GROUP D

Evaluation, Measurement, and Verification of Program Years 2023/2024 Strategic Energy Management Projects

California Public Utilities Commission
CALMAC ID: CPU0407.01
Date: July 16, 2026

Table of contents

List of figures

List of tables

Glossary of key terms and acronyms[1]

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

Baseline period – 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 using measure-level calculations. This methodology uses measure-specific formulas, inputs, and assumptions, etc., to calculate measure-specific savings. The overall site savings are then calculated by aggregating the energy savings of all implemented measures.

BRO measures – Implemented or planned SEM measures that are behavioral, retrocommissioning, and operations.

CCT – The Custom Core Template, an Excel-based tool that the SEM evaluation team uses to report site-specific data collection efforts, review participant documentation and methods, and document their methods and findings.

California Database for Energy Efficiency Resources (DEER) – 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.[2]

California Energy Data and Reporting System (CEDARS) – The database that securely manages California Energy Efficiency Program data reported to the Commission by investor-owned utilities (IOUs), Regional Energy Networks (RENs), and certain Community Choice Aggregators (CCAs).[3]

Capital project – A capital project typically entails larger dollar expenditures for large equipment like a new process line or a new central plant boiler or chiller. The company includes the project in a capital plan approved at the corporate level in advance of the project implementation.

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

Custom project review (CPR) – 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.

Energy Champion – The person who is responsible for the success of the SEM program at the facility. This individual is responsible for coordinating both with the SEM Coach and internally with any facility staff, including the Energy Team, Data Owner, and Executive Sponsor.

[1] Please refer to the Energy Efficiency Policy Manual for additional terms and definitions: https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/e/6442465683-eepolicymanualrevised-march-20-2020-b.pdf

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

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

Energy Coach – The Energy Coach is typically employed by the program implementer and acts as the connector between the facility and the Program Administrator (PA). The Energy Coach facilitates SEM engagement, supports data planning, verifies Scoping Report details, and ensures accurate documentation of capital projects and savings opportunities.

Energy Team – Typically, a cross-functional team (i.e. management, production, procurement, maintenance, HR) that meets regularly to discuss SEM activity at their site.

Energy Performance Improvement Action (EPIA) – A documented behavioral, retrocommissioning, operational, maintenance, or capital action that improves facility energy performance. Typically identified during SEM engagement, each EPIA must include savings estimates, costs, and interactions to ensure compliance with CPUC reporting and avoid double counting.

Executive Sponsor – The Executive Sponsor should be the highest-level manager available at the facility (typically the facility manager) and is responsible for ensuring the Energy Team has the resources it needs to succeed during the SEM program.

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

Gross realization rate (GRR) – The ratio of achieved 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 come to fruition.

Gross savings – 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.

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

Lifecycle savings – 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) – 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 – 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-year cycle and a top-down modeling approach in the 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 savings by the net-to-gross ratio.

Net-to-gross ratio (NTGR) – A ratio or percentage of net program savings divided by 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. NRAs occur separately from the routine adjustments made using independent variables in the adjustment model. NRAs 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) – 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 metered energy consumption (NMEC) – Refers to population- or site-level programs (HOPPs) that provide incentives based on metered energy consumption. This initiative fulfills the directive for utilities to quickly identify high energy-efficiency savings opportunities in existing buildings and include only CPUC-approved NMEC building programs

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 (e.g., Marin Clean Energy (MCE),[4] Pacific Gas & Electric (PG&E), Southern California Edison (SCE), Southern California Gas (SCG), San Diego Gas & Electric (SDG&E)).

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 staff memo issued on March 21, 2019, 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.

Self-reported attribution (SRA) – A NTGR research method that collects evidence of free-ridership through participant surveys (thus, the data is self-reported) about their motivations for installing a program sponsored measure.

Theory driven attribution (TDA) – An evaluation strategy for identifying program influence that integrates input from multiple perspectives to determine how well the program delivers the program in relation to its design as expressed in the program theory and logic models.

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.

[4] 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.

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

Total System Benefit – Is an expression, in dollars, of the lifecycle energy, ancillary services, generation capacity, transmission and distribution capacity, and GHG benefits of energy efficiency activities, on an annual basis.

DNV

1 EXECUTIVE SUMMARY

1.1 Introduction

This report presents the key findings of the evaluation of California's Strategic Energy Management (SEM) program conducted on behalf of the California Public Utilities Commission (CPUC) for Program Years 2023 (PY2023) and 2024 (PY2024). The evaluation assessed program impacts on electricity savings (kWh), natural gas savings (therms), and Total System Benefit (TSB). The CPUC adopted a new TSB goal metric and directed that the TSB replace energy and peak demand savings as the single goal metric for evaluating energy efficiency programs. Accordingly, this evaluation reports gross[5] and net savings[6] results for electricity, natural gas, and TSB. This evaluation is also the first SEM study to develop TSB realization rates.[7]

This report marks the second year of a two-year evaluation covering both PY2023 and PY2024. The prior report (PY2023)[8] detailed results for projects that completed their SEM cycles in PY2023, while this report presents results for projects that completed their SEM cycles in PY2024. This Executive Summary summarizes combined PY2023/PY2024 results that program administrators will use to update program-level gross realization rates (GRRs) [9] and net-to-gross ratios (NTGRs).[10]

Goals and objectives

The overall goal of this study is to evaluate TSB savings along with kWh, peak kW, and therms savings for SEM projects implemented in PY2024 to make meaningful recommendations for improving program performance through the evaluation of a sample of SEM projects. Specifically, the goals of this study are to:

- Develop first-year and lifecycle evaluated net and gross savings for the SEM program in PY2024 and integrate these results with the results from the PY2023 study to achieve a high level of precision.[11]
- Calculate TSB based on evaluated energy savings and compare them to PA forecasted values to quantify the TSB impact.
- Develop actionable recommendations to improve program performance in delivering energy efficiency savings.

---

[5] Gross savings are the changes in energy and power demand that resulted from energy efficiency program activities, regardless of what factors may have motivated the program participants to take actions.

[6] Net savings estimate is developed by first estimating the amount of "free ridership," which is the savings that would have occurred without the incentive being provided (e.g., because the customer indicates they would have purchased the equipment at full cost if the incentive had not been offered). We then subtract the amount of free ridership from the gross savings.

[7] TSB realization rate compares the Total System Benefits (TSB) that evaluators calculated using evaluated inputs—including evaluated energy savings, measure lives, and net-to-gross ratios—to the TSB that program administrators forecasted. The metric measures how closely evaluated lifecycle benefits—such as monetized energy, capacity, and greenhouse gas impacts accrued over the full measure life—align with forecasted benefits. Unlike the Gross Realization Rate, which compares only energy savings, the TSB realization rate evaluates alignment across the full system-wide benefit framework.

[8] DNV, *Evaluation, Measurement, and Verification of Program Year 2023 Strategic Energy Management Projects*. June 25, 2025. PY2023 SEM Impact Evaluation

[9] Gross realization rate (GRR) – A measure that compares the energy savings evaluators calculate with the energy savings Program Administrators (PAs) forecast.

[10] Net-to-gross ratio (NTGR) – A measure of how much of the installed SEM projects and their energy savings occurred because of the program. For example, an NTGR of 1.0 means the program caused all the savings, while an NTGR of 0.40 means 40% of the savings were driven by the program and the rest likely would have occurred without it.

[11] Relative precision – A measure of how precise an estimate is, expressed as the size of the error relative to the estimate.

Background

SEM is a unique program under the Energy Efficiency Portfolio, with a dedicated statewide Design Guide[12] and M&V[13] Guide[14] that all SEM program implementers must follow. Program evaluation requires adherence to the CA evaluation framework[15] and the documented approaches approved for SEM program implementation as described in the SEM Design and M&V Guides. The SEM M&V Guide designates top-down[16] modeling as the preferred method for calculating SEM savings. The alternative bottom-up approach estimates savings for each measure[17] individually using engineering methods, then aggregates them to determine total SEM savings and should only be used when implementers, following appropriate due diligence, determine that modeling limitations preclude feasible top-down analysis.

1.2 Methodology overview

The DNV team evaluated gross and net savings that the PAs forecasted for SEM projects reported in PY2024 to assess program performance. This study adheres to the International Performance Measurement and Verification Protocol (IPMVP)[18] and the California Evaluation Protocol.[19]

Gross methods

The DNV team selected 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 documented our approach to evaluating the SEM projects at each site and our findings in the Custom Core Template (CCT) tool, a macro-based Excel workbook that documents PA savings claims along with our assumptions and methods used to estimate savings. We also used the CCT to report evaluated savings, assess the data collected from sites, review the analysis methods used and GRRs reported, and document the reasons for discrepancies between forecasted savings and evaluated savings. Data collection consisted of participant interviews to determine which measures they installed and operate, photographs to verify installed measures, consumption data to estimate savings, and, in some cases, equipment trend data or performance logs to confirm operation.

Net methods

In this evaluation cycle, we employed two evaluation methods for net savings:

---

[12] 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.

[13] Measurement and Verification (M&V) – A process of using measurements to reliably quantify the actual energy performance of a facility, system, or measure, and to determine energy savings by comparing post-implementation performance to an established baseline, while accounting for relevant independent variables.

[14] 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

[15] TecMarket Works Framework Team, *The California Evaluation Framework*. January 24, 2006. [CA Evaluation Framework.pdf]

[16] Top-down savings calculation – A method that estimates SEM energy savings by comparing electric and natural gas use from utility bills before and after the program. It uses simple energy models to account for changes in weather, production, and other site activities, so the remaining reduction reflects energy efficiency measures installed through the program

[17] Measure – A documented action or project—behavioral, operational, retrocommissioning, or capital—that improves a facility's energy use.

[18] 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.

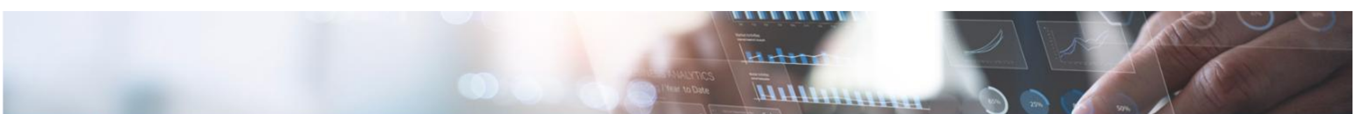
[19] The California Evaluation Protocol (CEP) is a set of guidelines and procedures developed by the CPUC for conducting evaluations of energy efficiency programs.

- **Capital measures.** We asked participants whether they would have installed a capital projects, which are high-cost projects that require company investment, without SEM intervention. Our evaluation team uses responses to develop a NTGR value between zero and 1, where zero indicates the program had no influence on the decision to install and 1 indicates the program influence was essential to installation. Under the SEM framework, capital measures, may apply a SEM NTGR of 1.0 “when program influence is evident.” The survey-based method is well suited to the assessment of capital measures that typically entail installation of equipment and a structured internal customer decision-making process.
- **Behavioral and other non-capital measures**. We used a theory-driven approach (TDA) to evaluate the overall design and delivery of the SEM program. TDA is a survey-based approach that assesses, on a scale from zero to one, how effectively program training and activities are delivered to customers, as well as the level of customer engagement, based on interviews with multiple stakeholders. This approach is well suited to assessing the organizational transformation intent of the SEM program through in-depth interviews with key stakeholders and decision-makers.

For all measures (capital and non-capital), we used survey instruments and methods approved in the PY2021/2022 evaluation.[20]

1.3 Evaluated program savings claims

This SEM impact evaluation focused on SEM participants who completed a two-year cycle in PY2023 and PY2024. We verified two-year aggregated savings for each reviewed site. The PY2024 sample frame included 103 customers (91 electric, 33 gas), and the combined PY2023/24 sample frame included an additional 42 PY2023 customers (31 electric, 12 gas).



1.4 Results

TSB results

Table 1-1 presents the TSB results by PA and statewide. We estimate $194.7 million in TSB savings compared to the forecasted estimate of $208 million. The team used revisions to gross and net savings per site to update the tracking values.

Table 1-1. PY2023/2024 First-year TSB by PA

Program Administrator	Forecasted TSB [a]	Evaluated savings	GRR	Relative precision
MCE	$1,335,155	$780,202	58%	13%
PG&E	$168,923,445	$165,697,220	98%	1%
SCE	$23,615,905	$15,838,361	67%	24%
SCG	$14,115,003	$12,341,175	87%	**24%**
Statewide	**$207,989,508**	**$194,656,960**	**94%**	**3%**

[a] Forecasted savings represent engineering estimates and do not include the realization rate that the PAs have applied in savings they presented in ED Tracking data.

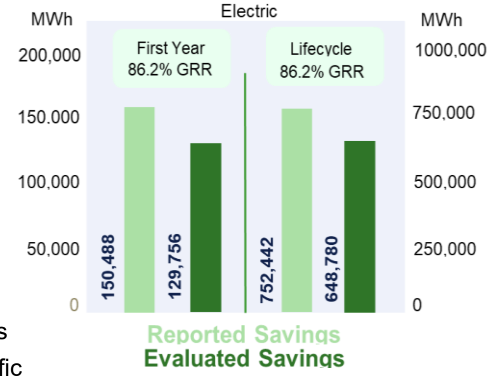
[20] DNV, *Strategic Energy Management (SEM) 2021-2022 Impact Evaluation*. May 25, 2024. PY2021/22 SEM Impact Evaluation

Gross savings results

This section presents the **overall electric and natural gas savings and GRRs for PY2023 and PY2024 combined**. We calculated all relative precisions in the tables that follow at the 90% confidence interval.[21]

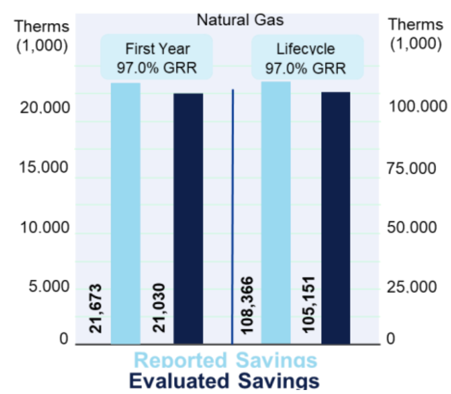
The evaluation met and exceeded the 90/10 precision target, achieving statewide relative precision of **7.0% for electric gross savings** and **2.8% for gas gross savings**, indicating a high level of confidence in the results.

Key electric gross findings



- As shown in the chart to the right, the SEM participants who completed a two-year cycle in 2023 and 2024 **achieved an aggregated first-year gross electric savings of 129,756 MWh** with a statewide[22] GRR of 86%. Updates to both bottom-up and top-down savings calculations to improve accuracy and consistency drove the electric first-year and lifecycle realization rates. These updates included excluding outlier baseline energy-consumption data from top-down models, updating weather variables using consistent and verifiable data sources, removing models that could not be validated due to invalid variables, and recalculating savings using longer and more representative production, equipment, and site-specific operating data, including confirmed operating setpoints from building management systems where available.
- PY2023/24 SEM program participants **achieved an aggregated lifecycle gross electric savings of 684,780 MWh** with a statewide GRR of 86%. The realization rates for lifecycle and first-year gross electric savings are the same for PY2023/24, as both PAs and evaluators consistently calculated lifecycle savings using a five-year effective useful life (EUL)[23] for all SEM claims which indicates that lifecycle gross realization rates reflect first-year savings (FYS) performance rather than differences in EUL assumptions.

Key gas gross findings



- As shown in the chart to the right, the SEM customers who completed a two-year cycle in 2023 and 2024 **achieved an aggregated first-year gross natural gas savings of 21,030,000 therms** with a statewide GRR of 97%. Updates to both bottom-up and top-down savings calculations to improve accuracy and consistency drove the natural gas first-year and lifecycle realization rates. These updates included replacing unsubstantiated bottom-up savings parameters with site-specific parameters (e.g., boiler efficiencies and pressure setpoints), recalculating production and equipment inputs using longer and more representative trend data, correcting proration errors, and refining top-down models by updating weather variables and excluding unsupported outlier data points.
- PY2023/24 SEM program participants **achieved an aggregated lifecycle gross natural gas savings of 5,804 thousand therms** with a statewide GRR of 97%. The realization rates for lifecycle and first-year gross gas savings are the same for PY2023/24, as both PAs and evaluators consistently calculated lifecycle savings using a five-year effective

[21] Relative precision at the 90% confidence level describes how large the uncertainty is compared to the estimate. For example, if savings equal 100 units with a ±10-unit range, the relative precision is ±10%, meaning we are 90% confident the true value lies between 90 and 110 units.

[22] This evaluation included SEM projects from all program administrators except SDG&E, as no SDG&E projects completed a full two-year SEM cycle in PY2023 or PY2024. Administrative timing shifts delayed SDG&E's reporting of SEM projects from PY2024 to PY2025.

[23] 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.

useful life for all SEM claims which indicates that lifecycle gross realization rates reflect first-year savings performance rather than differences in EUL assumptions.

Net savings results

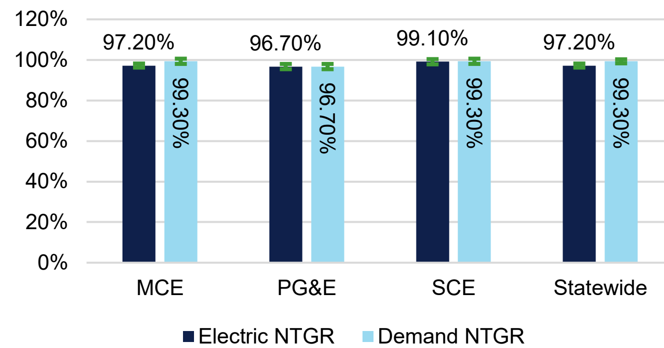
This section presents the **combined PY2023 and PY2024 electric and natural gas net savings and net-to-gross ratios (NTGRs)** for the evaluated PAs. We calculated all relative precisions in the tables that follow at the 90% confidence interval.

Key electric net findings

PY2023/24 SEM program participants achieved an aggregated first-year net electric savings of 126,087 MWh with a statewide NTGR of 97.2% with a relative precision of ± 1.0% at the 90% confidence interval. The relative precisions exceed targets for the following reasons:

- The survey-based capital measure NTGR showed relatively low variation of values ranging from 0.48 to 1.0. Capital measures contributed 16% of electric savings, a clear minority of savings.
- The TDA assessment of program delivery and customer engagement resulted in a NTGR of 1 which we applied to non-capital measures without variation. The non-capital measures contributed 84% of electric savings.
- The combination of applying a TDA NTGR of 1.0 to 84% of non-capital measures without variation and applying a survey-based NTGR calculated at the site level, which showed low variation across sites, to the remaining capital measures produced a total NTG of 97.2% with ±1.0% relative precision. The PY2023/24 statewide NTGR closely aligns with the PY2021/22 two-year result of 96.9%.

Figure 1-1. PY2023/24 FY electric statewide NTGRs



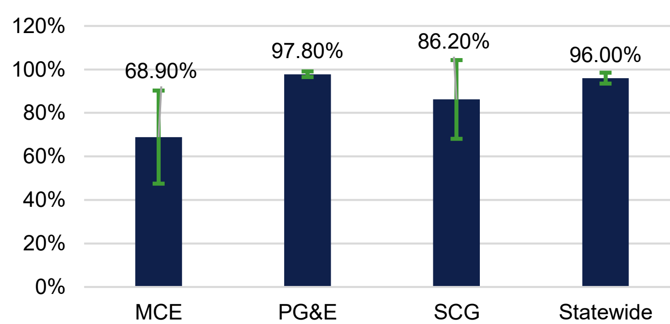
* We were unable to recruit any MCE sites with electric savings; therefore, we applied statewide results for MCE.
† SDG&E did not report any SEM savings in PY2023 or PY2024. This is due to administrative timing shifts that delayed their reporting to PY2025.
‡ Bands show absolute precision at the 90% confidence level.

Key natural gas net findings

PY2023/24 SEM program participants achieved an aggregated first-year net natural gas savings of 20.2 million therms with a statewide NTGR of 96.0% with a relative precision of ±1.0% at the 90% confidence interval. Natural gas precisions exceed targets for the same reasons as electric:

- The TDA assessment resulted in an NTGR of 1.0, which was applied to non-capital measures, accounting for 90% of natural gas savings.
- The site-level, survey-based capital NTGR covered the remaining 10% of gas savings and showed low variation across sites.

Figure 1-2. PY2023/24 FY natural gas statewide NTGRs



* SDG&E did not report any SEM savings in PY2023 or PY2024. This is due to administrative timing shifts that delayed their reporting to PY2025.
† Bands show absolute precision at the 90% confidence level.

- The combination of applying a TDA NTGR of 1.0 to the majority of savings and a site-level, survey-based NTGR to capital measures produced a PY2023/24 statewide NTGR that is within the error of the PY2021/22 two-year result of 96.5%

Both MCE and SCG had a single capital project negatively impacting their PA-specific capital project NTGR. However, because capital measures make up a small share of savings, the overall NTGR was higher. The preponderance of evidence framework established in the PY2018/19 SEM Impact Evaluation Study[24] supports the application of a 1.0 NTGR.

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

1.5 Conclusions and recommendations

Analysis methodology

The SEM M&V Guide designates top-down[25] modeling as the preferred method for calculating SEM savings because it is better suited to capturing the full impact of SEM Behavioral, Retrocommissioning, and Operational (BRO) measures.

In PY2024, approximately one-third of projects used top-down to estimate SEM savings, one-third used bottom-up, and one-third used a mixed approach, meaning participants applied both methods at different points across cycle years and fuel types. As in PY2023, limitations in data availability constrained many participants' ability to sustain valid top-down models across all fuels and cycle years, particularly the intermittent availability of production proxies and on-site generation data. Participants used top-down modeling when these drivers were available and reverted to bottom-up when they were not.

Compared to PY2021/22, PY2023/24 showed reduced use of top-down methods and increased reliance on bottom-up calculations and mixed approaches, reflecting ongoing data availability challenges in sustaining top-down modeling.

Recommendations 1–3

- **Prioritize building the data infrastructure that supports top-down modeling**: Implementers should make this a focus early in participant recruitment and onboarding by assessing whether key energy drivers (e.g., production and on-site generation) are being tracked with sufficient quality and continuity, and by working with participants to address gaps through improved logging practices, records from existing facility control systems, or considering installing new data logging systems where needed, to reduce reliance on case-by-case bottom-up exceptions.
- **Reassess top-down readiness at regular intervals**: Implementers should assess top-down readiness at enrollment and revisit it mid-year and ensure the availability of stable production proxies and valid operational data during the baseline period to avoid modeling challenges at mid-year.

---

24 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

25 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.

Recommendations 1–3

- **Document steps to restore top-down methodology**: Implementers should document steps to restore top-down modeling when they use bottom-up calculations due to data unavailability and include this information in the bottom-up justifications. This will reduce data limitations that require use of bottom-up calculation.[26] To identify progress of recommendations 1–3, future evaluations will measure percentage of top-down to bottom-up modeling used.

Savings calculation considerations

Modeling adjustments

Our top-down model corrections in PY2023/24 resulted in a 1.6% decrease in total forecasted savings. The DNV team reviewed all top-down models that SEM implementers used to calculate savings for projects implemented in PY2024. Several models required corrections to improve statistical significance, better reflect typical operations, and enhance the accuracy of savings calculations. We made these improvements through site-specific adjustments, including:

- Attempting to regenerate the implementer's models to verify their validity and then reviewing the selected variables to assess their relationship with facility energy use.
- Excluding non-routine events (NREs)[27] and periods of unusually low or high energy use when implementers did not provide supporting documentation and facility staff could not explain the cause.
- Requesting trend data for variables that better correlated with energy consumption and excluding savings when neither valid trend data nor substantiated bottom-up calculations were available.
- Applying a consistent screening framework to identify anomalous data points, including a ±3-standard-deviation statistical screen for baseline data and engineering reasonableness checks for reporting-period data. In several cases, implementers applied similar statistical screening but did not sufficiently document the rationale for retaining or excluding flagged data points, limiting transparency and requiring additional evaluator review.

Recommendations 4–6

- **Apply consistent screening and documentation of anomalous data points**: Implementers should consistently apply established screening thresholds (e.g., statistical outlier tests) to flag anomalous baseline and reporting-period data points, and document whether flagged points reflect normal operations or a non-routine event. Clear documentation of flagged points will preserve institutional knowledge and support transparent and accurate evaluation in future years. Future evaluations will track consistent flagging and documentation of anomalous data points as normal or NREs.
- **Review and validate anomalously high or low savings results**: Implementers should review the savings contribution of each reporting-period data point and investigate periods that produce unusually high or low savings, confirm the cause of anomalies, exclude non-routine events from savings calculations, and document retained data points to ensure reported savings reflect SEM-driven operational improvements rather than external factors. Future evaluations will track progress using the same metric described under Recommendation 4 above.

---

[26] The DNV team identified this recommendation in the PY2021/22 SEM Evaluation; however, the issue has persisted, with the share of bottom-up projects unable to use top-down modeling approach due to data issues increasing from 15% in PY2021/22 to 31% in PY2023/24.
DNV, *Strategic Energy Management (SEM) 2021-2022 Impact Evaluation, Page 8*. May 25, 2024. PY2021/22 SEM Impact Evaluation

[27] Non-routine event (NRE) – 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.

Recommendations 4–6

- **Ensure top-down model validity or fully adopt bottom-up calculations when necessary**: Implementers should validate the top-down model by confirming variables' relevance and statistical significance. If implementers are unable to establish a valid model, they should fully transition to bottom-up calculations and avoid combining incomplete elements of both approaches as this undermines the validity and defensibility of savings estimates. Future evaluations will measure how often implementers rely on incomplete elements of both top-down and bottom-up approaches to estimate savings.

Bottom-up calculation discrepancies

Our corrections to PY2023/24 bottom-up projects resulted in a 2.6% reduction in total forecasted savings. The DNV team reviewed all submitted calculations for the sampled projects and made corrections to methodologies, inputs, and parameters based on detailed engineering review and as-built information provided by facility personnel. Where available, we relied on extended trend data to better reflect typical annual operating conditions rather than short-term snapshots. These updates ensured that savings estimates were consistent with verified equipment performance and actual facility operations. Key adjustments included:

- For production-based EPIAs,[28] we recalculated key parameters using full-year data when available or the longest and most representative trend period otherwise, replacing values derived from limited datasets.
- For compressed air EPIAs, we updated system efficiencies using verified CAGI[29] performance data, revised as built discharge pressures for pressure reduction measures, and corrected savings where operating conditions differed from implementer assumptions.
- For boiler-related EPIAs, we updated boiler efficiency and loading using the best available information, including facility logs and extended operational data, to reflect typical annual operation.
- For weather-dependent EPIAs, we consistently applied the National Oceanic and Atmospheric Administration (NOAA) station-based weather data and reprocessed weather inputs to ensure accuracy and consistency.

Recommendations 7–9

- **Use representative operational data to support accurate parameter derivation**: Implementers should derive production-, equipment-, and site-specific parameters from representative operational periods, preferably annual trend data when available, or otherwise the longest period that reasonably reflects typical operations, as short or atypical periods can bias savings estimates.
- **Verify all equipment performance and operating parameters using substantiated sources**: Implementers should verify all equipment performance and operating parameters using substantiated sources such as manufacturer documentation, facility-recorded operational data, and trend logs, and ensure that savings inputs reflect actual as-built operating conditions rather than assumed values. This practice will improve the accuracy and consistency of savings calculations. Progress for recommendations 7-8 will be assessed by recording the frequency and relative magnitude of savings corrections due to unsubstantiated operating assumptions and bottom-up savings corrections attributable to non-representative operational datasets.

[28] Energy Performance Improvement Action (EPIA) – A documented operational or capital action that improves facility energy performance and includes required savings and cost information for CPUC reporting.

[29] The Compressed Air and Gas Institute (CAGI) is the industry association that provides standardized performance data for compressed-air equipment. CAGI performance sheets give verified manufacturer data such as CFM per kW, which is used to benchmark compressor efficiency. Their website is https://www.cagi.org/

Recommendations 7–9

- **Adopt a weather-data source and use it across all projects**: With National Oceanic and Atmospheric Administration weather data no longer available after August 2025, the program should identify and adopt an alternative weather-data source and use it across all projects, to ensure accurate and comparable weather-dependent savings calculations across the program. PAs should utilize statewide forums to identify a credible and consistent alternative weather-data source, with coordination with CPUC staff.

Savings proration for bottom-up EPIAs

The SEM M&V Guide[30] requires that implementers prorate the annualized savings for bottom-up EPIAs based on the documented implementation date, with reasonable consideration for seasonality and simplicity. This guidance directs implementers to attribute only the portion of savings realized within the current reporting period and to carry forward the remaining savings into subsequent periods.

During this evaluation cycle, we observed that implementers have begun applying the proration requirement; however, the application of proration was often unclear or inconsistent. Because they often treat proration as a straightforward step, sites apply it differently, and new members of the implementation teams may struggle to determine the correct carry-forward values. Several projects did not document the proration method used, and in some cases, implementers carried forward an entire year of savings rather than only the remaining portion after proration. These issues indicate that SEM participants would benefit from a standardized method for calculating and documenting prorated savings.

Recommendation 10

- **Develop a standardized proration tool:** Implementers should develop a simple, standardized proration tool for bottom-up EPIAs to ensure consistent calculation of prorated savings and clear identification of the savings to carry forward to the next reporting period. Future evaluations will assess whether implementers develop a standardized proration tool and the frequency and magnitude of proration corrections.

Program reporting and tracking data

Participant engagement and evaluation process improvement

During the initial recruitment phase, participants showed relatively strong engagement in evaluation interviews and early data collection, reflecting the broader reach of the SEM program and participants' familiarity with the program. However, early interviews indicated that some energy champions did not have full awareness of their facility's SEM activities, which limited the completeness of gross data collection and NTGR administration. As the evaluation progressed and we requested additional data to verify bottom-up savings, participant responsiveness declined significantly. These challenges became more pronounced at facilities that experienced staff turnover.

[30] Sergio Dias Consulting. "California Industrial SEM M&V Guide, Version 3.02." Section 8.2. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

Recommendations 11–12

- **Support continuity during energy champion transitions**: Implementers should ensure new energy champions can quickly become familiar with prior SEM activities and development history by directing them to existing project documentation, ongoing SEM Activity Logs, and holding a brief orientation meeting. Implementers should document this orientation either as a standalone summary or within the reporting period performance report. This approach maintains continuity at facilities when staff turnover occurs. Future evaluations will verify through the gross data collection interviews with new energy champions whether implementers provided adequate transition support.
- **Emphasize participant ownership of SEM data and documentation**: Implementers should emphasize participant ownership of trend and activity logs, production data, and operational parameters during SEM workshops and coaching to ensure timely support for evaluation data requests. Future evaluations will assess progress through gross savings data collection interviews by verifying energy champions' knowledge of SEM data sources and reduced reliance on implementers, and document it in the CCT. In addition, future evaluations will track the adoption rate of designated SEM Data Leads, consistent with recommendations in the SEM Design Guide.[31]

NTGR methods

The DNV team refined the SEM capital survey instrument developed in the PY2021/22 evaluation and adapted the 2018–2019 TDA instruments for the PY2023 evaluation based on the approved SEM Design Guide rather than an evaluator-developed Program Theory and Logic Model and used the same instrument for PY2024 with additional questions to support the Activity Report. This change does not omit the importance of the program logic models which should continue to be maintained as part of Implementation Plan (IP) submissions and are generally required for IP development.

The Expansion Study[32] linked program success to program delivery that follows the SEM Guides. The Combined TDA Score measures how well the program complies with the letter and spirit of the Guide. Projects that meet a minimum standard of compliance for SEM delivery (currently set at 50%) warrant the SEM NTGR. Compliance and competence standards 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 13–14

- **Use the current NTG survey instruments in future SEM evaluations**: Evaluators should use the TDA and SRA instruments we developed in this impact evaluation to ensure consistency with the findings and recommendations of the Expansion Study. Future evaluations should document the rationale for using alternative NTG survey instruments, if applicable, and obtain concurrence from CPUC staff prior to implementation.
- **Reassess the 50% minimum compliance threshold for SEM NTGR assignment**: Evaluators should revisit the current 50% minimum compliance standard and determine if it remains an appropriate threshold for assigning a site the SEM NTGR (currently 1.0). The 50% threshold may overrepresent meaningful compliance with SEM Guide requirements and require a potential update. Future evaluations should document the basis for retaining or revising the current 50% minimum compliance threshold for SEM NTGR assignment.

[31] Sergio Dias Consulting. "California SEM Program Design Guide Core SEM and Graduate Pathway, Version 2.1." Section 1.2.5 September 1, 2025. California SEM Design Guide: Core SEM and Graduate Pathway

[32] 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) 2024. This evaluation includes SEM projects that completed a full two-year cycle in program year 2024 (PY2024) and will account for the savings claims from the first year of the cycle reported in 2022 or earlier. This study is the second 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 2023 (PY2023) to produce an update to program gross realization rates and net-to-gross ratios.

The overall purpose of this study is **to evaluate Total System Benefit (TSB) energy and demand savings for SEM projects implemented in PY2024.** 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 9, 2025.[33]

2.1 Background

SEM is a unique program under the Energy Efficiency Portfolio. It has a separate statewide Design Guide[34] and M&V Guide[35] that all SEM program implementers must use. As such, the program evaluation requires adherence to the CA evaluation framework[36] and the documented approaches approved for SEM program implementation. Top-down[37] modeling approach is the preferred methodology to calculate SEM savings. The M&V guide requires participants to provide justification if bottom-up calculations[38] are used instead of top-down modeling approach. All SEM energy savings calculations are expected to leverage the existing conditions recorded during the baseline period. The DNV team will also align with the California Public Utilities Commission (CPUC) directive, as decided during PY2021/2022 SEM Impact Evaluation, to review and evaluate SEM projects in a manner consistent with other custom programs, despite SEM's unique design approach.

The CPUC has adopted the Total System Benefit (TSB) as the standard measure to evaluate the time-dependent advantages of energy efficiency. This metric aims to maximize the potential of energy, demand, and greenhouse gas reduction through measures and programs. Starting in 2024, TSB became the sole metric employed to assess energy efficiency programs. However, starting in 2022, Program Administrators (PAs) were required to report TSB alongside kilowatt-hours (kWh), kilowatts (kW), and therms savings. The CPUC and PAs' transition to TSB for evaluating program benefits offers valuable insights by estimating TSB in conjunction with gross savings and net savings.

---

[33] DNV, *Evaluation, Measurement, & Verification of Program Year 2024 Strategic Energy Management Projects Work Plan*, September 9, 2025. https://pda.energydataweb.com/#!/documents/4262/view

[34] 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.

[35] 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

[36] TecMarket Works Framework Team, The California Evaluation Framework. January 24, 2006. CA Evaluation Framework.pdf

[37] **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.

[38] **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

2.2 Evaluation objectives

The primary objectives of this study are to:

- Develop first-year and lifecycle evaluated net and gross savings for the SEM program in PY2024. We will combine the results from this study with the results from the PY2023 study to achieve a high level of precision for the combined two-year period of PY2023/2024.
- Calculate TSB based on evaluated energy savings results and compare them with PA claimed values for PY2023/2024.
- Develop meaningful, actionable recommendations to improve program performance in delivering energy efficiency savings.

The key research questions for this impact evaluation are:

- What are the PY2024 and PY2023/2024 first-year and lifecycle gross kWh, peak kW, and therm savings by sampling domain (e.g., analysis approach, sector, PA)?
- What is the overall Total System Benefit (TSB) and how does that compare to forecasted TSB claims by PA?[39]
- 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 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

[39] CPUC staff guidance (per D.21-05-031) defines the Total System Benefit (TSB) metric, which replaces traditional savings goals with a dollar-based value reflecting grid benefits. It also outlines calculation steps and reporting via the Cost Effectiveness Tool (CET) and CEDARS.
CPUC, Total System Benefit Technical Guidance, Version 1.2. October 25, 2021. Final Total System Benefit Technical Guidance Document.

- 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, including the planned sample design, achieved sample sizes, gross savings, measurements and verification (M&V) activities, net savings approach, and final expansion procedures.

To meet the 2026 bus stop (May 1, 2026) for PY2024 and PY2023/2024, DNV followed the custom evaluation protocols consistent with the custom reporting process. However, the SEM program is unique and has its own M&V framework outlined in the M&V guide for the SEM program. As a result, we defined the impact evaluation approach for SEM separately and evaluated it outside of the custom program framework.

DNV calculated TSB-based evaluated savings results and reported and evaluated TSB, and compared them with PA-reported TSB values. We conducted the analysis using the Cost Effectiveness Tool (CET) available on the CEDARS website. Presenting both reported and evaluated savings provides stakeholders with an accurate delivered value of TSB. For PY2024, the TSB analysis is informational only, with no guarantees of meeting precision targets for TSB, because SEM was not using TSB in some of the previous years' claims still captured in this impact evaluation.

3.1 Sample design

The SEM evaluation spans two years, PY2023 and PY2024. Thus, sampling occurred in two waves, one wave for each year. We drew the first wave sample in 2024 using final PY2023 claims data. We drew the second wave sample using the final PY2024 claims data. The two-wave sampling approach allows the DNV team to achieve the precision target of ±10% at 90% confidence interval over two years.

This report presents the combined results of both Wave 1 and Wave 2 SEM projects from claims reported in PY2023 (Year 1) and PY2024 (Year 2). To achieve the two-year precision target of 90/10, each of the two evaluation waves had 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. It groups participants 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 evaluates all claimed savings for Cycles 1, 2, and 3 participants, shown in Table 3-1. This approach aligns with the program design and activity plan.

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. These parameters were not captured in California Energy Data and Reporting System (CEDARS) data in the first phase of this evaluation but were updated and became available starting in PY2025.

3.1.1 Gross and net savings and sample design overview

Table 3-2 summarizes the key assumptions for the sample allocations. Subsequent sections show the final design and associated estimated precisions.

Table 3-2. Combined subject area sample design assumptions

Approach to estimating claims	Non-claimable improvements
Population	Tracking data set for program year, including all SEM projects reported in claims with positive, negative, or zero savings.
Sampling variable	Site MMBtu savings combining electric and gas source Btus
Sample design approach	Stratified ratio estimation
Explicit sampling strata	Tracking savings Program year (PY2024 program data / PY2023 program data)
Implicit sampling strata	As supported by the data, fuel, PA, measure type, analysis approach, CPR status, cohort, program cycle
Gross sample allocation	40 projects, allocated by size (tracking MMBtu) for best overall precision
NTGR annual sample allocation	All gross savings sample projects
Target parameters	GRR, NTGR
Analysis domains	PA, Fuel
Error ratios	By overall site MMBtu savings based on the findings of the PY2023 SEM Evaluation
Precision targets at 90% confidence (planned)	Gross MMBtu savings by energy unit: ±8% Gross MMBtu savings by energy unit (electric): ±10% Gross MMBtu savings by energy unit (gas): ±5%
Savings size stratification	Up to four levels based on savings (3 certainty strata for the largest sites).[40]
Contingency and back-up sample	All remaining sites that were not selected into the sample will be pre-sorted into random selection sequence as a backup list to account for projected ineligible and nonresponse rates. We have employed strategies as part of the sample design process to ensure the target precision is achieved. Specifically, a conservative error ratio of 1.0, which is based on the PY2023 results, has been applied to meet the evaluation goal of ±10 relative precision at the statewide level.

We optimally designed the sample allocation to minimize the sample required to achieve the target statewide precision requirements. We based the estimated precisions at 90% confidence for the total electric and gas lifetime gross energy savings (MMBtu) and the assumed error ratio.

3.1.2 Gross sample completions and response rates

Table 3-3 presents the population counts, sample design quotas, and final achieved sample for key analysis dimensions, including PA and fuel, for the gross sample. The electric fuel type stratum targeted a ±10% relative precision at 90% confidence interval, while the gas fuel type targeted a ±5% relative precision at 90% confidence interval. Achieving the precision targets of ±14% relative precision for each program year (PY2023 and PY2024) resulted in meeting the combined cycle precision target of ±10% relative precision for both PY2023 and PY2024. Overall, we recruited 90% of electric projects and 88% of gas projects targeted for the primary sample.

[40] Although the three largest strata were certainty strata or attempted census, in reality, we assumed that we would not be able to recruit all sites selected into the sample. For this reason, we created three separate certainty strata, so that if we were unable to recruit sites within any of the three cells we could weight the achieved sample in each stratum back to its respective stratum population.

Table 3-3. PY2024 gross sample response rate by fuel and PA

PA	Electric				Natural gas			
	Population (N)	Sample design quota	Final sample (n)[a]	% complete	Population (N)	Sample design quota	Final sample (n)[a]	% complete
MCE	2	1	1	100%	1	1	1	100%
PG&E	36	20	20	100%	27	20	18	90%
SCE[1]	53	9	6	67%	N/A	N/A	N/A	N/A
SoCalGas[2]	N/A	N/A	N/A	N/A	5	5	4	80%
Total	**91**	**30**	**27**	**90%**	**33**	**26**	**23**	**88%**

[a] The final electric and natural gas samples include 15 sites reporting savings for both fuels. As a result, the PY2024 gross sample consists of 12 electric-only sites, 9 gas-only sites, and 15 dual-fuel sites, which are reflected in both fuel-specific totals, totalling 35 sites for PY2024.
[1] Electric only.
[2] Gas only.

Table 3-4 presents the results of the PY2024 gross data collection. We evaluated 35 sites across the four PAs. The PY2024 sample was designed using a stratified analysis approach based on program year and size of tracked savings (MMBtu). The PA dimension was not included in the sample design due to limited population sizes at the PA level within PY2024. For estimation purposes, the finalized PY2024 sample and population were post-stratified by savings size, with PA applied as a post-stratification dimension.

Table 3-4. PY2024 gross sample post-stratification

PA	Stratum	Strata maximum	Accounts population	Forecasted savings MMBtu	Accounts sample	Weight per site
MCE	1	1,641	2	2,648	1	2.00
SCE	2	6,258	47	39,522	4	11.75
SCE	3	23,747	6	63,986	2	3.00
SCG	4	26,580	3	62,561	2	1.50
SCG	5	53,175	2	96,900	2	1.00
PG&E	6	56,043	36	467,723	18	2.00
PG&E	7	209,356	6	662,379	5	1.20
PG&E	8	1,084,520	1	1,084,520	1	1.00
Total			**103**	**2,480,238**	**35**	

This final evaluation includes results from PY2023 and PY2024. Table 3-5 presents the results of the gross data collection. A total of 49 sites were recruited across both years. The final combined two-year sample and population were post-stratified by size and PA. The sample was designed with a stratified analysis approach based on program year (PY2023 or PY2024 program data) and size of tracking savings (MMBtu) and did not include the dimension of PA. The PA dimension was not included due to the small population size for each program year. Combining the two program years allowed for post stratification based on PA as well as savings size.

Table 3-5. PY2024 & PY2023 combined gross sample post-stratification

PA	Stratum	Strata maximum	Accounts population	Forecasted savings MMBtu	Accounts sample	Weight
MCE	1	3,491	5	7,568	3	1.67
MCE	2	13,113	1	13,113	1	1.00
SCE	3	6,258	71	80,988	6	11.83
SCE	4	23,747	11	122,899	7	1.57
SCG	5	17,550	10	91,987	5	2.00
SCG	6	53,175	4	149,599	3	1.33
PG&E	7	56,043	36	467,723	18	2.00
PG&E	8	209,356	6	662,379	5	1.20
PG&E	9	1,084,520	1	1,084,520	1	1.00
Total			**145**	**2,680,776**	**49**	

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 included interviews with customer and implementer 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 and are described further in Section 3.5.3. 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 to understand their design and effectiveness and ensure consistent compliance with SEM principles.

Table 3-6. PY2024 disposition of interviews for SRA and TDA methods

Interviewee type	Interview purpose	Method	PY2024 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	40	27[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	10	8[2]
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	5 coaches, covering 11[3] sites

1. Sent email invitations to all 40 Energy Champions in the sample and completed 31 interviews (only 27 used in the analysis and reporting). 4 interviews dropped from the analysis due to incomplete data or late data coming in after weighting was complete
2. Sent email invitations to all Executive Sponsors for the sites we completed Energy Champion interviews. Most Energy Champions provided new contact information for their Executive Sponsor. Most Executive Sponsors did not respond to our emails inquiring for interviews. 3 Executive Sponsors were also Energy Champions and answered both sets of questions. Several executive sponsors effectively served as energy champions and brought a broader, site-level perspective. All executive sponsors interviewed demonstrated strong familiarity with the program and its impacts on their site. Executive sponsors not interviewed did not respond to multiple requests from evaluation team.
3. Sent email invitations to all of the Energy Coaches and got responses from only five prior to the analysis being complete. Two coaches reached out after analysis. Five energy coach interviews were completed, representing 10 sites. Energy coaches were highly engaged, and all but one served as coaches for sites included in the evaluation. This remaining coach was listed erroneously as a coach in the data provided.

Table 3-7 presents the net analysis recruitment table for the PY2024 results. Seventeen sites were included in the net analysis across the four PAs. Sites were eligible for inclusion only if they successfully completed the gross evaluation and could be assigned a defensible site-level NTGR.

Consistent with the NTGR framework, eligible sites fell into two categories: (1) sites that did not report any capital measures and were therefore assigned a site-level NTGR of 1.0 which is consistent with the result of the PY2024/2023 TDA analysis, and (2) sites that reported capital measures, completed an Energy Champion interview, and received a valid SRA score. While 27 Energy Champion interviews were used in the NTG analysis, only 17 of those sites both remained in the finalized gross evaluation sample and had usable SRA scoring, making them suitable for inclusion in the net analysis. We also excluded several for which we completed interviews because they were dropped from the gross evaluation sample due to missing data critical to gross savings estimation. In addition, a small number of sites experienced staff turnover such that the Energy Champion interviewed lacked sufficient tenure or institutional knowledge to support a defensible SRA score. As a result, the final net analysis reflects 17 evaluated sites.

Table 3-7. PY2024 net sample recruitment for NTGR

PA	Stratum	Strata maximum	Accounts recruited for gross	Forecasted savings MMBtu	Accounts recruited for net
MCE	1	1,008	1	1,007	0
SCE	2	23,747	6	36,640	1
SCG	3	53,175	4	132,882	2
PG&E	4	61,007	19	440,340	10
PG&E	5	209,356	4	525,114	3
PG&E	6	1,084,520	1	1,084,520	1
Total			**35**	**2,220,503**	**17**

The NTG team combined results from PY2024 and PY2023 into a single analysis with post-stratified results. A single site had gross negative evaluated savings in PY2023. This site was evaluated and the expansion was unit weighted, or given a weight of one, so that the results represent that site alone and are not projected to the rest of the population.

Table 3-. PY2024 & PY2023 combined net sample post-stratification

PA	Unit weight	Stratum	Strata maximum	Accounts population	Forecasted savings MMBtu	Accounts sample	Weight
MCE	No	1	13,113	6	20,681	2	3.00
SCE	Yes	2	(826)	1	(827)	1	1.00
SCE	No	3	6,258	71	80,988	4	17.75
SCE	No	4	23,747	11	122,899	3	3.67
SCG	No	5	17,550	10	91,988	3	3.33
SCG	No	6	53,175	4	149,599	2	2.00
PG&E	No	7	56,043	36	467,723	10	3.60
PG&E	No	8	209,356	6	662,379	3	2.00
PG&E	No	9	1,084,520	1	1,084,520	1	1.00
Total				**146**	**2,679,950**	**29**	

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—which other evaluation teams, such as CIAC, also used 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 40 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, they did not do so 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 we needed more details 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 we developed and the PAs accepted during the PY2023/2024 Study. Before starting 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 aligned with expectations? What are the reasons for the measure discontinuance?
3. Does the measure impact a single IOU meter? We will use this information 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, we collected any additional measure-specific information or customer feedback on their program participation.

For bottom-up and mixed[41] 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 early-October 2025 and concluded in early January 2026. These efforts engaged one lead recruiter and two active recruiters. 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 failed 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. We escalated cases to the PAs and CPUC as necessary.

We were able to fulfill our target of 40 sites recruited. Of these sites, 38 were part of the primary sample and 2 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 supported follow-up site-specific data requests after initial interviews when customers became non-responsive. After three to four unsuccessful attempts to obtain the requested data, the PAs conducted additional outreach and leveraged their existing relationships to obtain the information needed to complete the analysis. In total, 19 sites required escalation to PA and CPUC involvement, and the DNV team was able to obtain data for 14 sites within the established timeframe. Despite additional outreach and assistance from SCG, PG&E, and the CPUC, data necessary to complete the gross evaluation could not be obtained for the remaining five sites (three SCE, one PG&E, and one SCG). As a result, the DNV team was unable to calculate site-specific gross results for these sites, and they were excluded from the overall PY2024 study results calculations.

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 follow the table.

[41] 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[42]

Parameter	Top-down	Bottom-up	Mixed analysis (top-down & bottom-up)
Reviewed projects, n = 40	12	15	13
Percentage by count	30%	38%	33%
Percentage of electric savings	41%	41%	18%
Percentage of gas savings	17%	60%	24%
Percentage of overall MMBtu savings	21%	56%	22%

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, the SEM M&V guide designates this method 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 the DNV team completed for each savings calculation.

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.

[42] We provide a further breakdown of savings calculation methodologies by PA and fuel in APPENDIX G.

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 PY2021/22, participants continued to estimate demand savings for projects completed in PY2023/2024 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 PG&E, 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 PG&E 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 PG&E and MCE.

According to the PY2018/2019 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)"*.[43] 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.

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.

---

[43] 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

The DNV team estimate program savings 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 3.1.3, 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

In accordance with the CPUC's SEM program guidelines and the CPUC decision enforcing a standard 5-year SEM EUL, DNV assigned a 5-year EUL to all evaluated projects. To maintain consistency, when calculating the combined results for PY2023 and PY2024, the DNV team updated the PY2023 projects to reflect a 5-year EUL. This ensures alignment across program years and adherence to the current SEM program requirements.

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, we employed two evaluation methods for net savings:

- **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 that can be investigated by interviewing 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.
- **Theory-driven attribution (TDA)**. We used the theory-driven approach 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. We adapted the methods and survey instruments developed in the 2018/2019 SEM Study with further improvements and optimization.

We also note that the SEM Activity Report (targeting a July 2026 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.

3.5.1 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. They were adapted from the Commercial, Industrial, and Agriculture Custom (CIAC) NTG survey battery.[44] 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. CPUC Staff reviewed and approved these proposed revisions for application beginning with the PY2023 evaluation, and they were shared with PAs during a statewide ED/PA Update Meeting.

The PY2021/2022 Study defined the protocol for identifying capital measures. The protocol starts with a review of the Opportunity Register to inventory measures classified as a capital measure. The capital classification is then confirmed by customer staff in the NTG battery.

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 PY2018/2019 Study[45] 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 PY2018/2019 Study. As in the PY2018/2019 Study, individual scores were developed for a sample of completed projects. Evidence for scoring was collected through in-depth interviews of multiple stakeholders and

[44] 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

[45] 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 PY2018/2019 Study methods for the PY2023/2024 Study. The PY2023/2024 Study built upon the framework of the PY2018/2019 Study, improving it by incorporating subsequent research, most notably the "Group D Strategic Energy Management Expansion Study"[46] ("Expansion Study"). This study was commissioned by the CPUC as directed by Decision 23-02-002[47] 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 PY2018/2019 Study compared PA program delivery to a PTLM that the evaluator developed. This evaluation compares PA program delivery to the requirements identified in the SEM Design Guide based on the recommendations 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 Guides 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. The PTLM serves multiple purposes and is still a valuable tool in ensuring program alignment with the SEM Guides, and should be revised in concert 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).[48] 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 SEM 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 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.

---

46 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

47 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

48 The DGRs are listed in Table 4-12 with further descriptions in the Appendix.

Following the methodological precedent of the PY2018/2019 impact evaluation, each site outcome was tested using a preponderance of evidence (POE) standard 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.[49] 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 the site case weight and savings of the site.

The evaluation included a sensitivity analysis, which is reported in APPENDIX A, to examine the impacts of revising the SRA algorithm to better account for 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. These three 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 complied fully with the CPUC's Self-Report Guidelines. The lengthy interviews with Energy Champions also contributed significantly 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 – Other facility staff engaged with the day-to-day SEM activities at a site. The PY2024 evaluation effort did not attempt to reach out to this group due to the high levels of turnover found during the FY2023 evaluation and difficulty in reaching site staff.

We note this data collection will also support the SEM Activity Report, targeting a July 2026 publication.

---

[49] *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

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.

3.6 Total system benefit approach

Beginning in PY2024, the CPUC adopted Total System Benefit (TSB) as the primary performance metric for energy efficiency programs. TSB monetizes the lifecycle energy, capacity, and greenhouse gas benefits of a project, promoting high-value load reduction, long-duration savings, and fuel-agnostic program design.

For this evaluation, the DNV team applied a TSB estimation method consistent with the approach used for program-wide energy savings. The evaluated TSB values were developed through a multi-step process:

1. **Calculate evaluated TSB for each claim year**: For every project, the DNV team updated the claim-year TSB values by replacing forecasted inputs with evaluation results, including ex post energy savings, MAT or EUL adjustments informed by evaluation findings, and NTGRs. This produced an evaluated TSB value for each claim year associated with the project.
2. **Aggregate evaluated TSB to the project level**: The claim-year TSB values were summed to produce a project-level evaluated TSB that reflects the full set of evaluated benefits for that project
3. **Expand project-level TSB to statewide results**: Consistent with the treatment of energy savings, the evaluated project-level TSB values were expanded using impact sample design weights, generating a statewide TSB estimate for the SEM portfolio.

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 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.

4.1.1 TSB Savings

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

Table 4-1. PY2023/2024 TSB savings by PA

	First year (1,000 TSB)			
Program administrator	**Forecasted TSB [a]**	**Evaluated TSB**	**GRR**	**Relative precision**
MCE	1,335	780	58%	13%
PG&E	168,923	165,697	98%	1%
SCE	23,616	15,838	67%	24%
SCG	14,115	12,341	87%	24%
Statewide	207,990	194,657	94%	3%

[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 Electric 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 electric savings. Relative precisions are provided at the 90% confidence interval.

Table 4-2. PY2023/2024 Gross electric energy and demand savings by PA

	First year				Lifecycle			
Program Administrator	**Forecasted savings [a]**	**Evaluated savings**	**GRR**	**Relative precision**	**Forecasted savings [a]**	**Evaluated savings**	**GRR**	**Relative precision**
Energy (MWh)								
MCE	1,018	1,012	99.4%	0.9%	5,090	5,058	99.4%	0.9%
PG&E	89,706	86,585	96.5%	2.3%	448,532	432,927	96.5%	2.3%
SCE	59,764	42,159	70.5%	21.0%	298,820	169	70.5%	21.0%
Statewide	**150,488**	**129,756**	**86.2%**	**7.0%**	**752,442**	**438,154**	**86.2%**	**7.0%**
Demand (MW)								
MCE	0.1	0.2	126.2%	27.7%	0.7	0.9	126.2%	27.7%
PG&E	12.6	12.1	96.6%	2.3%	62.8	60.7	96.6%	2.3%
SCE	7.6	5.2	68.5%	22.9%	38.1	26.1	68.5%	22.9%
Statewide	**20.3**	**17.5**	**86.3%**	**7.0%**	**101.6**	**87.7**	**86.3%**	**7.0%**

[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 Gas savings

Table 4-3 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 4-3. PY2023/2024 Gross gas energy savings by PA

Program Administrator	First year				Lifecycle			
	Forecasted savings [a]	Evaluated savings	GRR	Relative precision	Forecasted savings [a]	Evaluated savings	GRR	Relative precision
Energy (1,000 Therms)								
MCE	172	89	51.7%	12.4%	860	445	51.7%	12.4%
PG&E	19,085	18,755	98.3%	1.3%	95,427	93,774	98.3%	1.3%
SCG	2,416	2,186	90.5%	25.0%	12,079	10,932	90.5%	25.0%
Statewide	**21,673**	**21,030**	**97.0%**	**2.8%**	**108,366**	**105,151**	**97.0%**	**2.8%**

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

4.1.4 Total MMBtu savings

Table 4-4 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 4-4. PY2023/2024 Gross total energy MMBtu savings by PA

Program Administrator	First year				Lifecycle			
	Forecasted savings [a]	Evaluated savings	GRR	Relative precision	Forecasted savings [a]	Evaluated savings	GRR	Relative precision
Energy (1,000 MMBtu)								
MCE	21	11	55.0%	12.0%	103	57	55.0%	12.0%
PG&E	2,215	2,170	98.0%	1.2%	11,073	10,852	98.0%	1.2%
SCE	204	144	70.5%	21.0%	1,019	719	70.5%	21.0%
SCG	242	219	90.5%	25.0%	1,208	1,093	90.5%	25.0%
Statewide	**2,681**	**2,544**	**94.9%**	**2.6%**	**13,404**	**12,721**	**94.9%**	**2.6%**

4.1.5 Discrepancy analysis

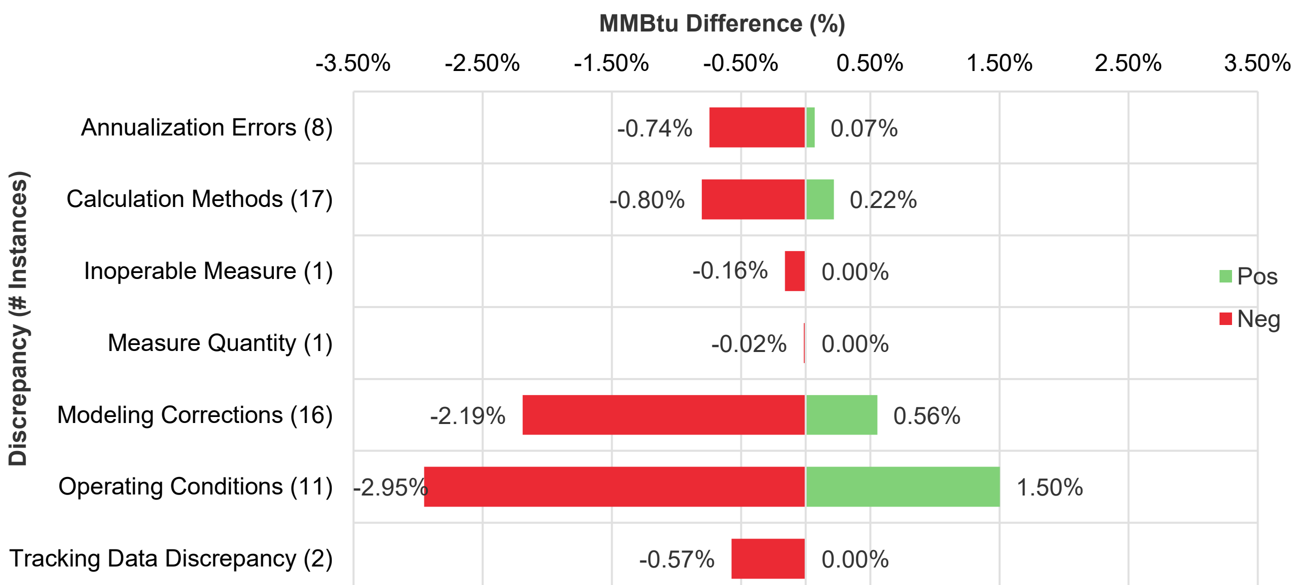
This section presents an analysis of what caused forecasted savings to differ from the evaluated savings estimates for the sampled projects. This analysis is based on the discrepancies associated with first-year gross savings and is calculated on a MMBtu basis. Table 4-5 summarizes the discrepancy categories that led the evaluated savings to differ from forecasted savings.

Table 4-5. Categories of savings discrepancies

Category	Description
Annualization errors	Differences attributed to the use of savings annualization, where partial-period savings are extrapolated to a full year. While permitted under prior SEM M&V guidance, this approach can misrepresent savings by not fully accounting for seasonal or non-representative operations.
Calculation methods	Differences attributed to evaluator-driven updates to savings calculation approaches or analytical procedures when necessary to accurately estimate savings based on site-verified data or correct limitations in the forecasted analysis (e.g., revised normalization techniques, corrected parameter application, or use of full-year 8,760-hour modeling)
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.
Measure quantity	Differences attributed to changes in the quantity of equipment installed, where the number of measures verified by evaluation teams is higher or lower than the quantity reported by the implementer (e.g., number of lights installed).
Modeling corrections	Differences due to evaluator updates to top-down models (e.g., outlier or NRE removal, weather data corrections, incremental savings adjustment, or variable refinement) to improve model accuracy and representativeness.
Operating conditions	Differences attributed to changes in actual facility operating conditions relative to assumptions used in the forecasted savings analysis. These differences are informed by evaluation-collected evidence (e.g., trend data, site documentation, photographs, or interviews) and may include changes in operating hours, production levels, load profiles, control set points, or evaluation period coverage.
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.

As the DNV team developed site-specific evaluated savings for each sampled site, we documented the drivers of any deviations between forecasted and evaluated savings using the identified discrepancy categories. We then quantified the contribution of each discrepancy category to the overall difference between forecasted and evaluated savings.

Figure 4-1. Key drivers behind overall GRR (MMBtu)



Further discussion on each discrepancy is provided below:

Annualization errors – This discrepancy primarily reflected the extrapolation of savings from short, seasonally unrepresentative periods (typically three to five months) to a full year, which often overstated forecasted savings and resulted in negative evaluation adjustments. Although these approaches followed prior SEM M&V guidance, evaluation results indicate that they produced inaccurate savings by failing to capture full-year operating conditions. In fewer instances, evaluation annualization of representative operating days increased savings where valid periods had been excluded. Overall, this category contributed a net reduction to evaluated GRR. Annualization errors were observed predominantly in PY2023, reflecting continued application of annualization practices from prior SEM M&V guidance, the impacts of which are documented in the PY2023 SEM evaluation findings. Overall, annualization errors contributed a net reduction of 18,074 MMBtu to evaluated GRR.

Calculation methods – These discrepancies were driven by a combination of corrected equipment specifications and updates to savings calculation inputs and assumptions. In some cases, site-collected data and document review showed that installed equipment characteristics differed from implementer assumptions, leading to revised specifications and associated changes in evaluated savings. In other cases, evaluators updated inputs and assumptions within the savings equations used by implementers while retaining the same overall calculation methods, based on verified as-built information collected at the site. Additional discrepancies resulted from the application of a different ex post M&V period and from normalization of reported data to align post-period measurements with baseline conditions or verified equipment counts. In fewer cases, data-driven refinements increased evaluated savings. Overall, this category produced a net negative impact on evaluated GRR. Overall, discrepancies in this category resulted in a net reduction of 15,732 MMBtu in evaluated GRR.

Inoperable measures – This discrepancy resulted in a negative adjustment to evaluated savings for a PY2024 site where verified site conditions showed that the measure was no longer in operation. In the single observed instance, the equipment was removed within five years of installation due to space limitations and reprioritization of facility space for other uses. As the measure was not operational during the evaluation period, associated savings could not be sustained and were therefore removed from evaluated GRR. Overall, discrepancies in this category resulted in a net reduction of 4,303 MMBtu in evaluated GRR.

Measure count – This discrepancy arose for a PY2024 site where the number of installed measures differed from implementer assumptions. In the single observed instance, site verification confirmed that fewer dryer doors were installed than assumed by the implementer, resulting in reduced evaluated gas savings. Overall, discrepancies in this category resulted in a net reduction of 403 MMBtu in evaluated GRR.

Modeling corrections – These discrepancies reflect evaluator-driven refinements to top-down models based on improved representation of actual facility conditions and analytical performance. Common drivers included removal of statistical outliers or non-routine events, correction of weather data sources, and refinement of independent variables to better reflect normal site operations. These adjustments were applied when implementer models relied on anomalous data, mismatched inputs, or variables that did not adequately explain observed energy use. While a limited number of corrections increased evaluated savings, most adjustments reduced savings by eliminating unsupported or unrepresentative influences, resulting in a net negative impact on the evaluated GRR. Overall, discrepancies in this category resulted in a net reduction of 43,866 MMBtu in evaluated GRR.

Operating conditions – These discrepancies reflect verified changes in site operations relative to forecasting assumptions, including changes in load profile, production, set points, and operating hours. Most instances resulted in negative adjustments where actual production, operating hours, load shapes, or control settings led to reduced evaluated savings relative to assumptions. In a limited number of cases, corrected set points increased evaluated savings where implementer assumptions understated performance. Overall, this category contributed the largest net reduction to evaluated GRR, underscoring the sensitivity of SEM savings to real-world operational conditions. Overall, discrepancies in this category resulted in a net reduction of 38,850 MMBtu in evaluated GRR.

Tracking data discrepancy – These discrepancies resulted from inconsistencies between program-reported savings and the savings calculated and documented by the implementer in project documentation submitted for evaluation. In the two observed PY2024 instances, evaluator-verified savings matched implementer calculations, but differed from program-reported values due to inaccuracies in tracking data. Overall, discrepancies in this category resulted in a net reduction of 15,369 MMBtu in evaluated GRR.

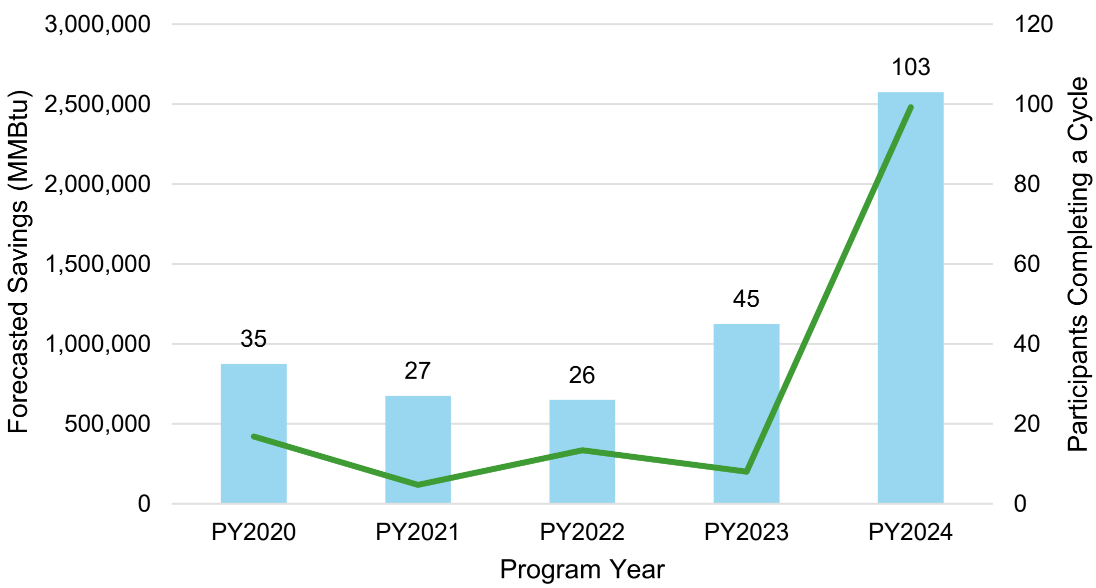
4.1.6 Comparison to previous evaluation findings

4.1.6.1 Program size and participation trends

Since PY2020, participation in the SEM program has shown moderate year-to-year variability, followed by a substantial increase in PY2024. The SEM program's multi-year structure—comprising three two-year cycles—has contributed to strong participant retention, as most facilities that joined at program launch have continued through successive cycles, with many now approaching program graduation. While a small number of participants have exited or paused participation, ongoing recruitment has more than offset these losses.

Figure 4-2 summarizes growth in SEM participation among sites that completed a cycle in each program year and were therefore eligible for evaluation, along with associated forecasted savings. After relatively stable participation levels between PY2020 and PY2022, the number of evaluated participants increased substantially in PY2023 and more than doubled again in PY2024, marking the largest growth observed to date. This growth was accompanied by a commensurate increase in forecasted savings, reflecting both higher participation and the cumulative impacts of multi-cycle engagement.

Figure 4-2. Historical SEM participation and forecasted savings



Recent growth also reflects program expansion beyond its traditional focus on industrial facilities. Historically, SEM eligibility was limited primarily to industrial participants, with only limited exceptions in certain jurisdictions. In PY2024, commercial participants represented approximately 17% of sites completing a cycle, though they contributed a smaller share of savings (approximately 3%). As commercial recruitment continues, participation by this segment is expected to increase gradually, with corresponding impacts on future evaluation cycles.

4.1.6.2 Gross savings realization rates

At a high level, statewide overall savings realization rates differed between the PY2021/2022 and PY2023/2024 SEM evaluations, reflecting both program evolution and changes in how forecasted savings aligned with evaluated results. While both evaluation periods exhibited net reductions from forecasted to evaluated savings, PY2023/2024 results indicate improved alignment overall. Table 4-6 summarizes statewide gross realization rates (GRRs) for electric, gas, and MMBtu across the PY2021/2022 and PY2023/2024 evaluation cycles.

Table 4-6. Comparison of statewide gross realization rates (PY2021/2022 vs. PY2023/2024)

Evaluation cycle	Total savings (MMBtu)	Overall RR (%)	Electric savings (MWh)	Electric RR (%)	Gas savings (1000 therms)	Gas RR (%)
PY2021/2022	453,575	85.0%	42,076	98.3%	3,101	79.0%
PY2023/2024	2,680,776	94.9%	150,488	86.2%	21,673	97.0%

Total System Benefits (TSB) impacts were not calculated for the PY2021/2022 evaluation and are therefore not included in this table.

Differences in savings realization rates across the two evaluation periods are attributable to changes in the relative contribution of key discrepancy drivers. In PY2021/2022, realization rates were primarily affected by annualization errors that overlooked operation seasonality by extrapolating savings from short or partially representative periods to full-year values, which frequently overstated forecasted savings when evaluated against longer data records. Modeling corrections related to calibration, variable selection, and treatment of anomalous periods further contributed to downward adjustments.

In PY2023/2024, annualization was a less dominant contributor, reflecting a gradual shift away from short-period extrapolation and greater reliance on full-year or otherwise representative data in savings estimation. At the same time, modeling corrections and operating-condition-driven adjustments became more prominent, driven by increased use of site-collected data, refined treatment of non-routine events, and improved representation of actual facility operations. While some adjustments resulted in increased evaluated savings where implementer assumptions were conservative, the net effect across discrepancy categories remained negative.

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. We have expanded savings from the site level to the population level using the site weights provided in the sample design, as described in Section 3.1.

We note that the statewide NTGR exceeded 96% with very good relative precisions of ±1%–2.5%. This reflects that a minority of MMBtu savings (14%) derived from capital measures (n=22 projects) achieved an average unweighted NTGR of 0.65, while the balance of the measures were factored into the analysis with a NTGR of 1 with no additional variation. PA results were similar, except MCE’s and SCG’s gas results, which were poorer.

4.2.1 Electric net savings

PY2023/2024 SEM program participants achieved an aggregated first-year net electric savings of 126,087 MWh with a statewide NTGR of 97.2%. These relative precisions exceed targets largely due to the application of the NTGR of 1 to the non-capital portion of savings (74% of electric FYS savings) and the relatively low variation of the SRA NTGR to the balance of savings from capital measures (26%). The PY2023/2024 statewide NTGR aligns closely with the PY2021/2022 two-year result of 96.9%.

Table 4-7 presents the first-year and lifecycle forecasted savings, evaluated gross and net savings, NTGR, and relative precision at the 90% confidence interval at the PA and statewide levels for electric savings.

Table 4-7. PY2023/2024 electric net savings and NTGR

Program Administrator	First-year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Energy (MWh)								
MCE[1]	1,012	983	97.2%	1.0%	5,058	4,915	97.2%	1.0%
PG&E	86,585	83,755	96.7%	1.3%	432,927	418,777	96.7%	1.3%
SCE	42,159	41,797	99.1%	1.3%	210,795	208,986	99.1%	1.3%
Statewide	**129,756**	**126,087**	**97.2%**	**1.0%**	**648,780**	**630,435**	**97.2%**	**1.0%**
Demand (MW)								
MCE[1]	0.2	0.2	99.3%	1.3%	0.9	0.9	97.2%	1.0%
PG&E	12.1	11.7	96.7%	1.3%	60.7	58.7	96.7%	1.3%
SCE	5.2	5.2	99.3%	1.3%	26.1	25.9	99.3%	1.3%
Statewide	**17.5**	**17.4**	**99.3%**	**1.0%**	**87.7**	**85.2**	**97.2%**	**1.0%**

1 We were unable to recruit any MCE sites with electric savings; therefore, we applied statewide results for MCE.
SDG&E did not report any SEM savings in PY2023 or PY2024. This is due to administrative timing shifts that delayed their reporting to PY2025.

4.2.2 Gas net savings

PY2023/2024 SEM program participants achieved an aggregated first-year net natural gas savings of 20.2 million therms with a statewide NTGR of 96.0%. These precisions exceed targets largely due to the application of the NTGR of 1 to the non-capital portion of savings (88% of natural gas FYS savings) and the relatively low variation of the SRA NTGR to the balance of savings from capital measures, which represents 12% of the gas saving claims. The PY2023/2024 statewide NTGR is within the error of the previous two-year PY2021/2022 result of 96.5%.

Table 4-8 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.

Table 4-8. PY2023/2024 natural gas net savings and NTGR

Program administrator	First-year				Lifecycle			
	Evaluated gross savings	Net savings	NTGR	Relative precision	Evaluated gross savings	Net savings	NTGR	Relative precision
Gas (Therms/1000)								
MCE	89	61	68.9%	21.4%	445	307	68.9%	21.4%
PG&E	18,755	18,350	97.8%	1.3%	93,774	91,751	97.8%	1.3%
SCG	2,186	1,885	86.2%	18.1%	10,932	9,424	86.2%	18.1%
Statewide	**21,030**	**20,198**	**96.0%**	**2.5%**	**105,151**	**100,991**	**96.0%**	**2.5%**

SDG&E did not report any SEM savings in PY2023 or PY2024. This is due to administrative timing shifts that delayed their reporting to PY2025.

4.2.3 Total MMBtu savings

Table 4-9 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 4-9. PY2023/2024 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 (1000 MMBtu)								
MCE	11	8	68.9%	21.4%	57	39	68.9%	21.4%
PG&E	2,170	2,119	97.6%	1.1%	10,852	10,594	97.6%	1.1%
SCE	144	143	99.1%	1.3%	719	713	99.1%	1.3%
SCG	219	188	86.2%	18.1%	1,093	942	86.2%	18.1%
Statewide	**2,544**	**2,450**	**96.3%**	**2.0%**	**12,721**	**12,250**	**96.3%**	**2.0%**

SDG&E did not report any SEM savings in PY2023 or PY2024. This is due to administrative timing shifts that delayed their reporting to PY2025.

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 4-10 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 14% of total program MMBtus, is less than PY2021/2022 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 4-10. Comparison of the PY2023/PY2024 NTGR with other NTGR results

Program and year	Electric NTGR	RP	NG NTGR	RP
SEM 2023/2024 final results	97.2%	1.0%	96.0%	2.5%
Capital, information only, electric and gas	0.71	N/A	0.71	NA
Recommended value	1.0	N/A	1.0	NA
SEM 2024 results	97.2%	1.0%	95.4%	2.1%
Capital, information only, electric and gas	0.65	N.C.	0.65	N.C.
SEM 2023 preliminary interim results	99.1%	0.9%	96.9%	1.3%
Capital only, information only	0.95	N.C.	0.75	N.C.
SEM 2021/2022	98.4%	0.6%	98.0%	2.5%
Capital only, information only	0.80	N.C.	0.74	N.C.
Recommended value	1.0	N.C.	1.0	N.C.
SEM 2018/2019	100.0%	N/A[1]	100.0%	N/A[1]
Core NTGR, info only	0.61	N.C.	0.61	N.C.
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%

[1] N/A = Not applicable. Relative precision is not applicable because results are based on a census sample rather than a statistical sample.
N.C. = Not calculated

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 4-11 shows the percentage of capital measures of program forecasted savings by PA. Capital measures, as reported in the Opportunity Registers, make up about 14% of PY2023/2024 forecasted MMBtu 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 unweighted average SRA-method NTGR of capital measures.

Table 4-11. PY2023/2024 contribution of capital measure to SEM program forecasted savings

PA	Total MMBtu savings	Percent of capital MMBtu	Total MWh savings	Percent of capital kWh	Total 1000 therm savings	Percent of capital therm	Ave unweighted capital NTGR
MCE	15,550	86%	207	33%	148	89%	0.44
PG&E	2,049,973	11%	75,987	21%	17,907	10%	0.83
SCE	95,640	39%	28,029	39%	0	N/A	0.86
SCG	180,043	30%	0	N/A	1,800	30%	0.75
Statewide	**2,341,206**	**14%**	**104,224**	**26%**	**19,856**	**12%**	**0.83**

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, 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 1) the SEM program was delivered in accordance with Design Guide principles, and 2) customer engagement in the program was high. Table 4-12 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 4-12. Average delivery and engagement scores by PA over the PY2023/2024 evaluation years

Requirements	Component	Average PG&E Score n=20	Average SCE Score n=10	Average SCG Score n=8	Average MCE Score n=3
1 – Kick-off meeting. Introducing SEM to the site, defining the roles at the facility, and laying out a program schedule.	Delivery	3.0	3.0	3.0	2.8
	Engagement	2.9	2.1	2.7	2.3
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	2.8	2.8	2.9	3.0
	Engagement	2.5	2.8	2.6	2.0
3 – Energy Use Mapping. Detailed mapping of all energy consumption at the facility.	Delivery	2.6	2.4	2.5	2.3
	Engagement	2.6	2.1	2.7	2.8
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	2.5	2.2	2.6	2.4
	Engagement	2.6	2.3	2.7	2.6
5 – Action Plan and Support. Identifying and creating action plans to meet SEM requirements including meetings, trainings, and workshops.	Delivery	2.6	2.5	2.6	2.0
	Engagement	2.8	2.1	2.6	2.0
6 – Energy Management Information System (EMIS). Developing a system to manage energy use and model energy savings across the entire facility.	Delivery	2.6	2.4	2.6	1.8
	Engagement	2.6	2.3	2.5	2.1
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	2.7	2.6	2.5	2.0
	Engagement	2.8	2.1	2.2	1.8
Average Delivery		2.7	2.5	2.7	2.3
Average Engagement		2.7	2.2	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 4-13 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 have high TDA indicators and high NTGR results, shown in Table 4-13. The MCE program 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 4-13. PY2023/2024 TDA and SRA NTGR results by site

Site ID	Delivery score	Engagement score	Combined TDA score	Final TDA NTGR score	Capital measure SRA NTG	BRO measure SRA NTG
PG&E_01	0.93	0.93	0.93	1	No Cap	0.80
PG&E_02	0.88	0.88	0.88	1	No Cap	0.91
PG&E_03	0.97	0.99	0.98	1	0.83	1.00
PG&E_04	0.91	0.75	0.83	1	1.00	0.59
PG&E_05	0.98	0.98	0.98	1	N.D.	0.78
PG&E_06	0.93	0.95	0.94	1	N.D.	0.54
PG&E_07	0.97	0.98	0.97	1	N.D.	N.D.

Site ID	Delivery score	Engagement score	Combined TDA score	Final TDA NTGR score	Capital measure SRA NTG	BRO measure SRA NTG
PG&E _08	0.94	0.88	0.91	1	N.D.	0.43
PG&E _09	0.80	0.74	0.77	1	0.48	1.00
PG&E _10	0.86	0.88	0.87	1	No Cap	0.90
PG&E _11	0.97	0.90	0.94	1	0.81	0.97
PG&E _12	0.97	0.86	0.91	1	No Cap	N.D.
PG&E _13	0.93	0.95	0.94	1	1.00	N.D.
PG&E _14	0.93	0.85	0.89	1	N.D.	N.D.
PG&E _15	0.94	0.91	0.93	1	0.58	0.57
PG&E _16	0.94	0.96	0.95	1	0.75	0.79
PG&E _17	0.79	0.70	0.74	1	0.67	N.D.
PG&E _18	0.94	0.98	0.96	1	0.77	0.81
PG&E _19	0.52	0.69	0.61	1	N.D.	0.60
PG&E _20	0.98	0.98	0.98	1	0.83	1.00
PG&E Summary	**0.90**	**0.89**	**0.90**		0.77	0.78
SCE_01	0.98	0.95	0.97	1	0.64	0.9
SCE_02	0.94	0.98	0.96	1	0.93	0.97
SCE_03	0.99	0.83	0.91	1	No Cap	0.73
SCE_04	0.94	0.88	0.91	1	1.00	0.97
SCE_05	0.92	0.74	0.83	1	No Cap	0.53
SCE_06	0.88	0.75	0.82	1	0.75	1.0
SCE_07	0.79	0.76	0.78	1	No Cap	N.D.
SCE_SCG_01[1]	0.90	0.93	0.91	1	No Cap	1.0
SCE_08	0.78	0.70	0.74	1	No Cap	0.51
SCE_09	0.99	0.93	0.96	1	0.19	N.D.
SCE Summary	**0.91**	**0.85**	**0.88**	**1.00**	**0.70**	**0.83**
SCG_01	0.84	0.88	0.86	1	0.90	1.0
SCG_02	0.91	0.79	0.85	1	1.00	0.61
SCG_03	0.86	0.83	0.84	1	0.77	N.D.
SCE_SCG_01[1]	0.90	0.93	0.91	1	No Cap	0.83
SCG_04	0.99	1.00	0.99	1	0.87	N.D.
SCG_05	0.97	0.96	0.97	1	N.D.	0.48
SCG_06	0.71	0.71	0.71	1	No Cap	0.62
SCG_07	0.75	0.55	0.65	1	0.00	N.D.
SCG Summary	**0.87**	**0.83**	**0.85**	**1**	**0.70**	**0.71**
MCE_01[2]	0.65	0.74	0.70	1	0.17	N.D.
MCE_02	0.60	0.71	0.66	1	0.72	0.70
MCE_03	0.79	0.74	0.76	1	N.D.	N.D.
MCE Summary	**0.68**	**0.73**	**0.71**	**1.00**	**0.45**	**0.70**

1 No gross evaluated savings due to missing data.
2 Commercial site
N.D. = No data. The respondent completed the TDA portion of the interview but lacked knowledge of specific capital or BRO measures
No Cap = The site did not have any capital measures

Conversations with SEM participants who reported having a successful program experience identified the Energy Treasure Hunt as the most beneficial and influential activity of the program, with attendees noting that it discovered previously unrecognized energy-savings opportunities. Energy Coaches similarly cited the Treasure Hunt to be the most important activity in the SEM program, allowing the implementer to build relationships with facility staff and providing insights into site-specific energy needs.

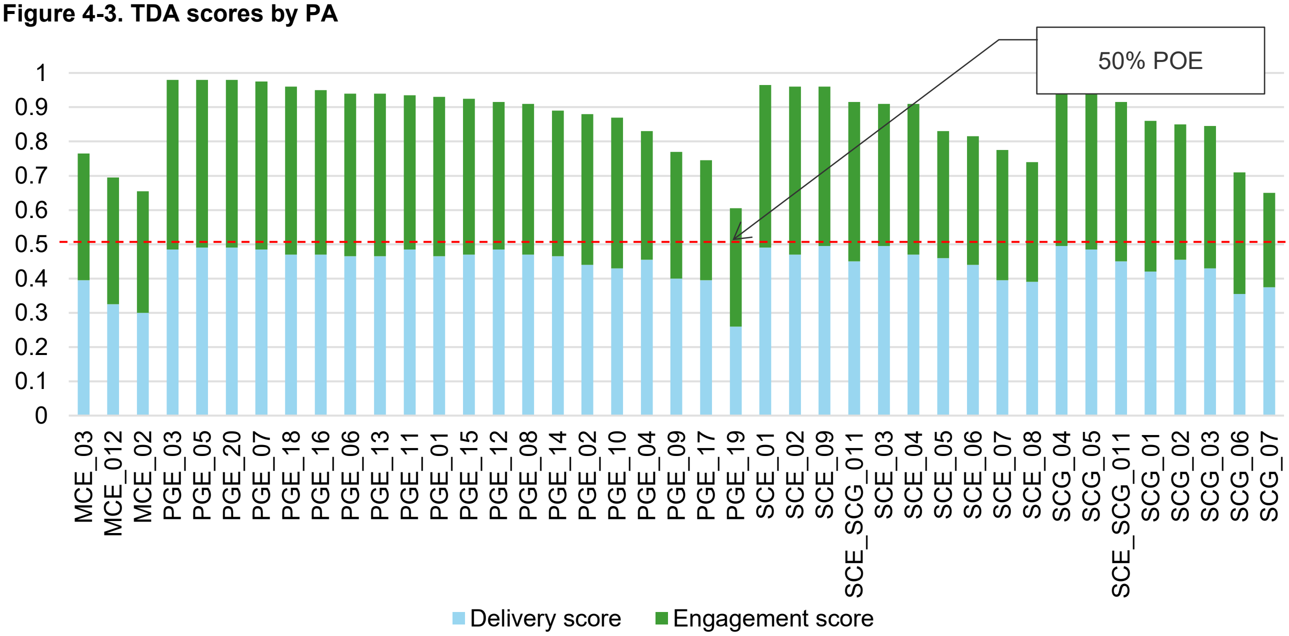
In contrast, the Energy Champions reporting less-than-successful experiences with the SEM program voiced disappointment with the Treasure Hunt, claiming the implementer did not have the expertise necessary to evaluate and diagnose their specific industrial equipment, and therefore could not properly identify opportunities to reduce energy waste. These and other findings will be expanded in the SEM Activity Report, targeting a July 2026 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, very few action plans were 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 PY2023/2024 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 4-3 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 rationale 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 4-3. TDA scores by PA



In contrast to the generally laudatory comments from SCE/SCG participants in 2023, 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.

Furthermore:

- PY2024 PG&E participants noted overall satisfaction with the SEM program and the impact it had on energy savings at their facility.
- Due to high turnover at many facilities, we could not reach the original decision makers regarding SEM projects, and therefore could not score certain TDA questions or generate NTG results for capital or BRO projects.

5 CONCLUSIONS AND RECOMMENDATIONS

This section summarizes all findings from the SEM impact evaluation study and highlights the implications of the findings and presents corresponding recommendations.

5.1 Analysis methodology

In PY2024, SEM participants continued to use a mix of top-down, bottom-up, and mixed methods to estimate savings. Like PY2023, the program shows a more distributed use of methods than in PY2021/2022, when top-down was the dominant approach. In PY2024, approximately one-third of projects used top-down, another third used bottom-up, and the remaining third used a mixed approach, meaning participants applied both methods at different points in the cycle, either across participation years or across fuel types.

As in PY2023, many participants faced data readiness gaps that limited their ability to sustain a valid top-down model across all fuels and reporting periods. Key drivers required for top-down modeling, including reliable production proxies, validated on-site generation data, and stable operational variables, were often available only intermittently. When these drivers were present, participants were generally able to use top-down; when they were not, participants reverted to the bottom-up approach.

The updated PY2024 bottom-up/mixed justification summary indicates that most decisions to use bottom-up or mixed approaches were driven by data limitations rather than discretionary choice. Common reasons included substantial changes in production or energy consumption, statistical failures in attempted models, and incomplete production or energy data. In several cases, sites attempted a top-down model before switching approaches. Additionally, more projects submitted the attempted model as part of their justification packages than in previous cycles. Table 5-1 provides a summary of the savings analysis methodologies used in PY2024.

Table 5-1. Breakdown of savings calculation methodologies

Parameter	Top-down	Bottom-up	Mixed analysis (top-down & bottom-up)
Reviewed projects, n = 40	12	15	13
Percentage by count	30%	38%	33%
Percentage of electric savings	41%	41%	18%
Percentage of gas savings	17%	60%	24%
Percentage of overall MMBtu savings	21%	56%	22%

The SEM M&V guide requires sites using bottom-up calculations to submit a Notification of Bottom-Up (NOBU), which documents the rationale for applying a bottom-up approach in lieu of a top-down method for calculating SEM savings. Table 5-2 summarizes the findings and conclusions from this review.

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

Justification for bottom-up/mixed analysis	Attempted model	Provided model	Quantity
Change in production/energy consumption	10[3]	9[3]	13[1]
Statistics	7[3]	4[3]	9[1]
No production data	6	2	7
No energy data	1	1	1
Installed projects resulted only in gas and not electric savings	0	0	1[2]

[1] Two bottom-up sites provided different justifications for using bottom-up calculations in each of its two reporting periods.
[2] One bottom-up site that achieved natural gas savings provided justifications for using bottom-up calculations for calculating electric energy savings but did not report any electric energy savings.
[3] One bottom-up site for which the top-down model was provided in the project files presented different justifications for using bottom-up calculations in each of its two reporting periods.

As presented in Table 5-2, top-down models for approximately 52% of the projects that used bottom-up calculations were included in the project files.

A key observation in PY2024 is that evaluated savings were less than the claimed savings, and a significant portion of this difference was driven by bottom-up EPIAs. The variability in bottom-up measure complexity, the inconsistent availability of supporting data, and the volume of measure-level calculations at some facilities contributed to these discrepancies. Bottom-up calculations require detailed and consistent documentation across sites to maintain accuracy and comparability—conditions that were not always met. In contrast, top-down approaches follow a standardized set of modeling steps, rely on whole-facility data, and tend to be easier for both implementers and site energy teams to understand and maintain. Table 5-3 presents the evaluation of share of top-down projects in the total evaluated projects.

Table 5-3. Historical share of analysis methodologies (by number of sites)

Parameter	PY2021	PY2022	PY2023[1]	PY2024[1]
Top-down	55%	38%	29%	30%
Bottom-up	27%	46%	43%	38%
Mixed (both top-down and bottom-up)	18%	15%	29%	33%

[1] Based on the review of sampled projects.

As presented in Table 5-3, compared to PY2021/2022, when nearly half of all analyses used top-down, the continued lower share in PY2024 reinforces the need to address these data-readiness issues earlier in the SEM cycle. The increased presence of mixed approaches reflects circumstances where top-down could not be sustained rather than any methodological preference. This further highlights the importance of establishing and maintaining the conditions required for successful top-down modeling and of clearly documenting why alternative methods were used when top-down was not feasible.

Implications

- The SEM M&V Guide designates top-down[50] modeling as the preferred method for calculating SEM savings because it is better suited to capturing the full impact of SEM Behavioral, Retrocommissioning, and Operational (BRO) measures.
- In PY2024, approximately one-third of projects used top-down to estimate SEM savings, one-third used bottom-up, and one-third used a mixed approach, meaning participants applied both methods at different points across cycle years and fuel types. As in PY2023, limitations in data availability constrained many participants' ability to sustain valid top-down models across all fuels and cycle years, particularly the intermittent availability of production proxies and on-site generation data. Participants used top-down modeling when these drivers were available and reverted to bottom-up when they were not.
- Compared to PY2021/2022, PY2023/2024 showed reduced use of top-down methods and increased reliance on bottom-calculations and mixed approaches, reflecting ongoing data availability challenges in sustaining top-down modeling.
- In PY2023/2024, more participants attempted top-down modeling and, in several cases, submitted the attempted models. This indicates that participants generally prioritize top-down and only turn to bottom-up when the top-down model cannot meet statistical or operational requirements. However, the fact that many top-down attempts still fail suggests that underlying data-readiness issues remain unresolved.

[50] 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.

Implications

- The information and data provided in the 16 abandoned top-down models that were available in the project files corroborated the details the participants provided to justify use of the bottom-up approach.
- Uncalibrated, missing, or unreliable on-site generation data remains one of the most common barriers preventing a statistically valid top-down model, particularly for gas-side modeling.
- As in prior years, bottom-up calculations showed greater variability in complexity, data quality, and documentation. Because bottom-up relies on numerous measure-level assumptions, inconsistencies can accumulate, leading to overstated savings or reduced comparability across sites.
- The ongoing decline in fully valid top-down models suggests that either (1) many participants lack the infrastructure or data governance needed for top-down, or (2) readiness assessments at enrollment are not adequately identifying these gaps. Without early intervention, the program will continue drifting away from its preferred methodology of top-down modeling.

Recommendations 1-3

- **Prioritize building the data infrastructure that supports top-down modeling**: Implementers should prioritize building the data infrastructure that supports top-down modeling consistently across fuels and reporting years to reduce reliance on case-by-case bottom-up exceptions.
- **Reassess top-down readiness at regular intervals**: Implementers should assess top-down readiness at enrollment and revisit it mid-year and ensure the availability of stable production proxies and valid operational data during the baseline period to avoid modeling challenges at mid-year.
- **Document steps to restore top-down methodology**: Implementers should document steps to restore top-down modeling when they use bottom-up calculations due to data unavailability and include this information in the bottom-up justifications. This will reduce data limitations that require use of bottom-up calculation.[51] To identify progress of recommendations 1-3, future evaluations will measure percentage of top-down to bottom-up modeling used.

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 Modeling adjustments

Our top-down model corrections in PY2023/2024 resulted in a 1.6% decrease in total forecasted savings. The DNV team reviewed all top-down models that SEM implementers used to calculate savings for projects implemented in PY2024. Several models required corrections to improve statistical significance, better reflect typical operations, and enhance the accuracy of savings calculations. We made these improvements through site-specific adjustments, including:

- Attempting to regenerate the implementer's models to verify their validity and then reviewing the selected variables to assess their relationship with facility energy use.
- Excluding non-routine events (NREs) and periods of unusually low or high energy use when implementers did not provide supporting documentation and facility staff could not explain the cause.

[51] The DNV team identified this recommendation in the PY2021/2022 SEM Evaluation; however, the issue has persisted, with the share of bottom-up projects unable to use top-down modeling approach due to data issues increasing from 15% in PY2021/2022 to 31% in PY2023/24.
DNV, *Strategic Energy Management (SEM) 2021-2022 Impact Evaluation, Page 8*. May 25, 2024. PY2021/2022 SEM Impact Evaluation

- Requesting trend data for variables that better correlated with energy consumption and excluding savings when neither valid trend data nor substantiated bottom-up calculations were available.
- For some models, statistical performance was poor due to weak correlations between selected variables and energy use, and alternative variables with similar performance yielded no savings. Although bottom-up calculations were reviewed, available data and documentation were not sufficiently rigorous to substantiate savings. As neither the top-down models nor the bottom-up analyses provided defensible evidence of savings, these models could not be validated and evaluated savings were set to zero.
- In several cases, we updated the weather variables by using NOAA[52] data and reconstructing the weather inputs from the raw observations. These updates created some differences in the top-down savings results. For most sites, the impact was small. However, a few sites showed larger differences between evaluated and claimed savings. In those cases, the project documentation indicated the use of NASA weather data or did not specify the source of the weather data.

Recommendations 4-6

- **Apply consistent screening and documentation of anomalous data points**: Implementers should apply a consistent screening method to flag baseline and reporting-period data points outside established thresholds and document whether flagged points reflect normal operations or a non-routine event. Clear documentation of flagged points will preserve institutional knowledge and support transparent and accurate evaluation in future years. Future evaluations will track consistent flagging and documentation of anomalous data points as normal or NREs.
- **Review and validate anomalously high or low savings results**: Implementers should review the savings contribution of each reporting-period data point and investigate periods that produce unusually high or low savings, confirm the cause of anomalies, exclude non-routine events from savings calculations, and document retained data points to ensure reported savings reflect SEM-driven operational improvements rather than external factors. Future evaluations will track progress using the same metric described under Recommendation 4 above.
- **Ensure top-down model validity or fully adopt bottom-up calculations when necessary**: Implementers should validate the top-down model by confirming variables' relevance and statistical significance. If implementers are unable to establish a valid model, they should fully transition to bottom-up calculations and avoid combining incomplete elements of both approaches as this undermines the validity and defensibility of savings estimates. Future evaluations will measure how often implementers rely on incomplete elements of both top-down and bottom-up approaches to estimate savings.

5.2.2 Bottom-up calculation discrepancies

Our corrections to PY2023/2024 bottom-up projects resulted in a 2.6% reduction in total forecasted savings. The DNV team reviewed all submitted calculations for the sampled projects and made corrections to methodologies, inputs, and parameters based on detailed engineering review and as-built information provided by facility personnel. Where available, we relied on extended trend data to better reflect typical annual operating conditions rather than short-term snapshots. These updates ensured that savings estimates were consistent with verified equipment performance and actual facility operations. Key adjustments included:

- For production-based EPIAs,[53] we recalculated key parameters using full-year data when available or the longest and most representative trend period otherwise, replacing values derived from limited datasets.

---

[52] NOAA (National Oceanic and Atmospheric Administration) is the U.S. federal agency that operates the national network of ground-based weather stations and provides quality-controlled, verified weather observations widely used for engineering analysis and energy modeling. Their website is https://www.noaa.gov.

[53] Energy Performance Improvement Action (EPIA) – A documented operational or capital action that improves facility energy performance and includes required savings and cost information for CPUC reporting.

- For compressed air EPIAs, we updated system efficiencies (CFM/kW) using verified CAGI[54] performance data, revised as built discharge pressures for pressure reduction measures, and corrected savings where operating conditions differed from implementer assumptions.
- For boiler-related EPIAs, we updated boiler efficiency and loading using the best available information, including facility logs and extended operational data, to reflect typical annual operation.
- For weather-dependent EPIAs, we consistently applied the National Oceanic and Atmospheric Administration (NOAA) station-based weather data and reprocessed weather inputs to ensure accuracy and consistency.

Recommendations 7-9

- **Use representative operational data to support accurate parameter derivation**: Implementers should derive production-, equipment-, and site-specific parameters from representative operational periods, preferably annual trend data when available, or otherwise the longest period that reasonably reflects typical operations, as short or atypical periods can bias savings estimates.
- **Verify all equipment performance and operating parameters using substantiated sources**: Implementers should verify all equipment performance and operating parameters using substantiated sources such as manufacturer documentation, facility-recorded operational data, and trend logs, and ensure that savings inputs reflect actual as-built operating conditions rather than assumed values. This practice will improve the accuracy and consistency of savings calculations. Progress for recommendations 7-8 will be assessed by recording the frequency and relative magnitude of savings corrections due to unsubstantiated operating assumptions and bottom-up savings corrections attributable to non-representative operational datasets.
- **Adopt a weather-data source and use it across all projects**: With National Oceanic and Atmospheric Administration weather data no longer available after August 2025, the program should identify and adopt an alternative weather-data source and use it across all projects, to ensure accurate and comparable weather-dependent savings calculations across the program. PAs should utilize statewide forums to identify a credible and consistent alternative weather-data source, with coordination with CPUC staff.

5.2.3 Savings proration for bottom-up EPIAs

The SEM M&V Guide[55] requires that implementers prorate the annualized savings for bottom-up EPIAs based on the documented implementation date, with reasonable consideration for seasonality and simplicity. This guidance directs implementers to attribute only the portion of savings realized within the current reporting period and to carry forward the remaining savings into subsequent periods.

During this evaluation cycle, we observed that implementers have begun applying the proration requirement; however, the application of proration was often unclear or inconsistent. Because they often treat proration as a straightforward step, sites apply it differently, and new members of the implementation teams may struggle to determine the correct carry-forward values. Several projects did not document the proration method used, and in some cases, implementers carried forward an entire year of savings rather than only the remaining portion after proration. These issues indicate that SEM participants would benefit from a standardized method for calculating and documenting prorated savings.

[54] The Compressed Air and Gas Institute (CAGI) is the industry association that provides standardized performance data for compressed-air equipment. CAGI performance sheets give verified manufacturer data such as CFM per kW, which is used to benchmark compressor efficiency. Their website is https://www.cagi.org/

[55] Sergio Dias Consulting. “California Industrial SEM M&V Guide, Version 3.02.” Section 8.2. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

Recommendation 10

- **Develop a standardized proration tool:** Implementers should develop a simple, standardized proration tool for bottom-up EPIAs to ensure consistent calculation of prorated savings and clear identification of the savings to carry forward to the next reporting period. Future evaluations will assess whether implementers develop a standardized proration tool and the frequency and magnitude of proration corrections.

5.3 Participant engagement and evaluation process improvement

During the initial recruitment phase of the evaluation, participants showed relatively strong engagement in evaluation interviews and early data collection, reflecting the broader reach of the SEM program and participants' familiarity with the program. However, early interviews indicated that some energy champions did not have full awareness of their facility's SEM activities, which limited the completeness of gross data collection and NTGR administration. As evaluation progressed and we requested additional data to verify bottom-up savings, participant responsiveness declined significantly. These challenges became more pronounced at facilities that experienced staff turnover.

Implications

- We observed a significant improvement in participant response rates and willingness to join the introductory evaluation interviews and initial data collection. However, response rates dropped sharply once we requested trend data, parameters, and inputs that implementers originally used to calculate savings for bottom-up measures.
- Participants relying too heavily on implementation partners limits the educational purpose of the SEM program and reduces the long-term value of past participation at these facilities.
- When facilities lack clear handover practices or do not store information on previous SEM activities and implemented measures, it becomes harder to support evaluation and verification in future years.
- When energy teams do not take ownership of implemented measures or the trend data and parameters used in savings calculations, they struggle to fulfill evaluation data requests, which in turn creates risks for meeting evaluation timelines.

Recommendations 11-12

- **Support continuity during energy champion transitions**: Implementers should ensure new energy champions can quickly become familiar with prior SEM activities and development history by directing them to existing project documentation, ongoing SEM Activity Logs, and holding a brief orientation meeting. Implementers should document this orientation either as a standalone summary or within the reporting period performance report. This approach maintains continuity at facilities when staff turnover occurs. Future evaluations will verify through the gross data collection interviews with new energy champions whether implementers provided adequate transition support.
- **Emphasize participant ownership of SEM data and documentation**: Implementers should emphasize participant ownership of trend and activity logs, production data, and operational parameters during SEM workshops and coaching to ensure timely support for evaluation data requests. Future evaluations will assess progress through gross savings data collection interviews by verifying energy champions' knowledge of SEM data sources and reduced reliance on implementers, and document it in the CCT. In addition, future evaluations will track the adoption rate of designated SEM Data Leads, consistent with recommendations in the SEM Design Guide.[56]

[56] Sergio Dias Consulting. "California SEM Program Design Guide Core SEM and Graduate Pathway, Version 2.1." Section 1.2.5 September 1, 2025. [California SEM Design Guide: Core SEM and Graduate Pathway]

5.4 NTGR methods and results

The DNV team refined the SEM SRA instrument developed in the PY2021/2022 evaluation and adapted the 2018–2019 TDA instruments for the PY2023 evaluation based on the approved SEM Design Guide rather than an evaluator-developed Program Theory and Logic Model and used the same instrument for PY2024 with additional questions to support the Activity Report. This change does not omit the importance of the program logic models which should continue to be maintained as part of Implementation Plan (IP) submissions and are generally required for IP development.

The Expansion Study[57] linked program success to program delivery that follows the SEM Guides. The Combined TDA Score measures how well the program complies with the letter and spirit of the Guide. Projects that meet a minimum standard of compliance for SEM delivery (currently set at 50%) warrant the SEM NTGR. Compliance and competence standards 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 13-14

- **Use the current NTG survey instruments in future SEM evaluations**: Evaluators should use the TDA and SRA instruments we developed in this impact evaluation to ensure consistency with the findings and recommendations of the Expansion Study. Future evaluations should document the rationale for using alternative NTG survey instruments, if applicable, and obtain concurrence from CPUC staff prior to implementation.
- **Reassess the 50% minimum compliance threshold for SEM NTGR assignment**: Evaluators should revisit the current 50% minimum compliance standard and determine if it remains an appropriate threshold for assigning a site the SEM NTGR (currently 1.0). The 50% threshold may overrepresent meaningful compliance with SEM Guide requirements and require a potential update. Future evaluations should document the basis for retaining or revising the current 50% minimum compliance threshold for SEM NTGR assignment.

[57] 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

APPENDIX A. NTGR APPROACH, ALGORITHM, AND FACTOR ANALYSIS

This appendix provides a more in-depth description of the NTGR method, including additional context and details of the survey methods. It also examines the potential impact of changes to NTGR algorithms in the sensitivity analysis conducted when the instruments were developed in the PY2021/2022 evaluation.

The PY2023/2024 net-to-gross (NTG) evaluation of the Strategic Energy Management (SEM) program built on the research from the PY2021/2022 evaluation, continuing the assumption held by D16-08-019 stating that a well-designed SEM program's holistic and long-term approach encourages the implementation of behavioral, retro-commissioning, and operational (BRO) measures as well as custom and capital measures. The decision concluded that all measures implemented through the SEM program that demonstrated SEM influence could apply the SEM default NGTR of 1.00. For the PY2021/2022 evaluation, DNV conducted NTG research for capital projects implemented through the SEM program using a self-reported attribution (SRA) approach. The DNV team interviewed Energy Champions and Executive Sponsors from firms participating in the program regarding their decisions to conduct capital upgrades. For this PY2023/2024 evaluation, we repeated the SRA effort for capital projects and conducted additional theory-driven attribution (TDA) research, introduced in the 2018–2019 Study, for non-capital or BRO measures. The PY2024 evaluation marks the completion of a comprehensive two-year evaluation covering both PY 2023 and 2024. The DNV team applied the same SRA and TDA methodologies in the PY2024 evaluation.

A.1 SRA for capital projects

Self-reported attribution (SRA) relies on direct feedback from program participants on two key metrics: the influence the program had on the installation of the capital project, and the intention the customer had to install the capital project if the SEM program had not existed. We interviewed Energy Champions to gather feedback on these two metrics using interview instruments and methods approved in the PY2021/2022 evaluation, with the improvements developed under an NTG improvement study.

The interview instrument and algorithms for scoring the survey responses follow California's standard non-residential 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 in the Group D NTG Methodological Update.

A.1.1 Capital projects interview guide

The capital project section of the interview instrument used in this study was developed in the previous PY2021/2022 Study and was adapted from the Commercial, Industrial, and Agriculture Custom (CIAC) NTG survey battery. 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. CPUC Staff and the DNV team streamlined this battery of questions as part of the Group D NTG Methodological Update. Table A-1 lists the program and non-program factors influencing decisions to install capital projects that remain in the revised battery. In this battery, the responder is asked how influential each of the factors were in making the decision to install the measure on a scale of 1–10.

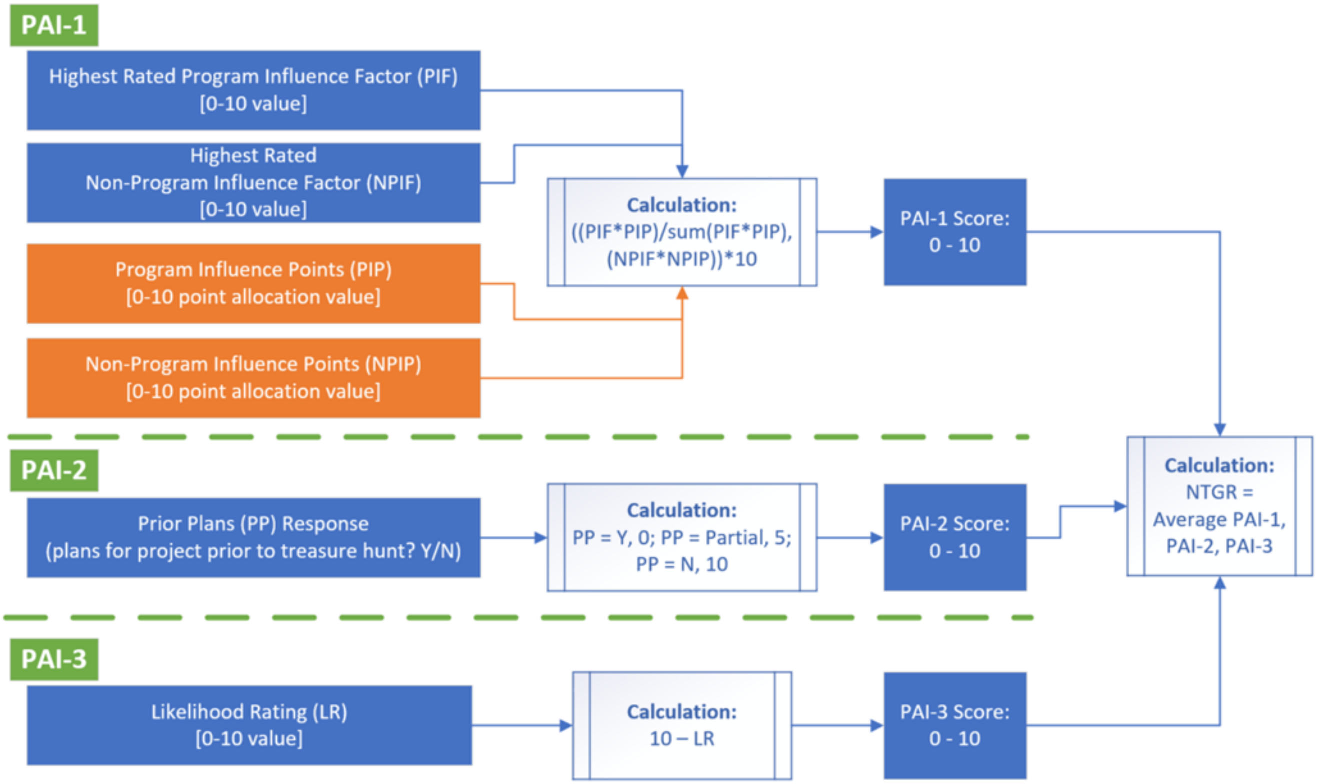
Table A-1. Program and non-program factors

Factors Program
Information provided through the Energy Treasure Hunt
Availability of the milestone and/or performance incentives
Training or information provided in Program workshops
Information provided in meetings with peers from cohorts or SEM participants
Technical support and information provided by the Energy Coach
The development of the Energy Management Information System
Recommendations from the Energy Team
Support from the Executive Sponsor
Non-Program
If the vendor who sold you the equipment made the recommendation to install it, how important was that recommendation on your decision?
Compliance with state or federal regulations, such as Title 24 building codes or Title 20 equipment codes, etc. In other words, was this specific operational or behavioral change incorporated due to these regulations?
Any other factors not discussed that were influential in your decision to install this equipment? Note, other factors provided by participants did not relate to the program and were therefore categorized as Non-Program Factors.

A.1.2 Capital projects algorithm

The DNV team updated the free-ridership algorithm in the PY2021/2022 Study, which CPUC Staff reviewed and approved. We used the same algorithm in the PY2023 Study as shown in Figure A-1.

Figure A-1. NTG scoring algorithm



The updated algorithm places more emphasis on the question, "How many points would you allocate to the program factors, and how many points would you allocate to the non-program factors?" and weights the influence score by that response. It is our opinion that this response provides more insight into the customer's decision-making process than simply applying the highest program and non-program factor influence scores.

Scoring sensitivity. PAI-2 addresses the reported "plans a facility had for a given project, prior to participating in SEM," referred to in the algorithm as the Prior Plan Response. This is a critical element in the free ridership calculation to show the true impact of SEM on capital and BRO projects. This concept is directly tied to the SEM definition of measures eligible to be claimed in the SEM M&V Guide, which specifies that an eligible measure excludes any measure that was identified and planned for implementation outside any SEM program.[58] Examples of evidence for planning include but are not limited to a budget having been allocated, purchase orders having been issued, or a project manager having been assigned. The PAI-2 metric was adapted for the SEM program by referencing the Treasure Hunt and asking Energy Champions a) whether the Treasure Hunt identified the project or opportunity, b) whether they knew about the project or opportunity prior the SEM Treasure Hunt identifying it, c) if they did know about the project or opportunity, if they had plans to install or adopt it, and finally d) if they did report prior plans, if they had a budget set aside for the project or opportunity. The quantified scoring algorithm for PAI-2 is illustrated in Table A-2.

[58] Sergio Dias Consulting. "California Industrial SEM M&V Guide, Version 3.02." Section 8.2.2. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

Table A-2. PAI-2 scoring outcomes

Question*	Response n=9**	Action	Score
Did SEM staff identify the project or opportunity during the treasure hunt?	Yes (5)	Ask follow-up	NA
	No (3)	Ask follow-up if necessary	No Plan = 10
	Don't Know (1)	Ask follow-up	NA
Were you aware of the project or opportunity before the SEM staff identified it during the treasure hunt?	Yes (3)	Ask Prior Plan follow-up	NA
	No, (6)	Ask follow-up if necessary	No Plan = 10
	Don't Know	Ask follow-up if necessary	No Plan = 10
Before SEM staff identified this opportunity, did you have plans in place to install or incorporate the project?	Yes (1)	Ask Budget follow-up	NA
	No (4)	Ask Budget follow-up	NA
	Don't Know	Ask Budget follow-up	NA
Before SEM staff identified this opportunity, did you have budget set aside for the project?	Yes (0)	End section	**Prior Plan = 0**
	No (1)	End section	**Partial Plan = 5**
	Don't Know	End section	No Plan = 10

* We did not always follow the skip logic for each of the prior plan questions. It was obvious during some of the conversations with Energy Champions that there was no prior plan to conduct the capital project without the SEM program. For these sites, we asked the first two questions in the table above, but did not always probe for prior plans to reduce interview burden on the participants.

** Thirteen Energy Champions completed the survey, but only nine had capital projects.

SEM participants who reported being aware of the project or opportunity prior to SEM and had plans to incorporate the project but lacked a budget, were assigned a partial plan score of 5. Participants who had both a plan and a budget set aside for the project were assigned a score of 0. All other participants with a prior plan were assigned a score of 10 for PAI-2. This also reflects the logic of the SEM M&V Guide regarding measure eligibility.

In this population, only three responders indicated that they had been aware of the measure before the Treasure Hunt. Two of these Energy Champions reported no plans to install the measure prior to participating in SEM. For the third site, a partial plan score of 5 was assigned because, although they intended to install the measure at some point, they had not budgeted any money for the work.

CPUC staff requested that the DNV team consider alternate scoring of the PAI-2 metric and the sensitivity of the outcome to changes. The following table examines the impact of changes to the algorithm scores that interprets "planned" measure differently. The DNV team is presenting alternate methods without necessarily advocating for them.

- Current score method: Taking a lead from the SEM Guide, a participating business may be aware of a measure and have had interest in the measure, but if it has taken no concrete steps to implement it, the response is scored 5 points, since intentions often fall short of actions.
- Alternate method 1: This method is more stringent; any evidence of prior planning should be scored as 0 points.
- Alternate method 2: This method is the most stringent, where any awareness of the measure prior to SEM engagement results in a score of zero.

The impact of re-scoring the response is illustrated in Table A-3. The number of responses falling into each category is included in parentheses.

Table A-3. PAI-2 alternate scoring outcomes

Question	Response (Count, n= 9)	Current score	Alternate method 1 More Stringent	Alternate method 2 Most Stringent
Did SEM staff identify the project or opportunity during the treasure hunt?	Yes (5)	NA	NA	NA
	No (3)	No Plan = 10	No Plan = 10	No Plan = 10
	Don’t Know (1)	NA	No Plan = 10	No Plan = 10
Were you aware of the project or opportunity before the SEM staff identified it during the treasure hunt?	Yes (3)	NA	NA	Aware = 0
	No, (6)	No Plan = 10	No Plan = 10	Not aware= 10
	Don’t Know	No Plan = 10	No Plan = 10	Not aware = 10
Before SEM staff identified this opportunity, did you have plans in place to install or incorporate the project?	Yes (1)	**NA**	**Prior Plans = 0**	
	No (4)	NA	No plans = 10	
	Don’t Know	NA	No plans = 10	
Before SEM staff identified this opportunity, did you have budget set aside for the project?	Yes	**Prior Plan = 0**		
	No (1)	Partial Plan = 5		
	Don’t Know	No Plan = 10		
Average weighted capital NTGR Score		92%	92%	89%

A.2 TDA for non-capital projects

The DNV team employed a theory-driven attribution (TDA) approach to evaluate the overall design and delivery of the SEM program and the incorporation of non-capital measures. The PY2018–2019 evaluation Study established TDA as the most appropriate methodology for determining whether the SEM programs were exhibiting program influence broadly. The original TDA methodology reviewed the Program Theory and Logic Model (PTLM)—which we built 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. That 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 were therefore able to assign the NTGR ratio of 1.0 to the SEM program.

The PY2023/2024 evaluation adapted the methodologies from the 2018–2019 Study with findings from the “Group D Strategic Energy Management Expansion Study” published May 23, 2024. The PY2018/2019 Study used the standard SEM PTLM, agreed upon and adopted by all program administrators, and compared the experience of the SEM participants with the input and outcome linkages defined in the PTLM. The Expansion Study concluded that the SEM Guide was patterned on a program delivery model that was linked to high NTGR in multiple jurisdictions. Thus, with approval from CPUC staff, the PY2023/2024 Study referred directly to the current version of the SEM Design Guide rather than a PTLM.

The TDA research compared the participant experience to the program requirements defined in the Guide. The DNV team began with the site-specific activity requirements defined in the SEM Design Guide and categorized them into seven Design Guide requirements, as shown in Table A-4.

Table A-4. SEM Design Guide requirements

Cycle 1	Guidebook site-specific activity	TDA category
Phase 1	Kick-off Meeting, Year 1	**Requirement 1 - Kick Off Meeting**

	Energy Management System Assessment (EMA), Year 1	**Requirement 2 - Energy Management Assessment (EMA)** (Multiyear)
	Energy Map/SEU Selection Support	**Requirement 3 - Energy Use Mapping**
	Treasure Hunt, Year 1	**Requirement 4 - Treasure Hunt** (Multiyear)
	Action Plan- Support	**Requirement 5 - Action Planning and Meetings** (Multiyear)
Phase 2	Planning for Year 2- Support	
Phase 3	OPTIONAL: Treasure Hunt, Year 2	(Included in Requirement 4)
	EMIS Support, Year 2	**Requirement 6 - Energy Management Information System (EMIS)**
Phase 4	EMA, Year 2	(Included in Requirement 2)
	Cycle 1 Completion and Next Steps, Year 2- Support	(Included in Requirement 5)
Cycle 2		
Phase 5	Kick-off Meeting, Year 3	**Requirement 7 - Future Year Planning** (Multiyear) *unless specified the Requirements 1-6 (such as the EMA and Treasure Hunt)*
	Energy Management System Assessment (EMA), Year 3	
	Planning Support, Year 3	
	OPTIONAL: Treasure Hunt, Year 3	
	Operational Control Support	
Phase 6	Employee Engagement Support, Year 3	
	Planning Support, Year 4	
Phase 7	Optional Treasure Hunt, Year 4	
	OPTIONAL: EMIS Support, Year 4	
	Employee Engagement Support, Year 4	
Phase 8	EMA, Year 4	
	Cycle 2 Completion and Next Steps Support	
Cycle 3		
Phase 5	Kick-off Meeting, Year 5	**Requirement 7 - Future Year Planning** (Multiyear) *unless specified the Requirements 1-6 (such as the EMA and Treasure Hunt)*
	Energy Management System Assessment (EMA), Year 3	
	Planning Support, Year 5	
	OPTIONAL: Treasure Hunt, Year 5	
	Operational Control Support	
Phase 6	Employee Engagement Support, Year 5	
	Planning Support, Year 6	
Phase 7	Optional Treasure Hunt, Year 6	
	OPTIONAL: EMIS Support, Year 6	
	Employee Engagement Support, Year 6	
Phase 8	EMA, Year 6	
	Cycle 3 Completion and Next Steps Support	

We gathered feedback from SEM stakeholders for each of the Design Guide requirements to ensure the SEM program was being delivered and received accordingly.

A.2.1 Data gathering

The DNV team conducted in-depth interviews with various stakeholders involved in the SEM program. These included two employees of the participating facility—the Energy Champion and the Executive Sponsor—and one employee of the implementation contractor, the Energy Coach, who is responsible for facilitating the facility through the program. These stakeholders each have unique perspectives on the delivery of the program at a specific site which is a key principle of the TDA approach. Since the responses are considered in concert, any response bias of the program vendor, such as the Energy Coach, is mitigated by customer reports. Note that program vendors are not asked about individual measure influences that weigh into the direct NTGR calculations, but instead, are asked factual questions about aspects of delivery and their assessments of the level of customer engagement. Table A-5 shows the number of completes we attempted to achieve and the actual number we completed.

Table A-5. Data collection summary

Stakeholder	Target interview completes	Actual interview completes	Disposition
Energy Champion **(1 to 1.5-hour interview)**	15 PY2023	13	Of the 15 possible interviews, one Energy Champion declined the interview, and another did not respond to the repeated email attempts and outreach. The team included CPUC staff, PAs, and implementer staff in the email correspondences to try to boost participation.
	36 PY2024	29	Of the more than 40 sites contacted, approximately half did not respond despite repeated outreach attempts. Sites that responded promptly were typically led by highly engaged and active energy champions or executive sponsors. Initial contact information provided to the evaluation team was inaccurate, including outdated listings for staff who had left the site or were no longer serving in energy champion roles. Midway through the evaluation cycle, the evaluation team received updated contact information identifying energy team members who were serving in informal energy champion roles. Throughout the evaluation process, the evaluation team copied CPUC staff on outreach emails and coordinated with implementers to help encourage site participation.
Executive Sponsor **(1 hour interview)**	5 PY2023	6	The team conducted 6 in-depth interviews, but one interview was at a site which did not include a completed Energy Champion interview. Therefore, the team did not use the feedback from this Executive Sponsor in the quantitative analysis.
	15 PY2024	8	Several executive sponsors effectively served as energy champions and brought a broader, site-level perspective. All executive sponsors interviewed demonstrated strong familiarity with the program and its impacts on their site. Executive sponsors not interviewed did not respond to multiple requests from evaluation team.
Energy Coach **(1 hour interview)**	10 sites PY2023	14 sites	Energy Coaches handle multiple participating SEM sites. The team was given contact information for 9 Energy Coaches and was able to contact all but one of them. The 8 completed interviews covered 14 sites participating in the SEM program.
	15 sites PY2024	11 sites	5 energy coach interviews were completed, representing 10 sites. Energy coaches were highly engaged, and all but one served as coaches for sites included in the evaluation. This remaining coach was listed erroneously as a coach in the data provided.
Energy Team Member **(10-minute survey)**	20 PY2023 Only	1	The team sent an online survey link to all Energy Team member contacts provided by the PAs. However, 7 sites did not provide team member contacts, and 7 of the remaining, 8 sites provided old or incorrect contact information due to the high employee turnover at these types of industrial facilities.

The in-depth interview guides were designed to determine if there was evidence that the SEM program was being implemented as intended and how well the SEM program was being received at each facility. We asked a battery of questions to gather feedback for each of the seven SEM Design Guide requirements shown in Table A-4. For each Design Guide requirement, we rated the strength of evidence on a 1–3 scale (strong evidence = 3, moderate evidence = 2, and weak evidence = 1) on 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 deliver or sell the SEM program to gain buy-in from each facility. We then averaged the 1–3 strength scores for delivery and engagement to arrive at the requirement score for each site, then divided the requirement score by 21 (the maximum possible score of the seven Design Guide requirements at a score of 3 each) to calculate the site's final TDA score (Table A-6).

Reminder: The final TDA score is not equivalent to the attribution score derived from the capital SRA research. Instead, it measures how well the delivery complies with the requirements and intent of the SEM Design Guide, since the Expansion Study[59] established a strong correlation with high NTGR for programs that followed the delivery specified in the SEM guides.

A.2.2 TDA results and site stories

The TDA methodology allowed us to gather detailed feedback from multiple sources about the experiences each participant had with the SEM program. We created a "story" for each site that helped support and justify the final TDA score. Table A-6 shows the TDA score and site story for each site.

Table A-6. Summary of site TDA story

Site and provider	Final TDA Score	Site story
Site A - PY2023	0.96	The site was excited to participate in the SEM program and was very engaged in the program activities through the first couple of program years. Staff resources were slim, and the Energy Champion did not have a large Energy Team to help with SEM projects or attend meetings and trainings. The site relied on the SENSEI software tool to manage projects, track energy use, and model energy savings but did not create a documented plan for backfilling SEM roles or pursuing energy efficiency projects after the program participation period ended. This lack of plan and engaged Energy Team became an issue in year 3 of the program, when the Energy Champion was pulled away from the site to open a new plant and no one was available to continue to champion the SEM program.
Site B - PY2023	0.91	The site had an employee who looked for energy efficiency opportunities and pursued them whenever possible, similar to an Energy Champion, prior to the SEM program. However, this Energy Champion reported learning a lot from the SEM trainings on what to look for and how to find deeper ways to reduce energy waste at the site. The dedicated facility staff was small due to other work duties, but the Treasure Hunt engaged the entire Energy Team and Executive Sponsor, and the SENSEI software gave them a great tool for managing projects and tracking energy use. The site did not create a plan to backfill SEM roles and when the Energy Champion retired, there was no one at the facility available to take his place. This Energy Champion has actually been called in to handle SEM activities and continue the SEM projects, but there needs to be a better plan for backfilling roles and continuing SEM savings after the program ends.
Site C - PY2023	0.97	The site was very engaged in the SEM program and was very pleased with the energy saving opportunities identified during the treasure hunt. The Energy Champion reported that the immersive nature of the program was hard on the site and the regular check-in meetings with the Energy Coach were key to keeping the site staff on task for completing the SEM activities. The SENSEI tool was also instrumental in quantifying the impact of the SEM projects and helped get additional facility staff to buy-in to the program requirements.
Site D - PY2023	0.84	The site was engaged with the program for the first couple of program years and conducted multiple treasure hunts to identify deeper opportunities to save energy. The facility staff available to help the Energy Champion was limited and the plans for backfilling SEM roles did not exist, so when the original Energy Champion changed

[59] 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

Site and provider	Final TDA Score	Site story
		roles, there was no backup staff member ready to take up the challenge of SEM and engage with the program properly. This led to a significant drop in engagement in the last year of the program and a lack of persistence after the program ended.
Site E - PY2023	0.66	The site was aware of their energy use prior to the SEM program but was very interested in learning more about energy efficiency through SEM participation. The Energy Champion expressed disappointment, however, when the program staff were unable to identify any deep retrofit opportunities and only recommended upgrades and maintenance to steam traps. The Energy Coach tells a slightly different story and reported that the program staff were only allowed to visit during off-season times and were unable to see the plant under full operation. The Energy Champion was also disappointed with the Excel-based gas model since the gas interval data was monthly and not granular enough to be of value. The site attended most of the trainings and workshops in the first year of the program, but decided it was not worth their time and dropped out of the program entirely after year 2.
Site F - PY2023	0.83	The site was interested in SEM but was skeptical that the program could really find enough energy efficiency opportunities to benefit their facility. The implementer was able to use the SENSEI software tool to quantify one of the large BRO projects identified through the program, which increased buy-in from facility staff to support and drive the SEM program forward. Staff attended the training workshops and were educated about SEM, but the site did not document a plan for teaching new staff on the importance of continuing to pursue the SEM activities when staff turnover occurred. When the original Energy Champion changed roles, the new Energy Champion was not as familiar with the program requirements and engagement dropped in years 3 and 4.
Site G - PY2023	0.82	The site's Energy Champion was excited about the program and very interested in ways for the facility to reduce energy waste and save money. There was not a lot of support from other facility staff due to work priorities and a relatively small staff size, so the Energy Champion was reliant on the SENSEI software tool to help manage projects and other SEM activities. The program found a number of energy efficiency opportunities during the Treasure Hunt and the site completed those with the strongest return on investment. The Energy Champion expressed some disappointment with the technical expertise of the implementation staff who did not know how to evaluate some of the site's major energy-consuming equipment and would have liked the program to bring in some third-party engineers to help identify deeper efficiencies. The program's inability to identify more opportunities beyond the low-hanging fruit caused the site to drop out after only the second program year.
Site H - PY2023	0.78	The site was very interested in pursuing opportunities to reduce energy use and cost since they are a very large energy consumer. They did not have a lot of resources available to devote to the program due primarily to other work priorities. The implementation team identified a couple of BRO projects, which the site completed, but did not have the expertise to fully understand the site's major energy consuming equipment and identify deep efficiencies. The site expressed disappointment with the programs lack of technical expertise to help a site such as theirs and decided to drop out after two years.
Site I - PY2023	0.86	The site found success through the SEM program but did not have enough resources available to build a strong Energy Team. Program engagement was slow at first due to work priorities and strict regulations at the site to maintain proper humidity and temperature levels for their products. The implementer was able to find opportunities to reduce energy waste, including some insulation maintenance that would not have been found for many years had the program Treasure Hunt not been so thorough.

Site and provider	Final TDA Score	Site story
Site J - PY2023	0.85	The site focused on safety and production over energy efficiency before the start of the SEM program, but their participation and engagement in the program showed them the importance of saving energy and the associated cost savings. The SENSEI software tool quantified the impacts of the efficiency projects identified through the treasure hunt and drove higher customer engagement in the second, third, and fourth years of the program. They did not have a lot of staff to devote to the program and did not make solid plans for continuing energy efficiency efforts after the end of the program. The site did, however, participate all four years and had only positive feedback about the program as a whole.
Site K - PY2023	0.91	The site's parent company has a strong focus on reducing energy costs and employes a Continuous Improvement team to promote energy efficiency at all of the companies in its portfolio. The specific site included in this sample has participated in an electric SEM program for four years, and in a gas SEM program for another four years. The site was more impressed by the electric SEM program that included the SENSEI software and granular interval data modeling energy savings. The gas SEM program provided an Excel-based tool that did not provide as much feedback to the customer. The gas SEM program staff brought in heat cameras to look for energy leaks and found significant savings in fixing equipment insulation. The customer was quite pleased with the aspect of the program.
Site L - PY2023	0.91	The site did not have a large staff and could not devote many resources to the SEM program. The limited staff who participated in the kick-off meeting and treasure hunt did not immediately buy-in to the program's ability to save them much energy. The implementer used the SENSEI software tool to quantify the impacts of the projects identified in the treasure hunt and significantly improved customer engagement. The Energy Champion used the SENSEI tool often for project management and tracking energy use and was disappointed when the tool was no longer available after the program period ended. They also did not have a documented plan in place for continuing to pursue energy efficiency opportunities after the program ended.
Site M - PY2023	0.70	The site's first Energy Champion was very engaged with the program and excited about finding and completing energy efficiency opportunities at the facility. However, other facility staff were not as engaged, and when the original Energy Champion left the company, the replacement champion did not engage with the program as much. The new Energy Champion attended some meetings with the Energy Coach but did not pursue most of the SEM activities. The one project the site completed through the program was a well-known project that did was not identified by the program, but rather something the facility had plans to do without SEM. Delivery of the SEM program for this site was poor and failed to engage enough staff to be successful.
Site N - PY2024	0.93	Before SEM, the site had no dedicated energy team—energy management was handled by site leadership while a centralized utilities group managed most systems across facilities; during SEM, leadership mainly provided escalation support and ensured access to data and resources. A strong kick-off identified a motivated Energy Champion who ultimately carried much of the workload and effectively served as Executive Sponsor. The site received annual EMAs (with no apparent improvement from year 1 to year 2) and leveraged existing building management, power monitoring, and data-collection systems; treasure hunts were highly engaging and surfaced new opportunities. Through the program, the site updated its energy policy and action plan, met with the coach biweekly, and attended 6–7 workshops (no cohort meetings). While no new EMIS was created, SEM strengthened energy management by enhancing an existing sustainability tracking tool and improving how the team set baselines, monitored performance, and refined assumptions before incentive submissions. The site now has multiple champions and strong upper-management involvement and is very likely to continue energy-reduction efforts; participants viewed SEM as clearly value-added because requests were purposeful and helped uncover non-obvious operational insights (e.g., linking washer performance to the water distillation system). The site joined after a PA

Site and provider	Final TDA Score	Site story
		representative introduced SEM as a good fit to support carbon goals and capture incentives, and the energy coach noted the right people were involved on calls but recommended broader participation from additional staff.
Site O - PY2024	0.88	This isn't new for us—we've pursued energy projects for 20+ years and CPUC incentives for 10+ years, staying highly engaged and consistently chasing incentives, which remain the main reason we participate; SEM and its vendors primarily help us support the work and track performance. We received EMAs each year, with scores improving from year 1 to year 2, and we use SEM tools and data because it's a better way to drive our goals. Treasure hunts were rated highly even though only the maintenance supervisor participated, and our involvement included one workshop plus monthly meetings with the energy coach. We did not develop an EMIS, but incentives have motivated the champion and team to "do better energy-wise," and while there's no formal transition plan, the maintenance supervisor is the likely backup if the champion turns over—ultimately, the focus is saving both money and energy.
Site P - PY2024	0.98	The Energy Champion role has rotated, but the current champion has been involved for five years and provides strong continuity; overall, the team is highly engaged, proactive, and says SEM has been eye-opening. The site received EMAs annually, though the evaluation team did not have access to the files. SEM energy models and dashboards are central to how the site operates—shops now track energy intensity per vehicle and attribute kWh/MMBtu to specific areas and processes, making it easier to explain baseline loads and non-zero energy use during low production to finance. Engagement has ramped up significantly after a slower first two years, especially in years 3–4, delivering substantial savings; the team has completed intensive multi-day treasure hunts (8–10 hours/day for three days), attended 6+ workshops, and meets with the coach at least four times per month. While the site did not develop a formal action plan, it built shop-level teams and champions, expanded from facility systems into production areas after early compressed-air success, and holds regular check-ins to prioritize opportunities and address barriers—though securing buy-in and capital for larger projects remains the main challenge. A successor champion has been identified, and the Process section could not be completed due to time constraints.
Site Q - PY2024	0.76	The organization manages 100+ multifamily properties with a small, non-engineering team, so SEM's free 1-on-1 coaching, rebate support, and local market expertise were especially valuable—particularly since they often rely on paid consultants. The site received annual EMAs and improved from year 1 to year 2, but overall use of SEM tools/EMIS was limited (an Excel tracker was used mainly for incentive calculations and high-level snapshots), in part because multifamily incentives are misaligned—landlords have few levers, limited capex, and little financial upside to improve in-unit efficiency. Engagement included one half-day treasure hunt (no follow-ups), monthly then quarterly coach meetings, and an action plan, but no workshops; the executive sponsor instead requested more scalable education (e.g., online modules). The program was discovered via email/utility bill and a PA intro meeting, and the site joined for the free audit/coaching with minimal downside. Energy coach feedback was very positive, highlighting the site's engagement and the coach's thoughtful, above-and-beyond education tailored to financial realities.
Site R - PY2024	0.83	The site held a kick-off meeting, but the current Energy Champion—who inherited the role 1.5 years ago—seemed only moderately familiar with SEM program details and was unsure how the site first learned about the program. The site received EMAs annually, though scores declined from year 3 to year 4, and while the Champion received a baseline energy map/model, they were unaware of any related data collection efforts. After an initial multi-hour treasure hunt attended by just two staff, additional treasure hunts were completed and opportunities were tracked via a task list. The team does not participate in cohort meetings, did not develop an action plan, meets with the energy coach monthly, and has not developed or actively used

Site and provider	Final TDA Score	Site story
		an EMIS; no transition plan or broader staff involvement was noted. The energy coach's feedback was positive, stating the site "did a great job."
Site S - PY2024	0.98	The site has 3+ staff closely involved in SEM and shows strong engagement, receiving annual EMAs with scores improving from year 3 to year 4. The team relies heavily on a baseline energy model, with the Energy Champion sharing daily electricity-per-production data that's embedded in routine operations. They conduct two treasure hunts per year (harvest and off-season) with 3–4 attendees, rated 8/10, and attend at least one workshop per quarter, valuing peer learning and coach check-ins for accountability. Using in-house decision science support, they built a predictive energy model to aid peak-demand shaving during harvest; while no EMIS was developed, they created an IDSM tool and a separate GHG tool (outside SEM). Incentives reinforced energy-waste reduction and delivered a strong ROI, and the team emphasized they operate more like an industrial facility than a boutique winery—benefiting from SEM staff with relevant industrial experience. Participation was driven by early success at a pilot site and the value of technical support (site visits, modeling, and linking production to energy use), with key decision factors being direct cost/savings for BRO leak repairs (rating 7) and a blended mix of aging assets, audits/regulatory drivers, and incentives for capital projects (rating 4). Energy coach feedback was very positive, noting that in-person treasure hunts deepened site familiarity and that modeling and financial justification support—plus in-person incentive check delivery—helped drive behavior change.
Site T - PY2024	0.94	The site has two highly engaged Energy Champions and a strong engineering team, received annual EMAs with scores improving from year 3 to year 4, and found SEM's savings calculations and energy-use models very useful—daily electricity-per-production tracking is now embedded in routine operations. A half-day treasure hunt drew 4–5 staff, with additional hunts in later years, and the site built its first action plan through the program, meets regularly with the coach, and attends workshops annually. While no formal EMIS was developed, the site relies on boiler SCADA for daily tracking and is nearing completion of a solar microgrid with real-time electric monitoring. Incentives totaling hundreds of thousands of dollars helped cover labor and improved understanding of plant performance and next steps. The program was introduced by executive sponsor S.A., a long-tenured mechanical/energy leader, primarily to address rising energy costs while advancing efficiency, learning, sustainability, and carbon-reduction goals. The energy coach noted the team knew the plant extremely well (even if less focused on food-process specifics) and saw no downside for the customer.
Site U - PY2024	0.97	The site's early efforts were more reactive than proactive, focusing on basic fixes rather than more innovative opportunities, but SEM helped keep the team aligned on targets and secure needed resources despite recent staff turnover. The site received annual EMAs with scores improving from year 1 to year 2, and—because operations are highly seasonal—used the off-season for preventive maintenance while verifying savings during peak season. Treasure hunts drew strong participation (15–20 staff), the team attended local workshops when possible, joined peer meetings, and maintained frequent touchpoints (weekly champion–coach meetings and biweekly executive sponsor check-ins) while developing and reviewing an action plan. With SEM support, the site upgraded its EMIS and related systems to be more reliable and user-friendly, improving data quality, opportunity tracking, and monitoring. The Energy Champion transitioned into the role with strong support and has a transition plan in place; the program is viewed as valuable for ongoing utility improvements (electricity, steam, water, controls) and for finding cost-

Site and provider	Final TDA Score	Site story
		and GHG-saving opportunities. Key drivers were a BRO focus on reducing waste and boosting boiler efficiency via hotter condensate return (rating 8) and a capital focus on better measurement/visibility to uncover inefficiencies and reduce natural gas and GHGs (rating 9). The energy coach noted the team was professional, responsive, and consistently provided the needed information.
Site V - PY2024	0.99	The Energy Champion is highly engaged—having served as an "unofficial" champion before SEM—and is supported by a knowledgeable engineering team. The site received annual EMAs with scores improving from year 1 to year 2 and uses SEM models to support annual environmental/sustainability audits. Treasure hunts were rated highly, with engineers participating and even uncovering previously overlooked safety issues. The team meets with the coach every 2–3 weeks, updated an existing action plan to include additional savings opportunities, and attended at least four cohort-style meetings. Since beginning installations in 2017/2018, the site has continued expanding initially underused capabilities, and SEM connected staff with technical experts when specialized knowledge was needed (e.g., boilers). While EMIS has been discussed, many EMIS-like functions already exist via the site's SCADA system. Participation was owner-supported—reflecting sustainability and brand pressures in the textile industry—with incentives as an added motivator; the energy coach described the team as very knowledgeable, pleasant, and responsive to reinforcement.
Site W - PY2024	0.91	The Energy Champion—initially "voluntold" into the role—remains highly engaged and has built an energy team that uses Fridays (supported by program incentives) to address energy issues; the Champion also serves as Executive Sponsor. The site wasn't sure whether it received EMAs, and scores declined from year 1 to year 2. SEM provided an energy model and supported a data collection plan, but a formal EMIS wasn't adopted because it doesn't fit the business model; instead, the site relies on an existing automation/trending system with minute-level data, graphics, and history, plus daily tracking in therms, kW, and dollars. Treasure hunts drew strong participation (6–7 staff), were rated highly, and continued in later years. The team developed an action plan, meets regularly with the energy coach, attended 4+ workshops (including some cohort sessions), and noted that energy management was not coordinated prior to SEM—leadership focused on big strategic projects but didn't actively manage ongoing operational energy use. Energy coach feedback was very positive: the site "did the work" and "asked all the right questions."
Site X - PY2024	0.97	The site was already highly proactive in energy waste reduction and continuous improvement, led by an engaged Energy Champion and utilities engineer. It received annual EMAs with scores improving from year 1 to year 2, and SEM provided an energy model plus a data collection plan, though the site has not implemented an EMIS (metering/real-time feedback has been discussed but not pursued). Because the site is "cutting edge," many treasure hunt recommendations were either already completed or offered limited incremental value, even though turnout was high. The team remains very engaged—meeting with the coach every other week (rated highly), attending workshops regularly, and occasionally participating in peer meetings despite privacy concerns—while integrating SEM action planning into existing plans and participating in a non-gas SEM program as well. The program added structure and continuity through organized project lists and action plans, supported by a semi-formal succession approach for the champion role. The site joined via a PA electric offering in 2020, saw strong early results, and expanded into natural gas based on continued success and rebate value; the energy coach praised the team for helping uncover overlooked energy impacts and bringing fresh perspectives.

Site and provider	Final TDA Score	Site story
Site Y - PY2024	0.74	Despite recent staff turnover, the current Energy Champion and their supervisor remain highly proactive about energy management—before SEM, the supervisor often “took it upon himself to save energy.” The site received EMAs annually, though scores declined from year 1 to year 2 amid leadership changes. The site received a data collection plan and an energy-use model, but overall engagement is unclear due to limited institutional knowledge and several “don’t know” responses. The initial treasure hunt had strong participation (including finance and maintenance) and was rated highly for being eye-opening, but no additional treasure hunts occurred after key staff left. The current EC did not attend workshops, recalled about six opportunities, and did not remember working with the energy coach (though prior ECs likely did). A successor EC has been identified, and the supervisor is expected to sustain the effort if the EC turns over. The site learned about SEM through the plant manager and joined primarily to reduce energy waste and save money; the energy coach described the team as great and very professional.
Site Z - PY2024	0.77	The site held a kick-off meeting and the Energy Champion was moderately engaged, but as the largest gas plant in the country, they felt the SEM team lacked sufficient gas-specific expertise to deliver high value. The site received EMAs annually with scores improving from year 3 to year 4,yet struggled with missing/backlogged data and keeping the energy model current—making it hard to track savings and secure incentives, even though the model was viewed as “how we got the money.” Treasure hunts were seen as low value and lasted only a few hours, staff attended about one workshop per year, and early coach meetings were frequent but rated poorly due to limited gas expertise and data issues. An EMIS was recommended but rejected by internal IT, and the site reported limited overall program value with no transition plan. Energy coach feedback echoed this, noting one strong contributor on a CO project but otherwise weak gas capability; the site joined after hearing positive results from peer facilities.
Site AA - PY2024	0.87	Energy Champion focused on maintaining a strong project pipeline, securing resources, and coordinating on measurement, validation, and verification through valuable monthly check-ins that sustained accountability and momentum. The site is extremely engaged: EMAs were completed annually (though files weren’t available to evaluators), a baseline map was provided but the facility’s size drives project-by-project tracking, and 20+ staff participated in a 2–3 day treasure hunt. The program added operational rigor through frequent coach meetings (weekly early on) and many workshops (no cohort sessions), helping the site become a top performer that hit the rebate cap, and it improved existing energy data into a more usable EMIS for decision-making. However, the team is reconsidering continued participation because without incentives the time and resource demands of analysis and verification make future projects less likely; incentives were especially critical for non-capital actions and enabled BRO leak-fix bundles that were highly feasible, operator-friendly, and strongly program-influenced (10/10). The program was introduced by the PA via site leadership, and the site has multiple engaged energy leads and sponsors, prompting them to recommend SEM to sister facilities.
Site AB - PY2024	0.94	SEM drove a clear cultural shift at the site—employees now continuously look for and report potential energy waste (even unusual sounds), reinforcing the idea that ignoring issues like air leaks has real opportunity cost. The site has highly engaged Energy Champions, a formal Executive Sponsor plus a former EC who remains involved, and a clear transition plan supported by knowledgeable staff. EMAs were completed annually, though scores declined from year 3 to year 4. SEM helped develop and refine the site energy model and expand reporting, building on pre-existing tracking and volt/energy mapping efforts (the site has EMIS-like capabilities but doesn’t label them as such). The site completed about four formal treasure hunts, but “treasure hunting” is now embedded in daily operations; workshops were less aligned (e.g., ammonia/refrigeration). Overall, the program delivered hundreds of thousands of dollars in value through improved uptime, preventive maintenance, tighter air/gas controls, and more standardized operations, with incentives a major participation and capital-project driver. Energy coach

Site and provider	Final TDA Score	Site story
		feedback was positive, citing strong technical support and broader implementer participation during treasure hunts.
Site AC - PY2024	0.91	The site held a kick-off meeting, though it's unclear whether executive leadership participated. Before SEM, they already had an energy management team and internal tools that allowed employees to submit and rank energy-saving ideas for resourcing. The site received EMAs annually with scores staying roughly flat from year 4 to year 5, and the site plus the company "transformation office" use SEM outputs to help prioritize projects—though they saw limited added value in treasure hunts and noted it's getting harder to find new opportunities. The site had a pre-existing (non-energy-specific) action plan and used SEM to build a more targeted project list, rating regular coach check-ins highly for help with savings calculations. They declined an EMIS due to post-acquisition resource constraints and concerns about sustaining IT/technical support, citing past internal tools that degraded over time and lost trust. The current Energy Champion inherited the role after a reorg; rebates remain a major driver because they help projects compete for scarce capital, and leadership prioritizes energy given its cost and sustainability impact in the gas industry. Overall, the site viewed SEM as extremely value-added—especially early training/education and manageable time demands—and the coach noted the team is easy to work with and helped keep the program burden low.
Site AD - PY2024	0.71	The Energy Champion is highly engaged, and the site already had a strong efficiency/sustainability foundation—previously using contractors for EE/LEED upgrades and maintaining an internal "energy team." The site tracks energy performance in the Gazebo/SENSEI tool, but it was unclear whether EMAs were received and scores appeared flat from year 1 to year 2. With 40+ buildings and some restricted areas, the treasure hunt felt less rigorous than paid audits and required significant site effort to uncover deeper opportunities. The site entered SEM without an action plan (focus had been water reduction) and reported limited value from coaching and site visits, with most technical analysis and M&V burden falling on the facility; EMIS was discussed but no implementation support was provided. While the site joined after learning about SEM through the utility—primarily to advance 2030 net-zero and emissions goals—dissatisfaction with technical support suggests a medium-to-low likelihood of continuing after the program.
Site AE - PY2024	0.94	The Energy Champion has led the program for seven years and remains highly engaged; the site completed annual EMAs with scores improving from year 1 to year 2 and received an energy map/model plus a data collection plan. Engagement has been strong, including a half-day treasure hunt (with a one-hour classroom session) attended by 4–5 staff, a formal action plan, regular coach meetings, cohort-style sessions, and several workshops per year, though no EMIS was developed. The organization is broadly committed, with two sites enrolled. The program was introduced by long-tenured executive sponsor and participation was driven by rising energy costs alongside efficiency, sustainability, learning, and carbon-reduction goals. The energy coach described the team as extremely prepared and professional—highly recommendable—while noting limited food-processing expertise was a minor challenge but also brought useful "fresh eyes."

Site and provider	Final TDA Score	Site story
Site AF - PY2024	0.96	The Energy Champion has a strong sustainability background, but executive sponsor support is limited, with SEM efforts often funded through the general plant budget. The site received annual EMAs with scores improving from year 1 to year 2, and while the Champion wants more granular energy mapping, limited metering and capital—prioritized instead for R&D and manufacturing cycle-time improvements—make that difficult; the Champion therefore relies on implementer reports to brief leadership. Treasure hunts were well attended (8–9 employees) and rated highly, and overall engagement has been strong, though IDSM was discussed but not pursued given metering constraints. There is no formal transition plan, but the Champion is highly qualified and the organization has additional sustainability reporting support and a broader sustainability crew at another facility. The site joined through the PA, viewing SEM as a way to cut costs and build staff engagement by training maintenance teams to spot day-to-day energy-saving opportunities.
Site AG - PY2024	0.65	Staff turnover (three departures) and competing priorities have constrained participation, leaving a newly assigned Energy Champion—who didn't attend the original kick-off—to manage the program with a small team (EC, site director, and one energy-focused staff member). The site was unsure whether it received EMAs and saw no clear score change from year 1 to year 2. Because data must be pulled manually, reporting became too time-intensive, and monthly coach meetings often showed little progress; many identified opportunities were also projects the site was already planning to complete anyway. As a result, the site is temporarily "graduating"/pausing SEM to focus on more urgent operational needs, planning to reevaluate in late 2026 and likely reengage in 2027. A key frustration was unclear program value—especially where incentives went and whether savings were being realized—though the implementer was described as persistent and supportive, and the energy coach noted the lack of progress was due to site constraints rather than engagement from the SEM team.
Site AH - PY2024	0.89	The team attended the kick-off and was already highly engaged and "cutting edge" in energy management, participating in both the electric and gas SEM tracks after joining through a PA offering in 2020. The site received annual EMAs with scores improving from year 1 to year 2 and remains strongly engaged—especially through highly rated check-ins with the energy coach—even though it prefers privacy and rarely joins peer meetings. Treasure hunts provide some value, but often surface ideas the team has already considered or views as too low-impact to prioritize. The site merged its existing action plan with the SEM plan and has discussed EMIS and IDSM but hasn't committed yet, largely because fuller EMIS capability depends on additional field metering that is still in progress. While there is no formal transition plan, the site was pursuing energy-waste reduction before SEM, and the energy coach praised the team for helping uncover overlooked energy impacts and bringing fresh perspectives.
Site AI - PY2024	0.93	The site has two Energy Champions and additional staff actively involved in SEM. It received annual EMAs, with scores roughly flat from year 3 to year 4, and overall engagement is strong. The initial treasure hunt was rated highly but was limited in scope (about 3 hours with 2–3 attendees) and was not repeated. The team attended cohort-style meetings and multiple workshops, which "planted ideas" and supported learning, and they rated biweekly coach check-ins highly—especially for confirming whether projects could qualify for incentives. An EMIS and IDSM were developed through SEM, and while the Champion still has questions, they praised the vendor's support. There is no formal transition plan, but participation remains active; the Champion joined the role by assignment to learn new practices and identify gas/electric savings in energy-intensive garlic and onion dehydration, improving costs and the bottom line. The energy coach described the site as a good partner, collaborating every two weeks to vet and add projects while providing practical feedback.

Site and provider	Final TDA Score	Site story
Site AJ - PY2024	0.95	The Energy Champion reports strong plant manager and upper-management involvement, with annual EMAs showing scores improving from year 3 to year 4. The Champion finds the energy maps, models, and graphs highly valuable, and the site rates treasure hunts extremely highly—dedicating a half day with about six participants from leadership and the energy team. Staff did not attend workshops or cohort meetings but met regularly with the energy coach. EMIS use is unclear, though the Champion noted the site would not have achieved its progress without the IDSM component. The site has successfully navigated at least one champion transition, has an engaged energy team that continually looks for savings, and the energy coach praised the site's professionalism and follow-through, noting that regular reminders and walk-throughs helped projects and smaller upgrades happen faster and more consistently.
Site AK - PY2024	0.74	The maintenance team effectively serves as the energy team and actively champions the program, with a supervisor helping surface opportunities, but leadership turnover has left the site without a current executive sponsor. The site was unsure whether it received EMAs, and changes in EMA tools from year 1 to year 2 make comparisons difficult. The team met once to review the energy model but rarely uses it, focusing instead on completing repairs and recognizing savings without deeper analysis; no EMIS was developed. Treasure hunts were short (about one hour) but rated very highly, and additional hunts were conducted in later years. While staff could not attend workshops or cohort meetings, they maintained frequent contact with the coach (often by text) and found the interactions productive—though technical support was limited to higher-level recommendations. The current lead inherited the role, is unsure how the program began, and joined because reimbursed costs and abundant utility-side opportunities made participation an easy decision; coach feedback was positive, noting strong responsiveness and communication.
Site AL - PY2024	0.96	The site completed a kick-off and has a highly engaged Energy Champion who works closely with the maintenance manager and has recommended similar programs to sister facilities in other states. The site received annual EMAs with scores improving from year 3 to year 4, reviews the energy model at least quarterly (often every 6 weeks) to account for operational changes, and conducts annual treasure hunts with about 10–12 staff. The team attends cohort meetings—valued for shared lessons and common challenges—meets regularly with the energy coach, and developed an action plan through the program. An EMIS was discussed but not pursued because the existing system is outdated and would require a costly overhaul, though the site has implemented IDSM. While there's no formal transition plan, multiple staff could step into the champion role; key challenges include corporate budget constraints and preventing over-reliance on the high-performing champion. The champion inherited SEM after the site joined in 2020, citing incentives, guidance, and sustainability alignment as key benefits, and the coach praised the program as an effective extension of the team—flagging only occasional delays in model updates due to stretched resources.
Site AM - PY2024	0.61	No kick-off was completed, but the new Energy Champion is highly engaged and eager to use SEM services, working with the plant manager, corporate staff, and an active Executive Sponsor to sustain involvement despite high turnover. The site was unsure about EMAs, and scores declined from year 3 to year 4. A site map and data collection plan were created, but the Champion doesn't know why the energy model hasn't been provided. Due to turnover and a rough transition, the Champion was unaware of many early program activities, hasn't participated in treasure hunts, and has only met the energy coach once, with no clear action plan or workshop participation. The site has no EMIS or IDSM yet and no formal transition plan, though one is likely to be developed. The Champion entered the program as the boiler supervisor to address leaks and inefficiencies, driven mainly by cost reduction and rebates; the energy coach noted no concerns.
Site AN - PY2024	0.98	Both Energy Champions are highly engaged and knowledgeable—one led sustainability efforts before SEM—and the engineering team is actively involved.

Site and provider	Final TDA Score	Site story
		The site received EMAs annually (though files weren't available to the evaluation team) and uses an energy model that tracks both electric and gas. Treasure hunts are a major activity: the site dedicates a full day, with at least four engineers attending the first hunt and additional staff joining later. While SEM did not result in a new or updated action plan, the team has attended several workshops (no cohort-style participation noted). EMIS development is underway with the implementer to align with ISO 5001 (about 85% complete), but the current champion is still ramping up, so understanding and internal buy-in are still developing. The site learned about SEM through a former employee who connected facilities engineering, corporate sustainability, and the PA, and incentives remain the primary driver—supported by ongoing weekly coordination with the PA. The energy coach noted the team can "walk the walk and talk the talk" across the site's systems.

A.3 Further TDA NTGR research

The SEM programs apply a composite NTGR of a capital measure NTGR using SRA methods and a non-capital measure NTGR using TDA methods. There are two areas of potential weaknesses in current methods:

- The current TDA threshold score for application of the SEM NTGR was established using a preponderance of evidence standard (of 50%). A better alignment with the findings of the "Group D Strategic Energy Management Expansion Study"[60] would establish a threshold based on minimum compliance to SEM Guides.
- In those cases where a program fails to qualify for the SEM NTGR, there is no alternate NTGR protocol.

A.3.1 TDA compliance threshold

The original TDA methods of the PY2018/2019 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. The threshold is applied twice, once when assessing whether each *site* met the evidence standard and a second time when assessing whether the *program* met the evidence standard.

As noted in Section A.2 the current evaluation cycle does not exactly replicate the PY2018/2019 Study, although it builds upon its framework using subsequent research. 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. Thus, the goal of the TDA research is to measure how well the program delivery reflects the SEM Guides. This raises the question of whether a program that scores 0.50 adequately reflects the SEM guides and warrants an NTGR of 1.0, since conceptually, it is only 50% compliant with the SEM delivery. Other disciplines have established minimum compliance or competencies that suggest a higher threshold is warranted. For example, the California BAR Exam has a 70% threshold for passing.

Understanding which threshold to use to determine whether the delivered programs, as measured by the TDA research, sufficiently reflect the SEM Guides remains a topic of debate and would require input from stakeholders across the California SEM landscape. For informational purposes, Table A-7 shows the number of sites in the current PY2023/2024 evaluation that met various thresholds of success and indicates that 95% of projects would exceed a 70% threshold for passing if that were selected for .

[60] 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

Table A-7. TDA qualifying threshold

Theoretical "passing" threshold:	0.5	0.6	0.7	0.75	0.8
2023 sites that would pass each threshold (n=11)	11	11	10	9	8
Percentage of sites passing each threshold	100%	100%	91%	82%	73%
2024 sites that would pass each threshold (n=27)	27	27	25	22	20
Percentage of sites passing each threshold	100%	100%	93%	81%	74%
2023/2024 sites that would pass each threshold (n=38)	37	37	35	31	28
Percentage of sites passing each threshold	100%	100%	95%	84%	76%

A.3.2 Alternate NTGR protocols

If a program fails to meet the TDA threshold, the program no longer qualifies for the SEM NTGR, which is currently 1.0. With failure, there is no protocol or precedent for assigning an alternate NTGR. The combined TDA scores presented in Table 4-13 are not equivalent to an NTGR.

As part of producing a well-rounded review of the SEM program, the DNV team 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.

APPENDIX B. ELECTRIC PROJECT RESULTS

Table B-1 presents electric project-level results, including the project sample weight, forecasted and evaluated first-year savings, and GRR. The sampling weights reflect the number of customers in the population that a sample customer represents for given strata. The sampling weights also incorporate sample and population characteristics not used for explicit stratification.

Table B-1. Combined PY2023/2024 electric project-level results

PA Claim ID	Evaluation Year	Analysis	Sampling weight	First year (kWh)		
				Forecasted	Evaluated	GRR
MC-3535	PY2024	Top-Down	1.67	65,869	64,147	97.39%
MC-194	PY2023	Top-down	1.67	142,018	142,408	100.27%
FPC18.01	PY2024	Mixed	2.00	3,537,596	3,537,596	100.00%
FPC18.08	PY2024	Top-Down	2.00	578,904	575,754	99.46%
FPC18.12	PY2024	Mixed	2.00	235,547	235,263	99.88%
FPC20.07	PY2024	Mixed	1.20	1,473,111	1,499,420	101.79%
FPC20.09	PY2024	Top-Down	2.00	3,104,564	3,104,565	100.00%
FPC20.11	PY2024	Mixed	2.00	5,669,524	5,626,602	99.24%
FPC20.14	PY2024	Top-Down	2.00	7,018,702	6,905,771	98.39%
FPC20.20	PY2024	Mixed	2.00	646,390	646,390	100.00%
FPC22.03	PY2024	Mixed	2.00	138,379	84,890	61.35%
FPC22.06	PY2024	Bottom-Up	2.00	525,753	493,473	93.86%
FPC22.07	PY2024	Bottom-Up	2.00	742,176	706,261	95.16%
FPC22.10	PY2024	Bottom-Up	2.00	175,467	133,101	75.86%
FPC22.13	PY2024	Bottom-Up	2.00	165,190	164,124	99.35%
FPC22.16	PY2024	Mixed	2.00	1,017,118	177,303	17.43%
MFR20.12	PY2024	Bottom-Up	2.00	1,451,602	1,568,007	108.02%
MFR20.13	PY2024	Mixed	2.00	2,088,283	2,080,172	99.61%
MFR20.16	PY2024	Mixed	2.00	3,171,281	3,424,515	107.99%
MFR20.18	PY2024	Bottom-Up	1.20	22,159,381	22,185,575	100.12%
MFR20.19	PY2024	Top-Down	1.20	17,879,344	15,754,352	88.11%
MFR22.24	PY2024	Bottom-Up	1.00	4,208,834	4,208,834	100.00%
7675.P2	PY2024	Bottom-Up	11.83	943,080	669,626	71.00%
7675.P9	PY2024	Top-Down	1.57	2,804,965	3,989,320	142.22%
7802.P3	PY2024	Bottom-Up	1.57	6,959,476	5,957,961	85.61%
EE_CALC_5970052	PY2024	Top-Down	11.83	30,542	43,142	141.26%
EE_CALC_5970115	PY2024	Top-Down	11.83	42	55	130.98%
EE_CALC_5970118	PY2024	Top-Down	11.83	52	66	128.55%
7321.P4	PY2023	Top-down	11.83	3,791	-242,365	-6393.15%
7335.P10	PY2023	Bottom-up	1.57	3,938,501	3,938,331	100.00%
7335.P1	PY2023	Bottom-up	11.83	21,824	13,013	59.63%

PA Claim ID	Evaluation Year	Analysis	Sampling weight	First year (kWh)		
				Forecasted	Evaluated	GRR
7335.P9	PY2023	Mixed	1.57	4,490,796	1,694,513	37.73%
7523.P1	PY2023	Bottom-up	1.57	2,964,794	-701,233	-23.65%
7523.P6	PY2023	Mixed	1.57	2,693,242	2,645,724	98.24%
7523.P8	PY2023	Mixed	1.57	3,178,182	3,210,330	101.01%

APPENDIX C. NATURAL GAS PROJECT RESULTS

Table C-1 presents the natural gas project-level results, including the project sample weight, forecasted and evaluated first-year savings, and GRR. The sampling weights reflect the number of customers in the population that a sample customer represents for given strata. The sampling weights also incorporate sample and population characteristics not used for explicit stratification.

Table C-1. Combined PY2023/2024 natural gas project-level results

PA Claim ID	Evaluation Year	Analysis	Sampling weight	First year (therm)		
				Forecasted	Evaluated	GRR
MC-3535	PY2024	Top-Down	1.67	7,823	7,823	100.00%
MC-120	PY2023	Bottom-up	1.67	9,455	9,455	100.00%
MC-354	PY2023	Bottom-up	1.00	131,126	53,883	41.09%
FPC18.01	PY2024	Mixed	2.00	255,544	176,001	68.87%
FPC22.13	PY2024	Bottom-Up	2.00	238,256	238,256	100.00%
MFR20.18	PY2024	Bottom-Up	1.20	847,033	850,569	100.42%
FPC20.20	PY2024	Mixed	2.00	198,721	202,308	101.80%
FPC18.12	PY2024	Mixed	2.00	172,262	142,913	82.96%
MFR20.13	PY2024	Mixed	2.00	137,550	137,550	100.00%
FPC20.12	PY2024	Mixed	1.20	771,178	771,178	100.00%
FPC20.07	PY2024	Mixed	1.20	732,999	735,819	100.38%
FPC18.08	PY2024	Top-Down	2.00	162,418	163,258	100.52%
FPC20.11	PY2024	Mixed	2.00	100,102	100,042	99.94%
FPC18.11	PY2024	Mixed	2.00	560,426	552,155	98.52%
FPC20.09	PY2024	Top-Down	2.00	90,351	65,714	72.73%
FPC20.17	PY2024	Mixed	2.00	94,401	91,382	96.80%
MFR22.24	PY2024	Bottom-Up	1.00	10,701,587	10,701,587	100.00%
MFR20.16	PY2024	Mixed	2.00	229,570	264,295	115.13%
FPC22.06	PY2024	Bottom-Up	2.00	117,974	110,731	93.86%
MFR20.11	PY2024	Top-Down	1.20	2,093,552	2,093,552	100.00%
FPC22.03	PY2024	Mixed	2.00	403,024	325,548	80.78%
7839.P1	PY2024	Top-Down	1.33	531,748	496,805	93.43%
7802.P2	PY2024	Top-Down	2.00	98,621	77,104	78.18%
7565.P3	PY2024	Bottom-Up	1.33	437,255	152,597	34.90%
7565.P1	PY2024	Mixed	1.33	261,194	217,441	83.25%
7396.P2	PY2023	Top-down	2.00	120,229	124,676	103.70%
7396.P3	PY2023	Mixed	2.00	175,499	92,355	52.62%
7638.P1	PY2023	Bottom-up	2.00	141,694	363,304	256.40%
7638.P6	PY2023	Top-down	2.00	34,193	23,025	67.34%

APPENDIX D. STANDARD HIGH-LEVEL SAVINGS

Table D-1. Gross lifecycle savings (MWh)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	5,090	5,058	0.99	0.0%	0.99
SEM	**MCE**	**Total**	**5,090**	**5,058**	**0.99**	**0.0%**	**0.99**
SEM	PGE	PGE	448,532	432,927	0.97	0.0%	0.97
SEM	**PGE**	**Total**	**448,532**	**432,927**	**0.97**	**0.0%**	**0.97**
SEM	SCE	SCE	298,820	210,795	0.71	0.0%	0.71
SEM	**SCE**	**Total**	**298,820**	**210,795**	**0.71**	**0.0%**	**0.71**
SEM	SCG	SCG	0				
SEM	**SCG**	**Total**	**0**				
SEM		***Statewide***	***752,442***	***648,780***	***0.86***	***0.0%***	***0.86***

Table D-2. Net lifecycle savings (MWh)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	5,090	5,057	1.00	0.0%	1.00	1.00	1.00	1.00
SEM	**MCE**	**Total**	**5,090**	5,057	**1.00**	**0.0%**	**1.00**	**1.00**	**1.00**	**1.00**
SEM	PGE	PGE	448,532	418,777	0.93	0.0%	1.00	0.97	1.00	0.97
SEM	**PGE**	**Total**	**448,532**	**418,777**	**0.93**	**0.0%**	**1.00**	**0.97**	**1.00**	**0.97**
SEM	SCE	SCE	298,820	208,986	0.70	0.0%	1.00	0.99	1.00	0.99
SEM	**SCE**	**Total**	**298,820**	**208,986**	**0.70**	**0.0%**	**1.00**	**0.99**	**1.00**	**0.99**
SEM	SCG	SCG	0							
SEM	**SCG**	**Total**	**0**							
SEM		***Statewide***	***752,442***	***632,820***	***0.83***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

Table D-3. Gross lifecycle savings (MW)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	0.7	0.9	1.26	0.0%	1.26
SEM	**MCE**	**Total**	**0.7**	**0.9**	**1.26**	**0.0%**	**1.26**
SEM	PGE	PGE	62.8	60.7	0.97	0.0%	0.97
SEM	**PGE**	**Total**	**62.8**	**60.7**	**0.97**	**0.0%**	**0.97**
SEM	SCE	SCE	38.1	26.1	0.68	0.0%	0.68
SEM	**SCE**	**Total**	**38.1**	**26.1**	**0.68**	**0.0%**	**0.68**
SEM	SCG	SCG	0.0				
SEM	**SCG**	**Total**	**0.0**				
SEM		***Statewide***	***101.6***	***87.7***	***0.86***	***0.0%***	***0.86***

Table D-4. Net lifecycle savings (MW)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	0.7	0.8	1.00	0.0%	1.00	1.00	1.00	1.00
SEM	**MCE**	**Total**	**0.7**	0.8	**1.00**	**0.0%**	**1.00**	**1.00**	**1.00**	**1.00**
SEM	PGE	PGE	62.8	58.7	0.93	0.0%	1.00	0.97	1.00	0.97
SEM	**PGE**	**Total**	**62.8**	**58.7**	**0.93**	**0.0%**	**1.00**	**0.97**	**1.00**	**0.97**
SEM	SCE	SCE	38.1	25.9	0.68	0.0%	1.00	0.99	1.00	0.99
SEM	**SCE**	**Total**	**38.1**	**25.9**	**0.68**	**0.0%**	**1.00**	**0.99**	**1.00**	**0.99**
SEM	SCG	SCG	0.0							
SEM	**SCG**	**Total**	**0.0**							
SEM		***Statewide***	***101.6***	***85.4***	***0.83***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

Table D-5. Gross lifecycle savings (MTherms)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	860	445	0.52	0.0%	0.52
SEM	**MCE**	**Total**	**860**	**445**	**0.52**	**0.0%**	**0.52**
SEM	PGE	PGE	95,427	93,774	0.98	0.0%	0.98
SEM	**PGE**	**Total**	**95,427**	**93,774**	**0.98**	**0.0%**	**0.98**
SEM	SCE	SCE	-2			0.0%	
SEM	**SCE**	**Total**	**-2**			**0.0%**	
SEM	SCG	SCG	12,079	10,932	0.91	0.0%	0.91
SEM	**SCG**	**Total**	**12,079**	**10,932**	**0.91**	**0.0%**	**0.91**
SEM		***Statewide***	***108,364***	***105,151***	***0.97***	***0.0%***	***0.97***

Table D-6. Net lifecycle savings (MTherms)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	860	307	0.36	0.0%	1.00	0.69	1.00	0.69
SEM	**MCE**	**Total**	**860**	**307**	**0.36**	**0.0%**	**1.00**	**0.69**	**1.00**	**0.69**
SEM	PGE	PGE	95,427	91,751	0.96	0.0%	1.00	0.98	1.00	0.98
SEM	**PGE**	**Total**	**95,427**	**91,751**	**0.96**	**0.0%**	**1.00**	**0.98**	**1.00**	**0.98**
SEM	SCE	SCE	-2			0.0%	1.00		1.00	
SEM	**SCE**	**Total**	**-2**			**0.0%**	**1.00**		**1.00**	
SEM	SCG	SCG	12,079	9,424	0.78	0.0%	1.00	0.86	1.00	0.86
SEM	**SCG**	**Total**	**12,079**	**9,424**	**0.78**	**0.0%**	**1.00**	**0.86**	**1.00**	**0.86**
SEM		***Statewide***	***108,364***	***101,482***	***0.94***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

Table D-7. Gross first year savings (MWh)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	1,018	1,012	0.99	0.0%	0.99
SEM	**MCE**	**Total**	**1,018**	**1,012**	**0.99**	**0.0%**	**0.99**
SEM	PGE	PGE	89,706	86,585	0.97	0.0%	0.97
SEM	**PGE**	**Total**	**89,706**	**86,585**	**0.97**	**0.0%**	**0.97**
SEM	SCE	SCE	59,764	42,159	0.71	0.0%	0.71
SEM	**SCE**	**Total**	**59,764**	**42,159**	**0.71**	**0.0%**	**0.71**
SEM	SCG	SCG	0				
SEM	**SCG**	**Total**	**0**				
SEM		***Statewide***	***150,488***	***129,756***	***0.86***	***0.0%***	***0.86***

Table D-8. Net first year savings (MWh)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	1,018	1,005	0.99	0.0%	1.00	1.00	1.00	1.00
SEM	**MCE**	**Total**	**1,018**	1,005	0.99	**0.0%**	**1.00**	**1.00**	**1.00**	**1.00**
SEM	PGE	PGE	89,706	83,755	0.93	0.0%	1.00	0.97	1.00	0.97
SEM	**PGE**	**Total**	**89,706**	**83,755**	**0.93**	**0.0%**	**1.00**	**0.97**	**1.00**	**0.97**
SEM	SCE	SCE	59,764	41,797	0.70	0.0%	1.00	0.99	1.00	0.99
SEM	**SCE**	**Total**	**59,764**	**41,797**	**0.70**	**0.0%**	**1.00**	**0.99**	**1.00**	**0.99**
SEM	SCG	SCG	0							
SEM	**SCG**	**Total**	**0**							
SEM		***Statewide***	***150,488***	***126,557***	***0.84***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

Table D-9. Gross first year savings (MW)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	0.1	0.2	1.26	0.0%	1.26
SEM	**MCE**	**Total**	**0.1**	**0.2**	**1.26**	**0.0%**	**1.26**
SEM	PGE	PGE	12.6	12.1	0.97	0.0%	0.97
SEM	**PGE**	**Total**	**12.6**	**12.1**	**0.97**	**0.0%**	**0.97**
SEM	SCE	SCE	7.6	5.2	0.68	0.0%	0.68
SEM	**SCE**	**Total**	**7.6**	**5.2**	**0.68**	**0.0%**	**0.68**
SEM	SCG	SCG	0.0				
SEM	**SCG**	**Total**	**0.0**				
SEM		***Statewide***	***20.3***	***17.5***	***0.86***	***0.0%***	***0.86***

Table D-10. Net first year savings (MW)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	0.1			0.0%	1.00		1.00	
SEM	**MCE**	**Total**	**0.1**			**0.0%**	**1.00**		**1.00**	
SEM	PGE	PGE	12.6	11.7	0.93	0.0%	1.00	0.97	1.00	0.97
SEM	**PGE**	**Total**	**12.6**	**11.7**	**0.93**	**0.0%**	**1.00**	**0.97**	**1.00**	**0.97**
SEM	SCE	SCE	7.6	5.2	0.68	0.0%	1.00	0.99	1.00	0.99
SEM	**SCE**	**Total**	**7.6**	**5.2**	**0.68**	**0.0%**	**1.00**	**0.99**	**1.00**	**0.99**
SEM	SCG	SCG	0.0							
SEM	**SCG**	**Total**	**0.0**							
SEM		***Statewide***	***20.3***	***16.9***	***0.83***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

Table D-11. Gross first year savings (MTherms)

Report name	PA	Standard report group	Ex ante gross	Ex post gross	GRR	% Ex ante gross pass through	Eval GRR
SEM	MCE	MCE	172	89	0.52	0.0%	0.52
SEM	**MCE**	**Total**	**172**	**89**	**0.52**	**0.0%**	**0.52**
SEM	PGE	PGE	19,085	18,755	0.98	0.0%	0.98
SEM	**PGE**	**Total**	**19,085**	**18,755**	**0.98**	**0.0%**	**0.98**
SEM	SCE	SCE	0			0.0%	
SEM	**SCE**	**Total**	**0**			**0.0%**	
SEM	SCG	SCG	2,416	2,186	0.91	0.0%	0.91
SEM	**SCG**	**Total**	**2,416**	**2,186**	**0.91**	**0.0%**	**0.91**
SEM		***Statewide***	***21,673***	***21,030***	***0.97***	***0.0%***	***0.97***

Table D-12. Net first year savings (MTherms)

Report name	PA	Standard report group	Ex ante net	Ex post net	NRR	% Ex ante net pass through	Ex ante NTG	Ex post NTG	Eval ex ante NTG	Eval ex post NTG
SEM	MCE	MCE	172	61	0.36	0.0%	1.00	0.69	1.00	0.69
SEM	**MCE**	**Total**	**172**	**61**	**0.36**	**0.0%**	**1.00**	**0.69**	**1.00**	**0.69**
SEM	PGE	PGE	19,085	18,350	0.96	0.0%	1.00	0.98	1.00	0.98
SEM	**PGE**	**Total**	**19,085**	**18,350**	**0.96**	**0.0%**	**1.00**	**0.98**	**1.00**	**0.98**
SEM	SCE	SCE	0			0.0%	1.00		1.00	
SEM	**SCE**	**Total**	**0**			**0.0%**	**1.00**		**1.00**	
SEM	SCG	SCG	2,416	1,885	0.78	0.0%	1.00	0.86	1.00	0.86
SEM	**SCG**	**Total**	**2,416**	**1,885**	**0.78**	**0.0%**	**1.00**	**0.86**	**1.00**	**0.86**
SEM		***Statewide***	***21,673***	***20,296***	***0.94***	***0.0%***	***1.00***	***0.97***	***1.00***	***0.97***

APPENDIX E. STANDARD PER-UNIT SAVINGS

Table E-1. Per unit (quantity) gross energy savings (kWh)

Report name	PA	Standard report group	Pass through	% ER ex ante	% ER ex post	Average EUL (yr)	Ex post lifecycle	Ex post first-year	Ex post annualized
SEM	MCE	MCE	0	0.0%	0.0%	5.0	842,939.9	168,588.0	168,588.0
SEM	PGE	PGE	0	0.0%	0.0%	5.0	10,068,077.3	2,013,615.5	2,013,615.5
SEM	SCE	SCE	0	0.0%	0.0%	5.0	2,570,671.0	514,134.2	514,134.2
SEM	SCG	SCG	0	0.0%	0.0%	5.0			

Table E-2. Per unit (quantity) net energy savings (kWh)

Report name	PA	Standard report group	Pass through	% ER ex ante	% ER ex post	Average EUL (yr)	Ex post lifecycle	Ex post first-year	Ex post annualized
SEM	MCE	MCE	0	0.0%	0.0%	5.0	842,939.9	167,507.9	167,507.9
SEM	PGE	PGE	0	0.0%	0.0%	5.0	9,739,001.3	1,947,800.3	1,947,800.3
SEM	SCE	SCE	0	0.0%	0.0%	5.0	2,548,606.3	509,721.3	509,721.3
SEM	SCG	SCG	0	0.0%	0.0%	5.0			

Table E-3. Per unit (quantity) gross energy savings (Therms)

Report name	PA	Standard report group	Pass through	% ER ex ante	% ER ex post	Average EUL (yr)	Ex post lifecycle	Ex post first-year	Ex post annualized
SEM	MCE	MCE	0	0.0%	0.0%	5.0	74,134.8	14,827.0	14,827.0
SEM	PGE	PGE	0	0.0%	0.0%	5.0	2,180,794.5	436,158.9	436,158.9
SEM	SCE	SCE	0	0.0%	0.0%	5.0			
SEM	SCG	SCG	0	0.0%	0.0%	5.0	780,891.0	156,178.2	156,178.2

Table E-4. Per unit (quantity) net energy savings (Therms)

Report name	PA	Standard report group	Pass through	% ER ex ante	% ER ex post	Average EUL (yr)	Ex post lifecycle	Ex post first-year	Ex post annualized
SEM	MCE	MCE	0	0.0%	0.0%	5.0	51,092.4	10,218.5	10,218.5
SEM	PGE	PGE	0	0.0%	0.0%	5.0	2,133,747.8	426,749.6	426,749.6
SEM	SCE	SCE	0	0.0%	0.0%	5.0			
SEM	SCG	SCG	0	0.0%	0.0%	5.0	673,133.9	134,626.8	134,626.8

APPENDIX F. KEY CONCLUSIONS AND RECOMMENDATIONS

Table F-1 presents the findings of the PY2024 SEM evaluation and the corresponding recommendations. Some recommendations may have been identified in previous SEM evaluations; however, because the underlying issues persist, they are reiterated here and, where appropriate, refined to provide more a targeted direction.

Table F-1. PY2024 key conclusions and recommendations

Item #	Page #	Findings	Comment	Recipient
1-3	43	• The SEM M&V Guide designates top-down modeling[61] as the preferred method for calculating SEM savings because it is better suited to capturing the full impact of SEM Behavioral, Retrocommissioning, and Operational (BRO) measures. • In PY2024, approximately one-third of projects used top-down to estimate SEM savings, one-third used bottom-up, and one-third used a mixed approach, meaning participants applied both methods at different points across cycle years and fuel types. As in PY2023, limitations in data availability constrained many participants' ability to sustain valid top-down models across all fuels and cycle years, particularly the intermittent availability of production proxies and on-site generation data. Participants used top-down modeling when these drivers were available and reverted to bottom-up when they were not. • Compared to PY2021/2022, PY2023/2024 showed reduced use of top-down methods and increased reliance on bottom-calculations and mixed approaches, reflecting ongoing data availability challenges in sustaining top-down modeling. • In PY2023/2024, more participants attempted top-down modeling and, in several cases, submitted the attempted models. This indicates that participants generally	• **Prioritize building the data infrastructure that supports top-down modeling:** Implementers should prioritize building the data infrastructure that supports top-down modeling consistently across fuels and reporting years to reduce reliance on case-by-case bottom-up exceptions. • **Reassess top-down readiness at regular intervals:** Implementers should assess top-down readiness at enrollment and revisit it mid-year and ensure the availability of stable production proxies and valid operational data during the baseline period to avoid modeling challenges at mid-year. • **Document steps to restore top-down methodology:** Implementers should document steps to restore top-down modeling when they use bottom-up calculations due to data unavailability and include this information in the bottom-up justifications. This will reduce data limitations that require use of bottom-up calculation. To identify progress of recommendations 1-3, future evaluations will measure percentage of top-down to bottom-up modeling used.	All PAs

[61] 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.

Item #	Page #	Findings	Comment	Recipient
		prioritize top-down and only turn to bottom-up when the top-down model cannot meet statistical or operational requirements. However, the fact that many top-down attempts still fail suggests that underlying data-readiness issues remain unresolved. • The information and data provided in the 16 abandoned top-down models that were available in the project files corroborated the details the participants provided to justify use of the bottom-up approach. • Uncalibrated, missing, or unreliable on-site generation data remains one of the most common barriers preventing a statistically valid top-down model, particularly for gas-side modeling. • As in prior years, bottom-up calculations showed greater variability in complexity, data quality, and documentation. Because bottom-up relies on numerous measure-level assumptions, inconsistencies can accumulate, leading to overstated savings or reduced comparability across sites. • The ongoing decline in fully valid top-down models suggests that either (1) many participants lack the infrastructure or data governance needed for top-down, or (2) readiness assessments at enrollment are not adequately identifying these gaps. Without early intervention, the program will continue drifting away from its preferred methodology of top-down modeling		

Item #	Page #	Findings	Comment	Recipient
4-6	45	Several top-down models required corrections to improve statistical significance, better reflect typical operations, and enhance the accuracy of savings calculations. We made these improvements through site-specific adjustments, including: • Attempting to regenerate the implementer's models to verify their validity and then reviewing the selected variables to assess their relationship with facility energy use. • Excluding non-routine events (NREs) and periods of unusually low or high energy use when implementers did not provide supporting documentation and facility staff could not explain the cause. • Requesting trend data for variables that better correlated with energy consumption and excluding savings when neither valid trend data nor substantiated bottom-up calculations were available. • For some models, statistical performance was poor due to weak correlations between selected variables and energy use, and alternative variables with similar performance yielded no savings. Although bottom-up calculations were reviewed, available data and documentation were not sufficiently rigorous to substantiate savings. As neither the top-down models nor the bottom-up analyses provided defensible evidence of savings, these models could not be validated and evaluated savings were set to zero. • In several cases, we updated the weather variables by using NOAA[62] data and reconstructing the weather inputs from the raw observations. These updates created some differences in the top-down savings results. For most sites, the impact was small. However, a few sites showed larger differences between evaluated and claimed savings. In those cases, the project documentation indicated the use of NASA weather data or did not specify the source of the weather data.	• **Apply consistent screening and documentation of anomalous data points:** Implementers should apply a consistent screening method to flag baseline and reporting-period data points outside established thresholds and document whether flagged points reflect normal operations or a non-routine event. Clear documentation of flagged points will preserve institutional knowledge and support transparent and accurate evaluation in future years. Future evaluations will track consistent flagging and documentation of anomalous data points as normal or NREs. • **Review and validate anomalously high or low savings results:** Implementers should review the savings contribution of each reporting-period data point and investigate periods that produce unusually high or low savings, confirm the cause of anomalies, exclude non-routine events from savings calculations, and document retained data points to ensure reported savings reflect SEM-driven operational improvements rather than external factors. Future evaluations will track progress using the same metric described under Recommendation 4 above. • **Ensure top-down model validity or fully adopt bottom-up calculations when necessary:** Implementers should validate the top-down model by confirming variables' relevance and statistical significance. If implementers are unable to establish a valid model, they should fully transition to bottom-up calculations and avoid combining incomplete elements of both approaches as this undermines the validity and defensibility of savings estimates. Future evaluations will measure how often implementers rely on incomplete elements of both top-down and bottom-up approaches to estimate savings.	All PAs

Item #	Page #	Findings	Comment	Recipient
7-9	46	We reviewed all submitted calculations for the sampled projects and made corrections to methodologies, inputs, and parameters based on detailed engineering review and as-built information provided by facility personnel. Where available, we relied on extended trend data to better reflect typical annual operating conditions rather than short-term snapshots. These updates ensured that savings estimates were consistent with verified equipment performance and actual facility operations. Key adjustments included: • For production-based EPIAs,[63] we recalculated key parameters using full-year data when available or the longest and most representative trend period otherwise, replacing values derived from limited datasets. • For compressed air EPIAs, we updated system efficiencies (CFM/kW) using verified CAGI performance data,[64] revised as built discharge pressures for pressure reduction measures, and corrected savings where operating conditions differed from implementer assumptions. • For boiler-related EPIAs, we updated boiler efficiency and loading using the best available information, including facility logs and extended operational data, to reflect typical annual operation. • For weather-dependent EPIAs, we consistently applied the National Oceanic and Atmospheric Administration (NOAA) station-based weather data and reprocessed weather inputs to ensure accuracy and consistency.	• **Apply consistent screening and documentation of anomalous data points:** Implementers should apply a consistent screening method to flag baseline and reporting-period data points outside established thresholds and document whether flagged points reflect normal operations or a non-routine event. Clear documentation of flagged points will preserve institutional knowledge and support transparent and accurate evaluation in future years. Future evaluations will track consistent flagging and documentation of anomalous data points as normal or NREs. • **Review and validate anomalously high or low savings results:** Implementers should review the savings contribution of each reporting-period data point and investigate periods that produce unusually high or low savings, confirm the cause of anomalies, exclude non-routine events from savings calculations, and document retained data points to ensure reported savings reflect SEM-driven operational improvements rather than external factors. Future evaluations will track progress using the same metric described under Recommendation 4 above. • **Ensure top-down model validity or fully adopt bottom-up calculations when necessary:** Implementers should validate the top-down model by confirming variables' relevance and statistical significance. If implementers are unable to establish a valid model, they should fully transition to bottom-up calculations and avoid combining incomplete elements of both approaches as this undermines the validity and defensibility of savings estimates. Future evaluations will measure how often implementers rely on incomplete elements of both top-down and bottom-up approaches to estimate savings.	All PAs

[62] NOAA (National Oceanic and Atmospheric Administration) is the U.S. federal agency that operates the national network of ground-based weather stations and provides quality-controlled, verified weather observations widely used for engineering analysis and energy modeling. Their website is https://www.noaa.gov.

[63] Energy Performance Improvement Action (EPIA) – A documented operational or capital action that improves facility energy performance and includes required savings and cost information for CPUC reporting.

[64] The Compressed Air and Gas Institute (CAGI) is the industry association that provides standardized performance data for compressed-air equipment. CAGI performance sheets give verified manufacturer data such as CFM per kW, which is used to benchmark compressor efficiency. Their website is https://www.cagi.org/

Item #	Page #	Findings	Comment	Recipient
10	47	We observed that implementers have begun applying the proration requirement;[65] however, the application of proration was often unclear or inconsistent. Because they often treat proration as a straightforward step, sites apply it differently, and new members of the implementation teams may struggle to determine the correct carry-forward values. Several projects did not document the proration method used, and in some cases, implementers carried forward an entire year of savings rather than only the remaining portion after proration. These issues indicate that SEM participants would benefit from a standardized method for calculating and documenting prorated savings.	**Develop a standardized proration tool:** Implementers should develop a simple, standardized proration tool for bottom-up EPIAs to ensure consistent calculation of prorated savings and clear identification of the savings to carry forward to the next reporting period. Future evaluations will assess whether implementers develop a standardized proration tool and the frequency and magnitude of proration corrections.	All PAs

[65] The SEM M&V Guide requires that implementers prorate the annualized savings for bottom-up EPIAs based on the documented implementation date, with reasonable consideration for seasonality and simplicity. This guidance directs implementers to attribute only the portion of savings realized within the current reporting period and to carry forward the remaining savings into subsequent periods. Sergio Dias Consulting. "California Industrial SEM M&V Guide, Version 3.02." Section 8.2. July 6, 2022. https://pda.energydataweb.com/api/view/2648/CA_SEM_MV_Guide_v3.02.pdf

Item #	Page #	Findings	Comment	Recipient
11-12	47	• We observed a significant improvement in participant response rates and willingness to join the introductory evaluation interviews and initial data collection. However, response rates dropped sharply once we requested trend data, parameters, and inputs that implementers originally used to calculate savings for bottom-up measures. • Participants relying too heavily on implementation partners limits the educational purpose of the SEM program and reduces the long-term value of past participation at these facilities. • When facilities lack clear handover practices or do not store information on previous SEM activities and implemented measures, it becomes harder to support evaluation and verification in future years. • When energy teams do not take ownership of implemented measures or the trend data and parameters used in savings calculations, they struggle to fulfill evaluation data. requests, which in turn creates risks for meeting evaluation timelines.	• **Support continuity during energy champion transitions:** Implementers should ensure new energy champions can quickly become familiar with prior SEM activities and development history by directing them to existing project documentation, ongoing SEM Activity Logs, and holding a brief orientation meeting. Implementers should document this orientation either as a standalone summary or within the reporting period performance report. This approach maintains continuity at facilities when staff turnover occurs. Future evaluations will verify through the gross data collection interviews with new energy champions whether implementers provided adequate transition support. • **Emphasize participant ownership of SEM data and documentation:** Implementers should emphasize participant ownership of trend and activity logs, production data, and operational parameters during SEM workshops and coaching to ensure timely support for evaluation data requests. Future evaluations will assess progress through gross savings data collection interviews by verifying energy champions' knowledge of SEM data sources and reduced reliance on implementers, and document it in the CCT. In addition, future evaluations will track the adoption rate of designated SEM Data Leads, consistent with recommendations in the SEM Design Guide.[66]	All PAs

[66] Sergio Dias Consulting. "California SEM Program Design Guide Core SEM and Graduate Pathway, Version 2.1." Section 1.2.5 September 1, 2025.
California SEM Design Guide: Core SEM and Graduate Pathway

Item #	Page #	Findings	Comment	Recipient
13-14	48	The Expansion Study[67] linked program success to program delivery that follows the SEM Guides. The Combined TDA Score measures how well the program complies with the letter and spirit of the Guide. Projects that meet a minimum standard of compliance for SEM delivery (currently set at 50%) warrant the SEM NTGR. Compliance and competence standards 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.	• **Use the current NTG survey instruments in future SEM evaluations:** Evaluators should use the TDA and SRA instruments we developed in this impact evaluation to ensure consistency with the findings and recommendations of the Expansion Study. Future evaluations should document the rationale for using alternative NTG survey instruments, if applicable, and obtain concurrence from CPUC staff prior to implementation. • **Reassess the 50% minimum compliance threshold for SEM NTGR assignment:** Evaluators should revisit the current 50% minimum compliance standard and determine if it remains an appropriate threshold for assigning a site the SEM NTGR (currently 1.0). The 50% threshold may overrepresent meaningful compliance with SEM Guide requirements and requires a potential update. Future evaluations should document the basis for retaining or revising the current 50% minimum compliance threshold for SEM NTGR assignment.	All PAs

[67] 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

APPENDIX G. RESPONSES TO COMMENTS ON DRAFT REPORT

Table G-1 presents DNV’s responses to the comments on the draft report that were received during the public review period.

Table G-1. Responses to comments on draft report

Comment #	From	Report content	Comment
1	PG&E	**Page 22**: In Section 3.2.5, Table 3-8, the table states: "60% of overall gas savings and 56% of overall MMBtu come from BU calculations".	It might be valuable to break this out by PA to identify any differences and highlight programs that may have developed effective practices for improving the percentage of successful TD models.
		DNV response	

Table G-2 below includes a breakdown of savings calculation methodologies by PA and fuel. In future evaluation reports, we will consider providing this level of detail consistently to better highlight differences in performance across programs.

Table G-2. Breakdown of savings calculation methodologies by PA and fuel

PA	Fuel	Top-down	Bottom-up	Mixed analysis
MCE	Electric	100%	0%	0%
MCE	Gas	100%	0%	0%
PG&E	Electric	46%	30%	24%
PG&E	Gas	14%	62%	24%
SCE	Electric	24%	76%	0%
SCE	Gas	N/A	N/A	N/A
SCG	Electric	N/A	N/A	N/A
SCG	Gas	47%	39%	14%

Comment #	From	Report content	Comment
2	SCG	**Page 44**: In Section 5.1, the report states the sentence, "without early intervention, the program will continue drifting away from its preferred methodology of top-down modeling,"	This statement could be interpreted as a requirement to use top-down M&V methodology. SoCalGas requests clarifying that implementers should continue to use whichever method best fits the site's operations, even if trends show less top-down usage year-over-year. Clarifying this statement helps explain that bottom-up methodology is still acceptable, given the situation.
		DNV response	

We acknowledge SoCalGas’s perspective, but we do not agree that the statement implies a requirement to use top-down M&V.

The report language is consistent with the SEM M&V Guide, which identifies top-down modeling as the preferred methodology as it can capture whole-facility impacts and interactive effects. The Guide also clearly allows for bottom-up methods when top-down methods cannot feasibly calculate accurate savings, as long as participants document and justify this determination.

The sentence in question describes an observed program trend and its’ potential implications for alignment with program design, not a prescription or mandate for a specific methodology. In particular, the statement reflects the finding that much use of bottom-up methods appears driven by gaps in early data infrastructure and readiness during the baseline period, rather than inherent site constraints. This context is important in understanding the observed shift away from top-down modeling.

Because this sentence is describing program conditions and not establishing a requirement, we propose no revision.

Comment #	From	Report content	Comment
3	**SCG**	**Page 9**: Section 5.1, **Recommendation 10**: "Develop a standardized proration tool: Implementers should develop a simple, standardized proration tool for bottom-up EPIAs to ensure consistent calculation of prorated savings and clear identification of the savings to carry forward to the next reporting period. Future evaluations will assess whether implementers develop a standardized proration tool and the frequency and magnitude of proration corrections."	SoCalGas suggests updates to the M&V Guide with a standardized calculation method to avoid inconsistencies across implementers in accordance with recommendation 10. While the M&V guide includes some proration requirements, further clarity can be beneficial.
DNV response			

We plan to pursue feedback from program administrators (PAs) and implementers on this recommendation through the Response to Recommendations (RTR) process and statewide forums. These discussions will help identify an appropriate path to developing a standardized proration approach, including potential timing considerations. The conversation will also include the authors of the Statewide M&V Guide to assess the feasibility of incorporating this recommendation into a future Guide update.

Comment #	From	Report content	Comment
4	**SCG**	**Page 32**: Section 5.1, the report states: "**Operating conditions** – These discrepancies reflect verified changes in site operations relative to forecasting assumptions, including changes in load profile, production, set points, and operating hours. Most instances resulted in negative adjustments where actual production, operating hours, load shapes, or control settings led to reduced evaluated savings relative to assumptions."	SoCalGas request specific examples of how operating conditions negatively affected GRR. This additional context can help future evaluations and help understand if these impacts are related to top-down or bottom-up savings
DNV response			

Operating condition discrepancies are more commonly observed in bottom-up savings estimates, as these approaches are measure-level and therefore more sensitive to site-specific operational inputs. During evaluation, differences between assumed and actual operating conditions such as load profiles, production levels, equipment performance, and operating hours frequently led to adjustments in evaluated savings. Below are some examples:

- **Load profile adjustments**: Actual temperatures, speeds, and loads differed from reported values, changing savings—often reducing them when impacted equipment operated with lower loads or less hours.
- **Process parameter differences**: Measured conditions (e.g., condensate return temperatures, firing rates) were less than predicted, reducing savings.
- **Operating hours:** Verified runtimes were shorter than reported, reducing savings.
- **Production levels:** Actual annual production was less or more variable than predicted, reducing normalized savings.

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.