned with the motivations designed within the logic models for each PA" and was therefore able to assign the NTGR ratio of 1.0 to the SEM program.

Methodology similarities to the 2018/2019 Study. As in the 2018/2019 Study, individual scores were developed for a sample of completed projects. Evidence for scoring was collected through in-depth interviews of multiple stakeholders and

---

[31] DNV, *Group D Evaluation, Measurement, & Verification of Program Year 2020/21 Commercial, Industrial, and Agricultural Custom Projects, Final Work plan,* May 20, 2022. https://pda.energydataweb.com/#!/documents/2629/view

[32] SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

from associated project files. The evidence was also organized into thematic groupings, which were scored. Like the prior study, preponderance of evidence criteria was applied to determine the final program TDA outcome.

Adaptions made to 2018/2019 Study methods for the PY2023 Study. This PY2023 Study built upon the framework of the 2018/2019 Study, optimizing and improving it by incorporating subsequent research, most notably the "Group D Strategic Energy Management Expansion Study"[33] ("Expansion Study"). This study was commissioned by the CPUC as directed by Decision 23-02-002[34] to determine whether the SEM NTGR (now at 1.0) and five-year EUL assumptions would be appropriate for non-industrial SEM and to develop recommendations for non-industrial SEM programs.

The Expansion Study concluded that the current CA SEM program as described in the California SEM Guides fosters a high level of participant engagement through a prescribed delivery patterned on other successful programs with high NTGR in other jurisdictions. The study also found that successful SEM programs in other jurisdictions include non-industrial participants using a similar program of engagement. The study recommended that non-industrial participants be included in the SEM program, with the same EUL and NTGR as the industrial sector, both following the SEM Guides.

The 2018/2019 Study compared PA program delivery to a PTLM developed by the evaluator. In this evaluation, PA program delivery is compared to the requirements identified in the SEM Design Guide based on the recommendations and findings of the Expansion Study. The Expansion Study conclusions enforced the importance of program delivery and indicated that if participants were actively engaged in the prescribed program, high NTGR outcomes would be achieved. Thus, the SEM Guidebooks are the benchmark for the participant program experience. This study relied on objective accounting of participant engagement and customer reports of the quality of the delivery. It should be noted that the PTLM serves multiple purposes and is still a valuable tool in ensuring program alignment with the SEM guides and should be included in with SEM guide updates.

3.5.2.2 Implementation specifics

We began with the site-specific activity requirements defined in the Design Guide and categorized them into seven Design Guide Requirements (DGRs)[35]. The in-depth interview guides were designed to determine for each of the categories 1) how well the implementer delivered the SEM program, and 2) how well the customer engaged with the program. We rated the strength of evidence, on a 1–3 scale (strong evidence = 3, moderate evidence = 2, and weak evidence = 1) that the DGRs were being met at the site. The Delivery score is an indication of how well the delivered program meets the SEM Guide requirements. The Engagement score is an indication of how well the customer engaged with the SEM program by providing staff resources and support to program activities, and how well the implementer was able to sell the SEM program to gain buy-in from each facility.

Site scoring. We identified seven requirement categories from the Cycle 1 and Cycle 2 Design Guide tables and assigned a strength of evidence score to each category for each site. The team then calculated a combined TDA score as the sum of the seven evidence scores divided by the total possible evidence score of 21 (the maximum score of 3 multiplied by 7), thus producing a Combined TDA Score value of between 0.0 and 1.0.

Program scoring. The Combined TDA score is not the same as an attribution NTGR. Instead, it measures how well the delivery complies with the SEM Guide. This is both the underlying rationale of TDA and consistent with the

---

[33] DNV, *Group D Strategic Energy Management Expansion Study*, Final Report, May 23, 2024. https://www.calmac.org/publications/CPUC_SEM_Expansion_Study_Final_Report_240731.pdf

[34] CPUC. Rulemaking 13-11-005-*Decision Addressing Energy Efficiency Third-Party Processes and Other Issues.* February 2, 2023. Decision number 23-02-002. https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M501/K931/501931085.PDF

[35] The DGRs are listed in Table 4-11 with further descriptions in the Appendix.

recommendations of the recent Non-Industrial Expansion Study which found a strong correlation between high program NTGR and the program activities specified in the SEM Guides.

Following the methodological precedent of the 2018–2019 impact evaluation, each site outcome was tested using a preponderance of evidence (POE) to determine whether there was sufficient evidence that the customer experienced the SEM prescribed delivery or not. This POE framework has precedence to determine, for example, whether an opportunity can claim an accelerated replacement measure application type.[36] In this test, each site TDA score was compared to the POE threshold of 0.5, where a TDA falling below 0.5 did not meet the evidence test, thus receiving a 0 and where at 0.5 or above, passed the evidence test, thus receiving a 1. While the TDA score is not equivalent to an attribution score derived from this, PA-specific and statewide program NTGR was calculated as the sum of the POE result (a 1 or a 0) weighted by life-cycle savings of the site.

The evaluation included a sensitivity analysis, which is reported in Appendix A to examine impacts of revising the SRA algorithm to better account the measure history and in revising the POE threshold.

3.5.3 NTG data collection

The NTG evaluation team conducted in-depth interviews with three different stakeholder groups involved in the implementation of SEM, 1) Energy Champions, 2) Executive Sponsors, and 3) Energy Coaches. An online survey was planned for a fourth group, the Energy Team. These four roles are defined in the SEM Guides.

Energy Champions – facility employees driving program participation at the site. We conducted 90-minute interviews with Energy Champions to gather feedback on capital project NTGR results and to inform the TDA research. The NTG data collection process was fully compliant with the CPUC's Self-Report Guidelines. The lengthy interviews with Energy Champions were also a major contributor to understanding how the SEM program is received by facilities and what hurdles participating sites experience when engaging with SEM activities.

Executive Sponsors – upper facility management responsible for allocating capital dollars and supporting the SEM effort. The evaluation team conducted 60-minute interviews with upper management staff, defined through SEM as Executive Sponsors, from a small sample of facilities participating in SEM. We only interviewed Executive Sponsors at sites where we also interviewed the Energy Champion to triangulate responses and gather multiple stakeholder feedback at a sample of sites. Responses from Executive Sponsors informed the TDA analysis results, including the quality of program delivery and the level of program engagement from site staff. The team also used the Executive Sponsor feedback to corroborate the influence of the SEM program on capital project decisions.

Energy Coaches – implementation staff in charge of coaching participating facilities through the SEM program over multiple years of participation. Each implementer has staff trained to coach participating facilities through the SEM program activities and requirements. We conducted 45-minute-long in-depth interviews with these SEM Energy Coaches to understand how the program is implemented and how well sites engaged with the program. This feedback informed the TDA results including the quality of program delivery and level of program engagement from the implementer's perspective.

Energy Team Members – includes other facility staff engaged with the day-to-day SEM activities at a site. Facilities participating in the SEM program appointed staff members to make up an Energy Team to support the Energy Champion in program activities. We planned to use short, 10-minute online surveys but had very little success getting completes. Contact

---

[36] *Early Retirement Using Preponderance of Evidence, Version 1.0,* July 16, 2014. https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/p/5325-projectbasis-eulrul-evidencev1july172014.pdf

information in provided project files was sparse for these stakeholders, and turnover at the sites made much of the sparse information obsolete. Thus, only one survey was completed.

We note this data collection will also support the SEM Activity Report targeting a July 2025 publication.

3.5.4 Final program NTGR

The final site NTGR combined the composite capital NTGR from the site-specific SRA responses and the TDA PA-specific program NTGR, proportional to the savings contribution of capital and non-capital measures at that site.

4 RESULTS

This section presents the results of the gross and net savings by key reporting dimensions. We have included reasons for any deviations between forecasted and evaluated gross savings. This section also discusses net savings results and ratios which also addresses capital measures and their contribution to overall program savings.

4.1 Gross savings and realization rates

The following subsections present the results of our gross savings analysis, starting with electric energy and demand savings. Both first year and lifecycle savings are provided for each PA and at the statewide level. Savings have been expanded from the site level to population level based on the sample design, as described in Section 3.1.1.

4.1.1 Electric savings

Table 4-1 presents the first year and lifecycle forecasted savings, evaluated savings, GRR, and relative precision at the PA and statewide levels for electric savings. Relative precisions are provided at the 90% confidence interval.

Table 4-1. PY2023 Gross electric energy and demand savings by PA

PA	First year				Lifecycle			
	Forecasted savings [a]	Evaluated savings	GRR	Relative precision	Forecasted savings [a]	Evaluated savings	GRR	Relative precision
Energy (MWh)								
MCE	471	473	100.3%	0.0%	2,357	2,622	111.2%	0.0%
SCE	29,418	17,963	61.1%	47.8%	147,091	95,076	64.6%	43.6%
Statewide	**29,890**	**18,436**	**61.7%**	**46.6%**	**149,448**	**97,698**	**65.4%**	**42.4%**
Demand (MW)								
MCE	0.1	0.1	100.3%	0.0%	0.4	0.4	111.2%	0.0%
SCE	3.6	2.0	55.2%	56.8%	18.0	10.6	58.6%	51.1%
Statewide	**3.7**	**2.1**	**56.1%**	**54.7%**	**18.4**	**11.0**	**59.7%**	**49.1%**

[a] Forecasted savings represent engineering estimates and do not include the realization rate that has been applied in savings presented in ED Tracking data

4.1.2 Gas savings

Table 4-2 presents the first year and lifecycle forecasted savings, evaluated savings, GRR, and relative precision at the PA and statewide levels for gas savings. Relative precisions are provided at the 90% confidence interval.

Table 4-2. PY2023 Gross gas energy savings by PA

PA	First year				Lifecycle			
	Forecasted savings [a]	Evaluated savings	GRR	Relative precision	Forecasted savings [a]	Evaluated savings	GRR	Relative precision
Energy (Therms/1,000)								
MCE	164	74	45.1%	9.5%	821	383	46.6%	42.5%
SCG	821	1,051	127.9%	46.3%	4,106	5,422	132.0%	45.4%
Statewide	**986**	**1,125**	**114.1%**	**42.5%**	**4,928**	**5,804**	**117.8%**	**42.5%**

[a] Forecasted savings represent engineering estimates and do not include the realization rate that has been applied in savings presented in ED Tracking data

4.1.3 Total MMBtu savings

Table 4-3 presents first year and lifecycle forecasted savings, evaluated savings, GRR, and relative precision at the PA and statewide levels for total MMBtu savings. Relative precisions are provided at the 90% confidence interval.

Table 4-3. PY2023 Gross total energy MMBtu savings by PA

PA	First year				Lifecycle			
	Forecasted savings [a]	Evaluated savings	GRR	Relative precision	Forecasted savings [a]	Evaluated savings	GRR	Relative precision
Energy (MMBtu)								
MCE	18,033	8,457	46.9%	11.4%	90,167	43,961	48.8%	41.7%
SCE	100,379	61,293	61.1%	47.8%	501,894	324,412	64.6%	43.6%
SCG	82,126	105,068	127.9%	46.3%	410,629	542,167	132.0%	45.4%
Statewide	**200,538**	**174,818**	**87.2%**	**32.5%**	**1,002,690**	**910,540**	**90.8%**	**31.2%**

4.1.4 Discrepancy analysis

This section presents an analysis of what caused forecasted savings to differ from the evaluated savings estimates for the sampled projects. This analysis is based on the discrepancies associated with first-year gross savings and is calculated on a MMBtu basis. Table 4-4 summarizes the discrepancy categories that led the evaluated savings to differ from forecasted savings.

Table 4-4. Categories of savings discrepancies

Category	Description
SEM-specific discrepancies	Differences attributed to annualization errors, non-routine adjustments, long-term and short-term operational changes, incremental savings adjustment.
Tracking data	Differences attributed to inconsistencies between savings claimed and savings calculated in the provided models and/or completion report. This also includes discrepancies in savings due to unexplained or non-documented changes.
Inoperable measure	Differences attributed to measures that were removed and were no longer in operation. This includes measures that were not installed, failure of installed equipment, and business closure.
Inappropriate baseline	Represents a difference in evaluated and reported baseline, including any baseline periods adjustment in the models used to estimate forecasted and evaluated savings. This also includes any savings deviation due to a different ISP, code, or pre-existing baseline.
Operating conditions	Collected trend data or photographs of setpoints informs different operating parameters, including hours of use, setpoints, efficiency, etc.
Calculation methods	Differences attributed to changes in calculation methodology between that used for forecasting savings and evaluation analysis. The evaluator only changed analysis methodology when necessary to accurately calculate savings such as employing an 8760 model.

As the DNV team calculated site-specific results for each sampled site, we noted the reasons for any deviation of the evaluated savings from forecasted for each site into those listed categories. We then calculated the contribution of each category of discrepancy to the overall difference between forecasted and evaluated savings.

.

Table 4-5 shows the number of instances a given discrepancy occurred, and its impact on overall gross MMBtu realization rates.

Table 4-5. Key drivers behind overall GRR (MMBtu)

Discrepancy Sub-category	Counts	Impact on RR		Impact (%)
Operating Conditions: Ex-Post M&V period different	1	0.00%	17.82%	17.8%
Operating Conditions: Same M&V methods, production changed	1	-10.22%	0.00%	-10.2%
Calculation Methods: Incorrect regression model	1	-7.80%	0.00%	-7.8%
Calculation Methods: Model Adjustments	3	-7.03%	0.00%	-7.0%
SEM-Specific Errors: Annualization Errors	6	-6.02%	0.45%	-5.6%
Operating Conditions: Change in operating hours	1	-0.02%	0.00%	0.0%
Calculation Methods: Errors found in IOU calculation method	2	0.00%	0.00%	0.0%
Total	**15**	**-31.09%**	**18.27%**	**-12.83%**

4.1.5 Comparison to previous evaluation findings

This report marks the first phase of a comprehensive two-year evaluation covering both PY 2023 and 2024. The DNV team will provide a comparative analysis of the combined results of PY2023 and PY2024 against the findings from previous evaluations findings as part of the PY2024 study report.

4.2 Net savings results and ratios

The following subsections present the results of our net savings analysis, starting with electric energy and demand savings followed by natural gas. Both first year and lifecycle savings are provided for each PA and at the statewide level. Savings have been expanded from the site level to population level using the site weights provided in the sample design, as described in section 3.1.

We note that the relative precisions of the SCE and SCG NTGR results presented in the next sections are very good, ±1-2%. This was accomplished with a small number of capital NTG site results (n = 10), because in the expansion of the results, sites with no capital measures were factored into the analysis with a NTGR of 1 with no additional variation. The relative precision of the MCE gas results is poorer due to the variable results of the capital SRA results.

4.2.1 Electric net savings

The following sections present the PY2023 results of the net savings analysis, starting with electric energy and demand savings. Both first year and lifecycle savings are provided for each PA and at the statewide level. The PY2023 electric NTGR value is similar to the PY2021/2022 results, however, these are based on SCE only, since the evaluated MCE sites did not claim electric savings. Both the TDA and capital attribution research support a NTGR close to 1.0.

Table 4-6. PY2023 electric net savings and NTGR

PA	First year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Energy (MWh)								
MCE[1]	473	469	99.2%	0.8%	2,622	2,592	98.8%	1.2%
SCE	17,963	17,820	99.2%	0.8%	95,076	93,966	98.8%	1.2%
Statewide	**18,436**	**18,289**	**99.2%**	**0.8%**	**97,698**	**96,558**	**98.8%**	**1.2%**

Demand (MW)								
MCE[1]	0.1	0.1	99.6%	0.6%	0.4	0.4	99.1%	1.3%
SCE	2.0	2.0	99.6%	0.6%	10.6	10.5	99.1%	1.3%
Statewide	**2.1**	**2.1**	**99.6%**	**0.6%**	**11.0**	**10.9**	**99.1%**	**1.3%**

1 No MCE sites with electric savings were recruited; therefore, statewide results were applied.

4.2.2 Gas net savings

Table 4-7 presents the first year and lifecycle forecasted savings, evaluated gross and net savings, NTGR, and relative precision at the PA and statewide levels for gas savings. Both the TDA and SRA research supports the current NTGR of 1.0. However, both the TDA and SRA outcomes were poor for MCE which is apparent in the natural gas NTGR and poor relative precision.

Table 4-7. PY2023 gas net savings and NTGR

PA	First year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Gas (therms)								
MCE	74	51	68.9%	18.5%	383	204	53.2%	48.5%
SCG	1,051	1,042	99.2%	0.9%	5,422	5,354	98.8%	1.5%
Statewide	**1,125**	**1,093**	**96.9%**	**1.3%**	**5,804**	**5,558**	**95.2%**	**2.5%**

4.2.3 Total MMBtu savings

Table 4-8 presents first year and lifecycle forecasted savings, evaluated savings, GRR, and relative precision at the PA and statewide levels for total MMBtu savings. Relative precisions are provided at the 90% confidence interval.

Table 4-8. PY2023 total MMBtu net savings and NTGR

PA	First year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Energy (MMBTU)								
MCE	8,457	5,828	68.9%	18.5%	43,961	23,396	53.2%	48.5%
SCE	61,293	60,806	99.2%	0.8%	324,412	320,625	98.8%	1.2%
SCG	105,068	104,242	99.2%	0.9%	542,167	535,425	98.8%	1.5%
Statewide	**174,818**	**170,876**	**97.6%**	**0.9%**	**910,540**	**879,446**	**96.3%**	**1.7%**

4.2.4 Other NTGR results

This section presents additional results and findings relevant to the final NTGR results.

4.2.4.1 Comparison with other NTGR research results

Table 4-9 compares the NTGR of recent impact evaluations for SEM, CIAC, and NMEC. The table shows that the SEM NTGR results have been consistent across the evaluations for both overall NTGR results and for capital only. The capital NTGR itself, where capital measures account for about 19% of total program MMBtus, is similar to the PY2021/22 evaluation. The SEM NTGR values are significantly better than CIAC and NMEC, partly explained by the predominance of BRO measures in the SEM measure portfolio, which are encouraged through a demanding customer engagement.

Table 4-9. Comparison of the PY2023 NTGR with other NTGR results

Program and year	Electric NTGR	RP	NG NTGR	RP
SEM 2023 preliminary interim results	99.1%	0.9%	96.9%	1.3%
Capital only, information only	0.95		0.75	Not calculated
Recommended value	1.0		1.0	
SEM 2021/2022	98.4%	0.6%	98.0%	2.5%
Capital only, information only	0.80		0.74	Not calculated
Recommended value	1.0		1.0	
SEM 2018/2019	1.0		1.0	
Core NTGR, info only	0.61	Not calculated	0.61	Not calculated
CIAC 2023 preliminary	55.2%	4.8%	39.5%	±4.7%
CIAC 2022	61%	±7%	76%	±0.1%
CIAC 2020/2021	42.7%	±10.4%	15.1%	±5.2%
NMEC 2023 preliminary	75.7%	±5.0%	76.0%	±7.0%
NMEC 2020/2022	45.9%	±11.0%	46.5%	±86.0%

4.2.4.2 NTGR of capital

As noted previously, SEM measures are tracked by the customer in the Opportunity Register. Each measure record includes fields for a measure description, the origin of the measure (for example, through the Treasure Hunt), estimates of savings and costs, and relevant to this discussion, the measure type, as well as other characteristics. The measure type is a label, which identifies whether a measure is a behavioral, retro-commissioning, maintenance, or capital type measure facilitating a direct mapping to a capital or non-capital classification.

We collected all the Opportunity Registers from the selected sites and combined them into one dataset. For the cases where the savings estimate field was empty, the engineers gave a rough estimate as described in Section 3.3. Table 4-10 shows the percentage of capital measures of program forecasted savings by PA. Capital measures, as reported in the Opportunity Registers, make up about 19% of PY2023 forecasted savings, with some differences by PA and fuel. Capital measures constituted 16% of the PY2021/2022 program MMBtu forecasted savings. The last column presents the average SRA-method NTGR of capital measures.

Table 4-10. PY2023 contribution of capital measure to SEM program forecasted savings

PA	Total MMBtu savings	Percent of capital MMBtu	Total kWh savings	Percent of capital kWh	Total therm savings	Percent of capital therm	Ave SRA capital NTGR
SCE	32,706	18%	1,682,073	18%		0%	0.83
SCG	60,336	13%	-	0%	75,693	13%	0.89
MCE	6,820	77%	27,125	19%	51,297	81%	0.44
Statewide	**99,862**	**19%**	**1,709,198**	**18%**	**126,990**	**19%**	

We found that the measure type field was well populated and accurately identified capital measures about 90% of the time based on follow-up interviews with site staff. Other associated fields intended to capture measure savings and costs were often not populated.

SEM capital measures constitute a solid but small portion of program savings. Thus, when blending the TDA NTGR of 1 for non-capital measures, the composite site and program NTGR are close to 1.

4.2.4.3 TDA results

The theory-driven attribution (TDA) research collected feedback from multiple stakeholders for each sampled site, including Energy Champions, Executive Sponsors, Energy Team members, and Energy Coaches from the implementer. The evaluation team aggregated the feedback into seven Design Guide Requirement (DGRs) categories and scored each category using a 1-3 scoring system (1 = very weak; 2 = moderate; 3 = very strong). The score indicates the strength of the evidence that the SEM program 1) was delivered in accordance with Design Guide principles, and 2) customer engagement in the program was high. Table 4-11 provides the unweighted average Delivery and Engagement scores for each of the DGRs and summarizes the average Delivery and Engagement score by PA for each DGR.

Table 4-11. Average delivery and engagement scores by PA

Requirements	Component	Average SCE Score n=8	Average SCG Score n=4	Average MCE Score n=2
1 – Kick-off meeting. Introducing SEM to the site, defining the **r**oles at the facility, and laying out a program schedule.	Delivery	3.0	3.0	2.5
	Engagement	2.3	2.5	2.0
2 – Energy Management Assessment (EMA). Assessing the site's current energy use policies and procedures. The implementer conducts these assessments each year of the program and has the option of sharing the results with the facility.	Delivery	3.0	3.0	3.0
	Engagement	3.0	3.0	3.0
3 – Energy Use Mapping. Detailed mapping of all energy consumption at the facility.	Delivery	2.8	2.5	1.5
	Engagement	2.8	2.8	3.0
4 – Energy Treasure Hunt. Walking the entire facility searching for ways to reduce energy waste either through capital project upgrades or process and behavioral changes.	Delivery	2.6	2.8	2.1
	Engagement	2.5	2.8	2.5
5 – Action Plan and Support. Identifying and creating action plans to meet SEM requirements including meetings, trainings, and workshops.	Delivery	2.5	2.4	1.6
	Engagement	2.8	2.8	1.5
6 – Energy Management Information System (EMIS). Developing a system to manage energy use and model energy savings across the entire facility.	Delivery	2.8	3.0	1.5
	Engagement	2.6	2.8	2.3
7 – Future Planning. Creating a plan for meeting program requirements in future program years as well as how to continue to pursue energy efficiency after program participation ends	Delivery	2.6	1.8	1.0
	Engagement	2.1	1.5	1.0
Average Delivery		2.8	2.6	1.9
Average Engagement		2.6	2.6	2.2

The team then averaged the Delivery and Engagement scores to arrive at a Compliance score for each DGR category. The team then divided the DGR sum for each site by the total available points (7 categories at 3 points each = 21 total points) to arrive at the final TDA result.

Table 4-12 presents the unweighted TDA results by site with a simple average by PA. The indicators are consistent with the final program NTGR results. Two of the PA programs (both delivered by Cascade Energy) have high TDA indicators and high NTGR results, shown in Table 4-6 and Table 4-7. The MCE program, delivered by CLEAResult, received lower TDA indicator values for the two gas-only sites, and a lower gas NTGR. The table also includes the capital and BRO SRA NTGs. The BRO SRA NTG is an indicator of NTG for comparison purposes.

Table 4-12. PY2023 TDA and SRA NTGR results by site

Site ID	Vendor	Delivery score	Engagement score	Combined TDA score	Final TDA NTGR score	Capital project SRA NTG	BRO project SRA NTG
SCE_01	Cascade	0.98	0.95	0.97	1	0.64	0.9
SCE_02	Cascade	0.94	0.98	0.96	1	0.93	0.97
SCE_03	Cascade	0.99	0.83	0.91	1	No Cap	0.73

Site ID	Vendor	Delivery score	Engagement score	Combined TDA score	Final TDA NTGR score	Capital project SRA NTG	BRO project SRA NTG
SCE_04	Cascade	0.94	0.88	0.91	1	1.00	0.97
SCE_05	Cascade	0.92	0.74	0.83	1	No Cap	0.53
SCE_06	Cascade	0.88	0.75	0.82	1	0.75	1.0
SCE_07	Cascade	0.79	0.76	0.78	1	No Cap	
SCE_SCG_01[1]	Cascade	0.90	0.93	0.91	1	No Cap	1.0
SCE Summary		**0.92**	**0.85**	**0.89**	**1.00**	**0.83**	
SCG_01	Cascade	0.84	0.88	0.86	1	0.90	1.0
SCG_02	Cascade	0.91	0.79	0.85	1	1.00	0.61
SCG_03	Cascade	0.86	0.83	0.84	1	0.77	
SCE_SCG_01[1]	Cascade	0.90	0.93	0.91	1	No Cap	0.83
SCG Summary		**0.88**	**0.86**	**0.87**	**1**	**0.89**	
MCE_01[2]	CLEAResult	0.65	0.74	0.70	1	0.17	No BRO
MCE_02	CLEAResult	0.60	0.71	0.66	1	0.72	0.7
MCE Summary		**0.63**	**0.73**	**0.68**	**1.00**	**0.45**	0.7

1 No gross evaluated savings due to missing data.
2 Commercial site

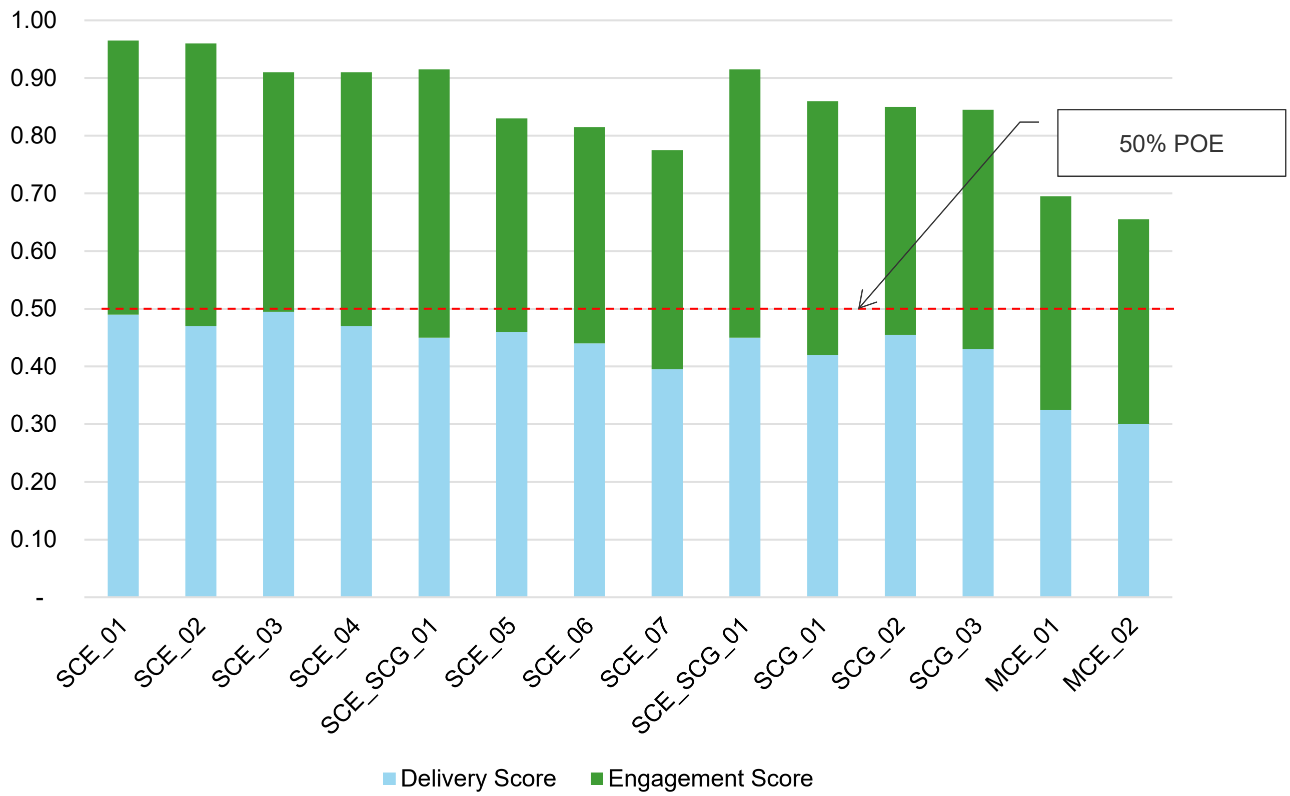
Conversations with SEM participants who reported having a successful program experience pointed to the Energy Treasure Hunt as the most beneficial and influential activity of the program. Energy Coaches also reported the treasure hunt to be the most important activity in the SEM program, allowing the implementer the chance to meet facility staff, create good working relationships, and learn about the site's energy needs. The Energy Champions reporting less than successful experiences with the SEM program voiced disappointment with the Treasure Hunt, claiming the implementer did not have the necessary expertise to evaluate and diagnose their specific industrial equipment, and therefore could not properly identify opportunities to reduce energy waste. This and other findings will be expanded in the SEM Activity Report, targeting a July 2025 publication.

The TDA effort also identified a gap in the program delivery regarding planning, and more specifically the documentation of detailed action plans and a plan to continue pursuing energy efficiency after the program ends. For example, there were very few action plans put in place to backfill SEM roles in case a person left the company or changed roles within the company. This happened multiple times at the facilities in the 2023-24 sample and none of the sites had staff able to backfill the role and continue to champion the SEM program.

4.2.4.4 Program TDA results

Figure 4-1 presents these TDA scores for each site for each PA (each bar represents a different site). The POE threshold of 0.5 overlays the figure and shows that all sites exceed that threshold. Applying the rational of the PY2021/2022 impact evaluation, program influence is more likely than not to be associated with the program implementation, thus the SEM Program NTGR of 1.0 remains.

Figure 4-1. TDA scores by PA



In contrast to the generally laudatory comments from SCE/SCG participants, the MCE participants noted:

- Expectations with the gas SEM program were unclear to one MCE participant, who reported disappointment with the types of opportunities identified in the Treasure Hunt. The Energy Champion expected a more detailed review of their gas-using equipment but said the implementation staff did not have the necessary expertise to provide valuable insight.
- The other MCE site was a school and there were very few opportunities for gas savings in the classrooms outside of the HVAC system. The maintenance staff was already aware that the HVAC units were over 30 years old and had plans to replace them even without the assistance of the SEM program.
- Both MCE participants were also disappointed with the granularity of the gas data that was only available at monthly intervals and did not allow for a detailed model of energy use.

4.2.4.5 Future SEM NTGR research

There are two considerations for future research.

Non-capital SRA research. As part of producing a well-rounded review of the SEM program, we estimated an informational non-capital NTGR using a simplified self-reported attribution approach (non-capital SRA). This simplified SRA method gathered responses from Energy Champions for the two largest savings non-capital measures to estimate a non-capital NTGR. However, the team used the capital project SRA algorithm, with slight modifications to the language of the capital question battery, adjusting it to refer to non-capital activities instead. The battery was not well-tailored to the decision-making associated with behavioral, retro-commissioning, or operational (BRO) projects. A BRO battery designed to address behavior changes and retro-commissioning-BRO type measures would more accurately account for the decisions a site considers when allocating resources for BRO upgrades.

As noted earlier, the TDA score is not equivalent to a NTGR score but instead measures customer engagement and delivery compliance with the SEM Guides. If TDA research shows that a program is not compliant, there isn't a fallback NTGR. A BRO battery could be developed and approved by stakeholders and deployed as a relatively low-cost risk mitigation strategy.

TDA threshold. The original theory-driven attribution results of the 2018/2019 Study used a binary 0.5 threshold (Yes, there is evidence, or no, there is not sufficient evidence) to show the preponderance of evidence indicates that the program was more likely than not to have influenced the outcome. However, the Expansion Study concluded the high NTGR was reasonable for programs that comply to the SEM Guidelines. This PY2023 Study TDA research scored the SEM programs based on their compliance with SEM Guidelines, promoting a discussion about whether a score higher than 0.5 is necessary for the program to be considered compliant. For example, two of the sites in the sample scored between 0.6 and 0.7 and also had the lowest gross and SRA results.

The evaluation team recommends revisiting the 0.5 threshold in the PY2024 evaluation for potential application in subsequent years. While the preponderance-of-evidence framework established in the 2018/2019 Study supports the application of a 1.0 to MCE's program. MCE's compliance, as measured by the TDA data, is on the edge of a "failing grade" at 70%.

5 CONCLUSIONS AND RECOMMENDATIONS

This section summarizes all findings from the SEM impact evaluation study and highlights the implications from the findings and recommendations from the DNV team.

5.1 Analysis methodology

The DNV team observed a further decline in the use of the top-down modeling approach by SEM participants, compared to PY2021/2022. Fewer sites used the top-down modeling approach, and the percentage of savings from top-down models for electric and gas has decreased relative to overall savings. The top-down modeling approach is better suited to capturing the full impact of SEM BRO measures, which is why the SEM M&V guide designates it as the preferred methodology for calculating SEM savings. The guide also mandates that participants justify the use of bottom-up calculations when they opt for this approach, a requirement followed by all participants who used bottom-up calculations. The SEM M&V guide also outlines examples where bottom-up calculations may be used, including but not limited to situations where energy consumption data or relevant variables are unavailable, when the number of energy meters exceeds 10 at the facility, when on-site generation metering is absent, or when production exhibits high variability. Finally, D16-08-019 in part justifies the NTGR treatment of SEM because SEM "uses NMEC and a dynamic baseline model to determine savings from all program activities at the facility...." As the program increasingly moves away from top-down estimates in savings, this treatment may be called into question.

The DNV team observed instances where SEM participants used both top-down and bottom-up calculations to estimate SEM savings for the same reporting period as an additional layer of verification. In cases where top-down savings appeared disproportionately high relative to the types and sizes of measures installed, participants opted to rely on the bottom-up method to claim savings. The DNV team considers this approach to be a best practice, ensuring accuracy and consistency in savings claims. The DNV team also observed an increase in the percentage of SEM participants who included actionable steps in their "Bottom-up justification" memos to address issues preventing the use of top-down modeling, compared to PY2021/2022. This approach is considered best practice. The DNV team classified the savings calculation methodologies used by program participants to calculate the SEM savings (for both electric and natural gas) over the two years of participation into three categories as summarized in Table 5-1.

Table 5-1. Breakdown of savings calculation methodologies

Parameter	Top-down	Bottom-up	Mixed analysis (top-down & bottom-up)
Reviewed projects, n=14	4	6	4
Percentage by count	29%	43%	29%
Percentage of electric savings	1%	40%	59%
Percentage of gas savings	25%	46%	29%
Percentage of overall MMBtu savings	13%	43%	44%

We reviewed the justifications provided by each site for using bottom-up calculations, which are summarized in Table 5-2.

Table 5-2. Bottom-up/mixed analysis calculation rationales

Justification for bottom-up/mixed analysis	Attempted model	Provided model	Quantity[1]
No correlation between energy consumption and relevant variables was identified	No	N/A	1
Challenges in developing production data, significant discrepancies in operational loads between baseline and reporting periods, and unmetered on-site energy generation	No	N/A	1

Justification for bottom-up/mixed analysis	Attempted model	Provided model	Quantity[1]
Model did not meet statistical significance requirements	Yes	No	1
Non-routine event and insufficient data provided by participant to implementers	No	No	1
Unreliable model data for relevant variables and unexplained fluctuations in consumption	Yes	Yes[2]	1
Missing and unreliable data for on-site energy generation, along with a lack of calibration of energy generation meters	No	N/A	1
Inaccurate energy data for two months of the performance period	No	N/A	1
Simplicity and consistency with other sites in the cohort	No	N/A	1
Expansion in facility operations and addition of production lines phased in over several months within reporting period	No	N/A	1
Lingering COVID-19 impacts and rebaselined model does not meet statistical requirements.	Yes	No	1
lack of granular natural gas data resulted in the model's inability to statistically represent energy usage accurately	Yes	No	1

[1] One bottom-up site provided two different justifications for using bottom-up calculations in the two reporting periods.
[2] The attempted model was not provided to the DNV team, but the data for the reporting period was supplied upon request.

As presented in Table 5-2, none of the projects that used bottom-up calculations provided the model to the DNV team for review.

Implications

- In PY2021/2022, a recommendation was made to prioritize top-down modeling; however, the DNV team observed a further decline in its use by SEM participants in the following period. This decline resulted in reduced site counts and lower savings across electric, gas, and overall consumption compared to PY2021/2022.
- Using bottom-up calculations to determine SEM savings may mask zero or negative savings, particularly for sites with claims beyond the first reporting cycle, which could experience SEM-related savings backsliding.
- Following the PY2021/2022 SEM evaluation report, it was noted that implementers typically review specific failed models and underlying reasons, while PAs review and approve the conclusions, except for some sites that may undergo further technical reviews. However, since attempted models in both PY2021/2022 and PY2023 were mostly not provided in the project packages made available to the DNV team, it remains unclear whether unreliable data or attempted models are typically provided and reviewed in-depth by the PAs.
- Unaccounted or unreliable data for on-site energy generation remains a significant factor driving the use of bottom-up calculations.
- Some reasons provided for switching from the top-down modeling approach to the bottom-up approach were found to be non-site-specific and therefore not justifiable.
- When the savings resulting from either the top-down modeling or bottom-up calculations are disproportionate to the measures installed, it should be flagged for further investigation, and consideration should be given to using the alternative savings calculation approach. Some SEM participants in PY2023 have already adopted this best-practice.
- Identifying actionable items, including necessary resources and timelines, to address issues preventing the use of top-down modeling helps expedite resolution and the return to top-down modeling. Some SEM participants in PY2023 have already adopted this best practice.

Recommendations

- Implementers should provide justifications for using bottom-up calculations that are site-specific, reasonable (in alignment with the SEM M&V guide examples), including any failed top-down models or unreliable data. PAs must conduct thorough reviews of root causes noted in bottom-up justifications before approving. This will minimize the unnecessary use of bottom-up calculations and ensure compliance with the SEM M&V guide.
- For sites using bottom-up calculations to claim savings beyond the first reporting cycle, evidence and documentation of savings persistence from the previous cycle should be collected and provided in the project files package. Following this recommendation will allow validation of the savings persistence of SEM measures regardless of the analysis approach used.
- Implementers should continue to include identified actionable items, along with necessary resources and timelines, in bottom-up justifications to address issues preventing the use of top-down modeling, thereby expediting resolution and facilitating a return to top-down modeling.
- Implementers must document in detail any failure by SEM participants to provide site-specific variable data (e.g., production or occupancy data), specifying whether the issue is due to staff turnover, non-compliance with SEM program requirements, or other reasons, and outline steps or plans to address the issue before upcoming cycles.
- Implementers should prioritize identifying and addressing issues that impede the creation of a valid top-down model as early as possible during SEM participation. For example, for sites with on-site energy generation, ensure that SEM participants are informed that meters for on-site generation must be calibrated, and the data provided by these meters must be tested and proven valid during the baseline period.
- When using a bottom-up approach, SEM participants should take the following actions:
  - Continue providing thorough documentation to justify calculating the SEM savings using bottom-up calculations.
  - Use on-site metering and trend data to determine the most accurate values for parameters used in measure-level calculations. Using as-built values leads to accurate savings estimation.
  - Provide thorough documentation of all inputs and parameters used in bottom-up calculations.
  - Expect and prepare to fulfil data requests made by the evaluators to validate measure-specific parameters.

5.2 Savings calculation considerations

This subsection summarizes the DNV team’s findings regarding the top-down models and bottom-up calculations used by SEM participants.

5.2.1 Savings annualization

Savings annualization refers to prorating the savings calculated within a short period of time to represent a full year of savings. Annualization is often used when SEM projects were installed late in the year and consequently, the full annual impact of those savings would not appear in the billing analysis. In some cases, annualization may be used when certain periods are considered unrepresentative of typical facility operations and are therefore excluded from the modeling analysis consideration. The typical observed annualization period ranges from three to five months within the final five months of the reporting period in consideration. This approach was required by older versions of the SEM M&V guide.[37] The current version of the guide limited the use of annualization to only when the model is being retired or a customer will not be participating in the SEM program after the current Reporting Period, with PA authorization.[38] The DNV team acknowledges

---

37 Sergio Dias Consulting. “California Industrial SEM M&V Guide, Version 2.01.” Section 11.5.1. September 12, 2020. https://pda.energydataweb.com/api/downloads/2525/CA_Industrial_SEM_M%26V%20Guide_v2.01.pdf

38 Sergio Dias Consulting. “California Industrial SEM M&V Guide, Version 3.02.” Section 1.4. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

that program participants who used annualization followed the SEM M&V guidelines. However, the annualization approach often overlooks the seasonality in the typical annual operation for facilities which could result in inaccurate savings estimation.

Overall, the use of annualization based on insufficient or unrepresentative periods of facility operation in PY2023 resulted in a 6% decrease in total forecasted savings, compared to an 8% decrease in PY2021/2022, reflecting significant improvement from PY2021/2022. This improvement is primarily due to sites using longer periods for annualization in PY2023. Further details on the impacts of the savings annualization are summarized in Table 5-3.

Table 5-3. Savings annualization impacts

Parameter	PY2021/22	PY2023
Sites with savings annualization discrepancy	19	5
Total number of sites	47	14
Percentage by count	40%	36%
Contribution to overall deviation from forecasted savings	57%	36%

Overall, the DNV team found that savings annualization led to overestimation of SEM savings in approximately 60% of sites using this approach compared to 70% in PY2021/2022.

Implications

- SEM participants demonstrated significant improvement in utilizing longer annualization periods for sites with seasonal operation in PY2023 compared to PY2021/2022.
- Savings annualization carries a significant risk of savings miscalculation, as operations and production during the annualization period may not represent typical yearly conditions. This risk is increased when savings are annualized in the first year of the cycle, and bottom-up calculations are used for the second year.
- Savings annualization is inconsistent with SEM's performance-based approach to estimating savings through billing analysis and creates challenges in reconciling savings in subsequent years.

Recommendations

- Implementers should follow the SEM M&V guide, which recommends limiting annualization to only when the model is being retired or a customer will not be participating in the SEM program after the current reporting period, with PA authorization. Therefore, annualized savings should be rejected when annualization is being used outside of these two reasons as they are likely to produce inaccurate annual savings, such as seasonally impacted savings, or measures with that fluctuate over time, such as shutdown-type measures. Following this recommendation will result in more accurate savings estimations.
- Implementers should continue to investigate further when savings from either top-down modeling or bottom-up calculations are inaccurate or disproportionate to the measures installed and consider using the alternative savings calculation approach (either top-down or bottom-up) if the disparity cannot be explained or resolved. Following this recommendation will ensure that the claimed savings accurately reflect the performance of installed measures, leading to a more precise representation of the SEM program's impact.
- If SEM participants decide to use bottom-up calculations for the second reporting period after using the top-down modeling approach in the first, implementers should consider recalculating the savings from the first reporting period using the bottom-up calculations and truing up the savings at the end of the cycle. Following this recommendation will minimize the use of annualization based on a limited time period, reducing the risk of inaccurate savings estimate.

5.2.2 Modeling adjustments

The DNV team performed model adjustments that accounted for 45% of the difference between forecasted and evaluated savings, up from 27% in PY2021/2022. The team reviewed all top-down models used by SEM participants to calculate savings for projects implemented in PY2023. However, several models required adjustments to enhance statistical significance, better reflect typical operations, and improve savings accuracy. These improvements were achieved through site-specific adjustments, including:

- Refined weather variables (e.g., adjusting the changepoint and removing dependent variables).
- Removed dependent variables and variables that showed no correlation with energy consumption.
- Removed and added indicators depending on their added value and impact on the model statistical significance.
- Excluded non-routine events (NREs) and periods of unexplainable energy consumption spikes and dips.
- Switched some models from daily basis to weekly to better correlate energy consumption with relevant variables.
- Removed variables that showed no correlation to energy consumptions.
- Reviewed variables used to determine between the consolidated variables that are directly connected (such as production of different units) or only including variables that improve the model statistical significance.

Recommendations

- Continue to follow the SEM M&V guidelines on creating top-down models and assessing their validity.[39] Below are some examples of the steps to take in ensuring the M&V guidelines are followed:
  - Ensure that the model is reflective of the facilities' typical operation for both baseline and reporting periods and that claimed savings correspond to.
  - Optimize the model's statistical fit by testing various changepoints for weather variables and selecting the most appropriate point for each site.
  - Continue to provide collected granular data for all tested variables to allow PAs and evaluators the opportunity to verify implementers findings and conclusions.
  - Continue to only consider relevant independent variables in the model savings calculations.
  - Continue to provide thorough documentation to justify any periods that are excluded from the model savings calculations.
- Avoid using hard-coded values in the savings calculations. The use of hard-coded values prevents the participants, PA reviewers, and evaluators from tracking the sources of the used values and complicates the process of updating and validating model results.

5.2.3 Bottom-up calculations discrepancies

The DNV team reviewed all submitted bottom-up calculations for sampled projects as described in Section 3.2.5.2. The DNV team adjusted the calculation methodology, savings calculation inputs, and parameters as needed based on an in-depth engineering review and as-built data collected from facility personnel. Overall, these adjustments to the bottom-up calculations led to a 3% increase in forecasted savings. Below is a summary of key adjustments made:

- Revised the savings calculation methodology to incorporate measure-specific parameters rather than unsubstantiated savings factors.
- Updated savings calculations by normalizing savings to baseline operations, ensuring a valid comparison between baseline and as-built performances.
- Collected operational data and trend information to update load and hour assumptions in forecasted savings calculations for targeted systems, based on as-built operation data.
- Verified installed equipment specifications using invoices, photographs, and confirmation from facility personnel.
- Verified that the measures included in forecasted savings calculations were installed and operational.

Recommendations

- Implementers should continue using measure- and site-specific parameters with documented references and substantiation for all inputs, to the extent feasible. This will result in more accurate savings estimations.
- Implementers should continue normalizing baseline production and occupancy profiles based on as-built operations to result in calculated savings that reflect only installed measures and improvements.
- Implementers should include any trend or metered data used for forecasted savings estimation in project files which will result in more accurate savings impact analysis results.
- Implementers should collect invoices, photographs, and any available documentation to substantiate assumptions and parameters used in forecasted savings estimations. Following this recommendation will result in more accurate savings estimations.

---

[39] Sergio Dias Consulting. "California Industrial SEM M&V Guide, Version 3.02." Sections 4, 6, and 7. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

5.3 Program reporting and tracking data

Since PY2020, participation in the SEM program has remained steady despite restrictions on facility types, as it has primarily been available to industrial sector facilities, with limited exceptions in certain jurisdictions. The DNV team expects participation to accelerate significantly in the coming years with the planned expansion to the commercial sector. Additional details on the history of SEM participation rates and forecasted savings per program year are provided in Table 5-4.

Table 5-4. History of SEM participation

Parameter	PY2020[40]	PY2021	PY2022	PY2023
Number of participants	35	27	26	45
Total forecasted savings (MMBtu)	421,188	118,118	335,457	200,538

The unique structure of the SEM program allows a single participant to submit multiple claims over multiple years—typically within a two-year cycle, though instances of three-year spans have been observed. Therefore, as the program expands, it is essential for the CPUC to continue collaborating with PAs to enhance program claims reporting and tracking. Starting in PY2025, the CPUC and PAs have agreed to introduce a new field, "SEM Cycle Status," into CEDARS[41] to indicate each claim's program year (1, 2, or 3) and reporting period (first or second). This new field will enhance the tracking of multiple claims and total forecasted savings per cycle for each participant. Additionally, it will enhance the monitoring of the SEM population for each program year to track program trends and support future evaluations. The CPUC and PAs have also agreed to pursue further improvements, including the potential addition of a field to track overall site status (e.g., site dropped, site paused participation).

Implications

- PAs have consistently demonstrated a commitment to improving SEM program reporting and data tracking.
- Some improvement opportunities identified in prior studies and discussions have been implemented as of PY2025.
- The absence of reporting on certain aspects of SEM claims, including the analysis methodology (top-down or bottom-up), represents an area for further enhancement. Incorporating this reporting requirement would facilitate more precise tracking of program trends and improve the rigor of future evaluations by enabling the development of more refined sampling strata.

---

40 SBW Consulting, *2018-2019 Strategic Energy Management (SEM) Program Impact Evaluation*, Final Report, January 21, 2022. https://pda.energydataweb.com/#!/documents/2582/view

41 California Energy Data and Reporting System (CEDARS), "Welcome to CEDARS," cedars.sound-data.com, https://cedars.sound-data.com/

Recommendations

- PAs should continue working in collaboration with CPUC staff to enhance SEM program reporting and data tracking. This continuous collaboration will ensure that as the program expands, the data tracking systems are developed to effectively monitor and support this growth, allowing for more accurate tracking of the program's expansion and overall impact.
- The CPUC staff should continue to review, refine, and implement improvement opportunities for program tracking identified in prior studies and statewide discussions, which will result in better informed program and policy outcomes.
- The CPUC staff and PAs should prioritize exploring the addition of an "analysis methodology" field when reporting claims in CEDARS to indicate whether each claim used top-down modeling or bottom-up calculations. Following this recommendation will improve tracking of top-down vs. bottom-up methodologies, allowing for better monitoring of program trends and enhanced stratification of sampling for evaluation.

5.4 NTGR methods and results

NTGR methods

The SEM SRA instrument developed in the PY2021/2022 evaluation was optimized through revisions to reduce the number of program and non-program influence factors. The 2018–2019 TDA instruments were adapted for the PY2023 evaluation referencing the approved SEM Design Guide rather than an evaluator-developed Program Theory and Logic Model (PTLM).

SEM, like no other current program, seeks to change organizational behavior in a manner that produces persistent and ongoing savings through a series of prescriptive activities. TDA is an essential method for assessing whether these broad objectives have been met. However, TDA does not directly produce a NTGR. In addition, the nuanced picture produced by TDA of the site delivery and engagement is boiled down to a one or a zero. While this method was used to meet a 50% threshold of whether the program influenced the participants actions in order to earn a NTG of 1, if a PA's program-level TDA score fails to meet the preponderance of evidence criteria, there is no lower fallback NTGR.

Finally, the Expansion Study[42] tightly linked the success of the program to following the SEM Guides. The Combined TDA Score is a measure of how well the program is complying with the letter and spirit of the Guide. Another perspective is that a project should meet a minimum standard of compliance to be considered a SEM delivery warranting the SEM NTGR. Standards for compliance or competency typically meet a higher bar than 50% which is an accepted threshold for preponderance of evidence but may be too low a threshold for compliance.

[42] DNV, *Group D Strategic Energy Management Expansion Study*, Final Report, May 23, 2024. https://www.calmac.org/publications/CPUC_SEM_Expansion_Study_Final_Report_240731.pdf

Recommendations

- We recommend using the SRA and TDA survey instruments developed and refined in this impact evaluation for the PY2024 evaluation cycle, which will result in better alignment of program tracking objectives with current CPUC data tracking policies.
- We recommend the development of a non-capital, BRO-oriented SRA NTGR battery to provide additional perspectives on program influence. Following this recommendation will provide an alternate source of NTGR should a PA fail to demonstrate a high level of influence on program outcomes.
- We recommend reconsidering the algorithm threshold for converting a TDA score to a NTGR score for applications after the PY2024 evaluation. Following this recommendation will result in more accurate NTGR considerations for the unique characteristics of SEM.

NTGR results

The two MCE projects (of four) missed delivery and engagement metrics and scored notably lower than the other PAs on the site-specific metrics and the gas NTGR. Notably, the worst-performing site was an MCE commercial site. In contrast to the generally laudatory comments from SCE/SCG participants, the MCE participants noted:

- Expectations with the gas SEM program were unclear to one MCE participant, who reported disappointment with the types of opportunities identified in the Treasure Hunt. The Energy Champion expected a more detailed review of their gas-using equipment but said the implementation staff did not have the necessary expertise to provide valuable insight.
- The other MCE site was a school and there were very few opportunities for gas savings in the classrooms outside of the HVAC system. The maintenance staff was already aware that the HVAC units were over 30 years old and had plans to replace them even without the assistance of the SEM program.
- Both MCE participants were also disappointed with the granularity of the gas data that was only available at monthly intervals and did not allow for a detailed model of energy use.

Recommendations

- We recommend that MCE consider these findings in conjunction with the forthcoming SEM Activity Report (targeting a July 2025 publication) to review the delivery with the current vendor to re-enforce the SEM Design guide. Following this recommendation will result in enhanced performance calculations for MCE commercial sites in terms of participant engagement.

About DNV

DNV is an independent assurance and risk management provider, operating in more than 100 countries, with the purpose of safeguarding life, property, and the environment. Whether assessing a new ship design, qualifying technology for a floating wind farm, analyzing sensor data from a gas pipeline, or certifying a food company's supply chain, DNV enables its customers and their stakeholders to manage technological and regulatory complexity with confidence. As a trusted voice for many of the world's most successful organizations, we use our broad experience and deep expertise to advance safety and sustainable performance, set industry standards, and inspire and invent solutions.