GROUP D

Site-Level Normalized Metered Energy Consumption (NMEC) Impact and Net-to-Gross Evaluation, Program Year 2023

California Public Utilities Commission
CALMAC ID: [CPU0385]
Date: June 10, 2025

Information	
Project sponsor	Tyler Mahone
Project manager	Stephanie Whalley
Telephone number	+1 (510) 891-0446
Mailing address	1999 Harrison St., Suite 2150, Oakland, CA 94612
Email address	Tyler.mahone@dnv.com; stephanie.whalley@dnv.com
Report location	

Legal Notice

Table of contents

List of figures

List of tables

Glossary of key terms and acronyms

This section can be deleted if there are only a handful of acronyms or fewer.

Baseline period – The 12-month period leading up to the energy efficiency intervention or retrofit.

Calculated savings – For NMEC projects, a sum of the initial claimed savings and trued-up savings found in CEDARS. Calculated savings is expected to equal normalized savings.

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.[1]

California Energy Data and Reporting System (CEDARS) – This database securely manages California Energy Efficiency Program data reported to the Commission by investor-owned utilities, regional energy networks (RENs), and certain community choice aggregators (CCAs).[2]

Coefficient of determination (R-squared or R^2) – A model goodness-of-fit statistic, the proportion of the variation in the dependent variable (in this case, energy consumption) explained by the regression model. The higher the R^2, the better the model explains variation in the dependent variable.

Coefficient of variation of the root mean square error (CV(RMSE)) – A model goodness-of-fit statistic, a measure of variability (the square root of the consumption model's squared error) relative to the average value of the variable (in this case, average energy consumption) used to determine how well the model predicting the variable (in this case, baseline consumption) fits the data. The lower the CVRMSE, the better the model fit.

Custom project review (CPR) – The process of selecting custom projects, submitted biweekly by the program administrators, for review of all forecasted savings parameters and project documents.

Disposition – Usually, the CPUC Project Review document that summarizes any issues or comments related to project eligibility, baseline, savings calculation, or program influence documentation.

Documented realization rate (DRR) – The ratio of the evaluation-verified savings relative to the savings forecasted in the project documentation.

Early opinion – Review that allows the PAs to request from CPUC staff clarification of custom-project policies or rules before submitting a project.

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

Forecasted savings – An engineering-based savings estimate calculated before installation.

Fractional savings – The percent of annual energy usage saved through program participation. For NMEC projects, the rulebook recommends that projects have a forecasted fractional savings of at least 10%.

---

[1] California Public Utilities Commission (CPUC). *Resolution E-5152*. August 5, 2021. http://www.deeresources.com/files/DEER2023/Resolution%20E-5152%20DEER2023%20Complete.pdf.

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

Fractional savings uncertainty (FSU) – CV(RMSE) combined with percent savings, this statistic is similar to relative precision in that it measures the uncertainty around the expected savings. As the value FSU decreases, confidence in the estimated savings level increases.

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 PAs will come to fruition.

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

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

International Performance Measurement and Verification Protocol (IPMVP)[3] – A standardized approach to measuring and verifying energy efficiency investments. IPMVP incorporates M&V best practices in a non-prescriptive framework, allowing it to be applied flexibly based on a measure's application and the available information.

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 to account for the untapped useful life of the outgoing equipment.

Measure – A 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.

Measure application type (MAT) – The installation basis for efficiency claims. 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.

Net savings – The savings realized after accounting for free-ridership, calculated by multiplying 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; used to estimate and describe the free-ridership that may be occurring within energy efficiency programs.

Non-routine adjustment (NRA) –Adjustments used to account for the effects of non-routine events (NREs), when the NRE unsuitably distorts the baseline or reporting period adjustment models. NRAs are made separately from routine adjustments, which use 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) – A change not related to the energy efficiency intervention but affecting energy use in the baseline or the reporting period and which 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.

Normalization – A process by which consumption estimates from two different periods are put on a common basis. Baseline and performance period model predictions are observed at common values for the model's independent variables,

[3] Efficiency Valuation Organization (EVO). https://evo-world.org/en/.

e.g., temperature variables from typical meteorological year (TMY) data. Normalization accounts for differences in underlying drivers of consumption during the baseline and performance periods.

Normalized mean bias error (NMBE) – A model goodness-of-fit statistic that can indicate whether a model is overestimating or underestimating energy use.

Normalized metered energy consumption (NMEC) – High opportunity programs or projects 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 using a program and project approach where incentive payment and claimed savings are based on NMEC and include only approved NMEC building programs.

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

Occupied/unoccupied split – Within the standard time-of-week and temperature (TOWT) model structure, the use of two models to account for changes in occupancy over the course of a week. This enables the model to capture an occupancy-temperature interaction.

Parameter – Output of a regression model. For NMEC models, parameters measure how fuel consumption changes in response to a change in a given independent variable.

Peak demand – The maximum level of metered demand during a specified peak demand period for installed or implemented measures. CPUC Resolution E-4952 approved the Database for Energy-Efficient Resources (DEER) for 2020 and revised the DEER Peak Period definition to 4:00 p.m. to 9:00 p.m. effective January 1, 2020

Performance period – The 12-month period following the energy efficiency intervention or retrofit, during which savings are realized.

Program administrator (PA) – An entity tasked with the functions of portfolio management of energy efficiency programs and program choice, i.e., Marin Clean Energy (MCE),[4] Pacific Gas & Electric (PG&E), Southern California Edison (SCE), Southern California Gas (SCG), and San Diego Gas & Electric (SDG&E).

p value – The probability that a given parameter's true value is different from zero.

Relative precision – A ratio of the error bound divided by the value of the measurement itself. This provides the error on a relative basis, 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.

Savings delta – The difference between normalized savings and forecasted savings.

Spline – A model that is a collection of lines with different slopes that change at defined points (nodes), allowing for more flexible response to the given independent variable than a constant linear relationship.

Temperature node – In a temperature spline model, a boundary temperature at which the slope changes.

Time-of-week and temperature (TOWT) model – A standard regression model approach whereby fuel consumption is modeled against temperature, included as a spline and a set of time-of-week indicator variables, generally at the daily- or hourly-level. May be split into occupied and unoccupied models. Other variables may also be included.

[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 initial claimed savings and the normalized savings.

Typical meteorological year (TMY) – A data set of temperatures representing a typical year and used to normalize NMEC models to weather conditions. The CALEE CZ data sets are the standard used for NMEC.

1 EXECUTIVE SUMMARY

This report presents DNV's evaluation of the Site-Level Normalized Metered Energy Consumption (site-level NMEC) Programs for program year (PY) 2023 on behalf of the California Public Utilities Commission (CPUC). The evaluation determines how much electric, peak demand, and natural gas energy use was reduced by the NMEC programs.

1.1 STUDY BACKGROUND

NMEC is a set of statistical tools and approaches that estimate the energy consumption impact of energy efficiency programs by comparing pre- and post-intervention meter data. While most other energy efficiency programs claim final savings based on deemed[5] or calculated results, NMEC programs calculate and claim final savings based on measured impacts at the meter. This evaluation estimates energy savings for NMEC projects at the individual commercial site level. The gross evaluation provides savings estimates[6] for site-level NMEC projects with true-up claims[7] in PY2023. The net-to-gross (NTG)[8] evaluation provides program attribution for site-level NMEC projects with initial claims in PY2023.

Site-level NMEC programs are relatively new and the CPUC and Program Administrators continue to develop rules and guidelines. This evaluation builds on the PY2020 – PY2022 evaluation and the site-level NMEC Evaluability Study.[9] Guided by the site-level NMEC workplan,[10] DNV will produce two reports for this evaluation:

1. **PY2023 Site-level NMEC Impact Evaluation Report:** This report covers the gross and net evaluation of fully claimed projects from PY2023.
2. **PY2023 Site-level NMEC Additional Research Report**: DNV will provide a second report after the Impact Evaluation Report, which will cover the early gross evaluation projects and the additional research questions. The early gross evaluation will include projects with initial claims in PY2022 and PY2023 that have not yet been trued up. When possible, the early evaluation will include most of the same activities as the typical full gross evaluation. The additional research questions seek to fill information gaps regarding how site-level NMEC programs are functioning.

[5] Deemed refers to researched, vetted, and predictable savings for EE technologies and services with well-established properties. This contrasts with custom savings for EE technologies and services that require unique calculations and do not use predefined values.

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

[7] Site-level NMEC projects typically have two claims, with an initial claim made at the time of installation followed by a true-up claim after a 12-month performance period.

[8] Net savings are changes in energy use attributable to a particular energy efficiency program. They include savings from participants who would not have purchased energy-efficient technologies without the program. Savings attributable to participants who would have purchased energy-efficient technologies with or without the program influence are excluded from net savings. These participants, whom the program did not influence, are considered free riders.

[9] DNV, *Site-Level NMEC Evaluability Study, Program Years 2020-2021*, calmac.org, December 7, 2023, https://www.calmac.org/publications/Site-Specific_NMEC_Evaluability_Study_Report_-_Final.pdf.

[10] CPUC, *Site-level Normalized Metered Energy Consumption (NMEC) Evaluation, Measurement, & Verification Workplan, Program Year 2023*, pda.energydataweb.com, December 12, 2024, https://pda.energydataweb.com/api/view/4101/PY2023%20CPUC%20Site-level%20NMEC%20Evaluation%20Workplan%20-%20Final.pdf.

1.2 EVALUATION OBJECTIVES

For this evaluation DNV estimated the gross[11] and net savings of site-level NMEC projects. DNV also assessed the application of NMEC program requirements as outlined by the Rulebook for Programs and Projects Based on NMEC,[12] which includes the CPUC's specific requirements for NMEC programs and measurement and verification (M&V) plans.

The objectives of this evaluation are:

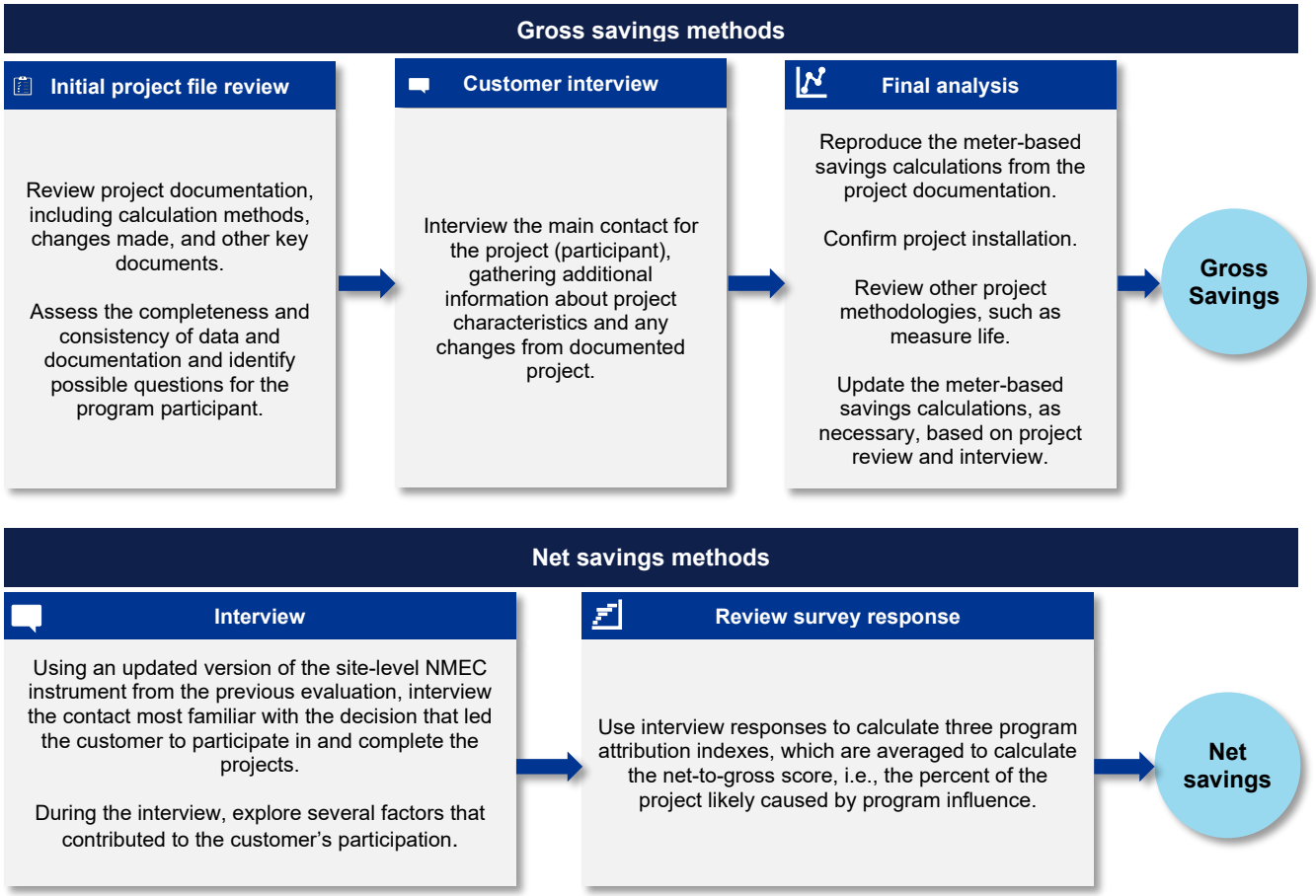
- Estimate gross kWh, peak kW, and therm savings for site-level NMEC projects with true-up claims in PY2023.
- Estimate net kWh, peak kW, and therm savings for site-level NMEC projects with initial claims in PY2023.
- Assess the methods used by program administrators (PAs)[13] and their third-party implementers to estimate meter-based savings.
- Provide timely feedback to the CPUC, PAs, and other stakeholders, facilitating program improvements and supporting future program design efforts.
- Provide meaningful and actionable recommendations to improve program performance in delivering energy efficiency savings.

1.3 STUDY APPROACH

This study included both a gross and net savings evaluation. For the gross evaluation, DNV reviewed important project documentation and, where necessary, updated the project's approach to calculate project-level savings. For the net-to-gross evaluation, DNV investigated how much the program influenced the participant's decision to make energy efficient improvements. Figure 1-1 illustrates the gross and net savings methodology in more detail.

11 Gross savings are a measure of change in energy use due to EE programs, regardless of why customers participated.
12 CPUC, *Rulebook for Programs and Projects Based on Normalized Metered Energy Consumption, version 2.0*, January 7, 2020, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.
13 A program administrator is an entity managing a portfolio of energy efficiency programs and program choice.

Figure 1-1. Site-level NMEC gross and net savings methods



1.4 EVALUATED PROJECTS

This evaluation included three projects from PY2023 that were ready for gross evaluation and 48 projects that were ready for NTG evaluation. The two populations differ because DNV conducts the evaluations at different project stages. DNV expects each site-level NMEC project to make two claims in the California Energy Data and Reporting System (CEDARS) tracking database. First, at the time of project installation, an NMEC project makes an engineering-based, forecasted savings claim. Second, a year later, after the performance period, the project calculates its meter-based normalized savings and makes a true-up claim representing the difference between the initial claim and the results of the performance-based measurement. The initial claim and true-up claim sum to the total meter-based estimate of savings.[14] The gross evaluation requires finalized projects and thus includes projects that have made that second, true-up claim in PY2023. The net evaluation only requires projects to be installed, and interviews are most fruitful when conducted as soon after installation as possible. The net evaluation includes all projects with initial claims made in PY2023 (Table 1-1).[15]

[14] Projects make true-up claims at least a year after the installation of the project. These claims are positive or negative savings differences, adjusting the initial claim up or down so that it aligns with the meter-based normalized savings.

[15] DNV followed this same schedule in last year's evaluation. The three projects included in this year's impact evaluation were included in last year's net evaluation.

Table 1-1. Site-level NMEC evaluation population and savings claimed in the tracking data

Program administrator*	Projects	First year savings			Lifecycle savings	
		kWh	kW	Therms	kWh	Therms
Gross evaluation population						
SCE	1	915,788	243	0	10,806,303	0
SoCalREN	2	124,473	-129	0	1,730,175	0
Total	**3**	**1,040,261**	**114**	**0**	**12,536,478**	**0**
Net evaluation population						
PG&E	17	12,834,707	1,517	321,168	103,716,772	2,135,814
SCE	25	2,070,845	147	-68	26,154,162	-799
SoCalREN	2	874,782	53	0	2,919,282	0
SDG&E	2	816,100	265	0	9,793,203	0
Total	**46**	**16,596,435**	**1,982**	**321,100**	**142,583,419**	**2,135,015**

* Southern California Edison (SCE); Southern California Regional Energy Network (SoCalREN); Pacific Gas & Electric (PG&E); San Diego Gas & Electric (SDG&E)

1.4.1 Gross savings

DNV's gross evaluation looked at the three projects that submitted true-up claims in PY2023.[16] The PY2023 Additional Research Report that DNV will produce later in the year will include a larger population of sites that made initial claims in PY2022 and PY2023 but have not yet submitted true-up claims. This second report will include additional details and recommendations drawing from that larger population.

Table 1-2 and Table 1-3 present the evaluated electricity and peak demand savings.[17] The tables show two sets of realization rates: a gross realization rate (GRR), which compares the evaluated savings (shown in the "Verified" column in the tables) with the savings claimed in the tracking data ("Claimed"), and a documented realization rate (DRR), which compares the evaluated savings with the savings provided in the project documentation ("Documented"). If projects report savings correctly, the two realization rates will be the same. In two out of three projects, the documented and claimed savings matched. This is a notable improvement from last year's evaluation as both the GRR and DRR are closer to 100% compared to the PY2020-2022 first-year savings DRR of 81.2% and GRR of 70.9%. One of SoCalREN's projects mistakenly zeroed out its savings in the tracking data, resulting in the difference between that PA's overall Documented and Claimed savings and their DRR and GRR.

Table 1-2. Gross electricity savings in kWh

Program administrator	Claimed*	Documented**	Verified†	GRR††	DRR‡
First-year savings					
SCE	915,788	915,788	916,970	100.1%	100.1%
SoCalREN	124,473	280,888	192,945	155.0%	68.7%
Statewide	**1,040,261**	**1,196,676**	**1,109,915**	**106.7%**	**92.7%**
Lifecycle savings					
SCE	10,806,303	10,806,298	11,095,336	102.7%	102.7%
SoCalREN	1,730,175	3,763,570	3,131,259	181.0%	83.2%
Statewide	**12,536,478**	**14,569,868**	**14,226,595**	**113.5%**	**97.6%**

*Claimed savings are those initially claimed by projects in the tracking data.
**Documented savings are the projected savings provided in project documentation.
†Verified savings are the evaluated project savings.
††Gross realization rate (GRR) compares Verified savings with the Claimed savings.
‡Documented realization rate (DRR) compares Verified savings with Documented savings.

[16] Impact evaluations are usually conducted on a statistical sample, but, given the small eligible population, DNV evaluated all complete projects.
[17] The projects trued-up in PY2023 did not claim gas savings.

Table 1-3. First-year gross demand savings in kW

Program administrator	Claimed	Documented	Verified	GRR	DRR
SCE	243.2	243.0	250.8	103.1%	103.2%
SoCalREN	-128.9	35.8	29.3	-22.7%	82.0%
Statewide	**114.2**	**278.8**	**280.1**	**245.2%**	**100.5%**

1.4.2 Net savings

Table 1-4 shows the evaluated net-to-gross ratios (NTGRs) for electric energy savings (76.6%), demand savings (75.2%), and natural gas savings (76.0%) among the 46 projects with initial claims made in PY2023. In this table, "Claimed" refers to those savings claimed in the tracking data. The NTGRs are higher than in last year's evaluation, which found NTGRs of 45.9%, 41.7%, and 46.5% respectively. The increase in NTGRs from the previous evaluation to this one is the result of two primary differences: 1) stronger program influence as reported by this year's sample compared to last year's sample and 2) updates to the NMEC NTG scoring methodology. At the end of 2024, as part of a larger effort to update all evaluation NTG methodologies, the CPUC adopted updates to the site-level NMEC NTG methodology to better align with NMEC program delivery and goals.[18] The update streamlined the instrument, removed overlap between different indicators in the algorithm, and removed any penalty for organizational sustainability policies or goals.

Table 1-4. Net electric energy savings results by PA

Program administrator	Projects	First-year net savings				Lifecycle net savings	
		Claimed	Net	NTGR	RP%*	Net	RP%*
Energy (kWh)							
PG&E	12	12,834,707	9,673,519	75.4%	±7.0%	78,347,649	±4.0%
SCE	24	2,070,845	2,058,834	99.4%	±0.0%	25,976,314	±0.0%
SoCalREN	2	874,782	558,986	63.9%	±0.0%	1,931,105	±0.0%
SDG&E	2	816,100	413,600	50.7%	±0.0%	4,963,196	±0.0%
Statewide	**40**	**15,780,335**	**12,291,339**	**76.6%**	**±5.0%**	**106,255,068**	**±3.0%**
Demand (kW)							
PG&E	12	1,517	1,167	76.9%	±10.0%		
SCE	18	147	147	100.0%	±0.0%		
SoCalREN	1	53	45	86.1%	±0.0%	NA	
SDG&E	2	265	134	50.7%	±0.0%		
Statewide	**33**	**1,982**	**1,494**	**75.2%**	**±8.0%**		
Natural gas (therms)							
PG&E	5	321,168	244,120	76.0%	±7.0%	1,626,209	±4.0%

* DNV used tracked savings in the ratio estimation.
** Relative precision is at the 90% confidence level.

[18] The CPUC also adopted an updated NTG survey instrument that better aligns with the current NMEC program design. The revised instrument is included in Appendix B of this report. The instrument and methodology updates incorporated PA and stakeholder feedback.

1.5 KEY FINDINGS & RECOMMENDATIONS

1.5.1 Gross and net savings findings and recommendations

NTGRs have increased from last year's evaluation, driven by both increases in reported program influence and by methodological changes.

Increases in reported program influence: The NTG survey asked respondents to divide a total of 10 points across two types of influential factors, those that are program-related and those that are non-program-related. The relative allocation of these points indicates the importance of all program influences relative to all non-program influences in the decisions to implement projects. In last year's evaluation, respondents gave an average of 5.1 points to program influences, indicating that program and non-program influences were equally important in their organizations' decisions to do the EE projects. This year, respondents gave an average of 7.5 points to the program influences, indicating that program influences were more important than non-program influences. Respondents particularly highlighted programs' technical support with one saying, "They offer an extension of our staffing. I don't have time to go and do research on what programs are available and what aligns most with our school district's needs. So having ongoing meetings with them to touch base gives me an extension of my capacity." Better program involvement with the customers improved program influence.

Decision timing: Additionally, no respondent in the current evaluation indicated that their organization had decided to do the project before interacting with the program, compared to 40% of respondents last year. Similarly, in response to a new question in this year's survey instrument, 80% of respondents indicated that their organization first learned about the opportunities included in their EE projects from program or utility staff. One respondent said, "[The biggest strength of the program was the] awareness that those audits brought to us. The audits bring a lot of issues to our attention which are hard for us to recognize ourselves."

Methodology update: To assess the impact of the methodology update on NTGR estimates, DNV calculated NTGRs for this year's sample using the previous methodology. Table 1-5 compares the NTGRs using the current and the old methodology. Using the previous methodology, NTGRs for this year's sample range from 60.0% to 62.8%. Approximately half of the increase in NTGRs from PY2020-22 to PY2023 is attributable to the changes in scoring methodology.[19]

Table 1-5. NTGR methodology and sample comparison

NTGR	PY2020-2022 evaluation Old methodology	PY2023 evaluation Old methodology	PY2023 evaluation New methodology
Statewide electricity	45.9%	61.3%	76.6%
Statewide demand	41.7%	61.0%	75.2%
Statewide natural gas	46.5%	60.2%	76.0%

[19] Note it is not possible to run the PY2020-2022 evaluation results through the new methodology because questions were asked differently to respondents in that evaluation. However, the Group D NTG Methodology Update_FinalMem-01072025-CLEAN memorandum provides a sensitivity analysis (Section 3.3.1) to the extent possible. The sensitivity analysis found for NMEC that the new methodology increased NTGRs by about 9%. The new methodology tends to make high NTGRs higher and low NTGRs lower when compared to the previous methodology, which pushed NTGRs towards 0.5.

The accuracy of savings claims in the tracking database system improved compared to PY2020-2022. However, incorrectly entered savings claims remained an issue in PY2023.

The NMEC savings claim process is more complicated than the typical custom claim process as it must accommodate the final, meter-based savings estimate, which projects calculate a year after implementation. Projects claim engineering-based, forecasted savings during implementation. A year later, after the performance period, projects calculate the meter-based normalized savings and enter into tracking a true-up claim that represents the difference between the two values. The two claims should sum to the final, meter-based savings estimate. The novel claims process for NMEC led to a reporting inaccuracy for one out of three projects during PY2023. This was an improvement compared to PY2020-2022, when incorrectly entered savings claims in the tracking database system were the largest source of savings discrepancies. In the PY2020-2022 evaluation, seven out of 22 sampled projects incorrectly entered savings.

The PY2023 project with the reporting inaccuracy (mentioned above) submitted two true-up claims, which over-adjusted the initial claimed savings. Together, the three claims for this project resulted in zero first-year kWh savings.

During the Response to Recommendations process following the PY2020-PY2022 evaluation report, the PAs indicated that they were working to improve the claims process. Future evaluations will continue monitoring this issue.

Recommendation

PAs should continue to focus on improving the timeliness and accuracy of site-level NMEC claims, following existing NMEC reporting guidance:[20] make the initial claim in the quarter measures are installed, use a ProjectID that can be tracked across project years, and make the true-up claim in the quarter in which the performance period is completed.

Models should reflect empirical conditions within pre- and post-installation periods and be normalized across the two periods.

NMEC savings claims rely on statistical models that characterize site-level consumption and allow for the comparison of pre- and post-installation consumption at the site on an apples-to-apples basis. To be eligible for NMEC, the pre-installation model needs to demonstrate the ability to explain site-level consumption well enough that expected savings will be discernible. The process also needs to encompass all changes in site-level consumption that are not program related.

Reliance on modeling algorithms rather than empirical schedule data: While the NMEC Rulebook is not prescriptive regarding the details of modeling decisions, it does recognize the central role of quantitative empirical data in the NMEC modeling process. All three evaluated sites used a popular site-level modeling package that decides how to characterize the site solely on statistical fit. While this approach has merit, NMEC implementers should also consider the wider context of empirical data. For example, implementers should consider known site schedules and override model-based outcomes where indicated by empirical and contextual information. Similarly, the modeler needs to decide if schedules affect only non-weather-correlated, baseload consumption or also affect weather correlated (HVAC) consumption. Empirical data has a role to play in those decisions as well.

Accounting for known pre-to post-installation changes: Meter-based savings methods must be informed by any empirical data relevant to energy consumption during the pre- and post-installation periods. For example, a project with different operating schedules before and after installation, a change that is unrelated to program participation, should explicitly address the implications for energy consumption. In two instances, a building shortened its operating hours in the period after project installation. With fewer open hours, energy consumption will likely drop even without any other

[20] This is explained in reporting guidance published by Energy Division as NMEC Reporting Guidance 04242020.pdf that was distributed to the PAs.

intervention. Neither of the two projects with this issue attempted to address it, in effect, claiming as savings consumption reductions actually due to a decrease in operating hours. The evaluator's adjustment to reflect the schedule changes resulted in lower verified savings.

Recommendations

- PAs should ensure that models align with empirical data regarding how the building operates and how savings will be achieved. This will help enhance model accuracy and PA savings estimates.
- The PAs should ensure that all projects address any changes to site operations so that projects claim only savings due to energy efficiency improvements rather than, for example, a reduction in hours.

Projects used Effective Useful Life (EUL) values without providing justification.

Two out of three projects reviewed during the gross evaluation had insufficient documentation to justify certain measure-level EUL values.[21] A measure's EUL indicates how long the first-year savings will persist and must be based on documentation, just as with non-meter-based custom projects. EULs cannot be measured using one-year of post-installation consumption data because EULs are generally longer than one year. Evaluators must carefully review EUL, as the resulting lifetime savings are important for calculating cost-effectiveness and total system benefit. For two replacement measures, projects used an EUL of 15 years, without providing justification, while the DEER-based EUL for both measures is 20 years. For these three projects, the evaluated EULs were higher than the claimed EULs, which resulted in an increase in lifetime savings.

Recommendation

Measure documentation should include a description of the measure, its EUL, and its respective DEER EUL ID (or other similar citation) to justify each measure's EUL.

1.5.2 Documentation findings and recommendations

Project documentation did not sufficiently catalogue either existing equipment or installed equipment.

DNV expects that project documentation will explain the project's initial plans, including pre-existing equipment types and condition, as well as any changes that occurred during project development, implementation, or the performance period. Some projects made a good effort to provide clear documentation. Documentation for all projects, however, could have been improved. The following areas present the greatest room for improvement.

Viability documentation: Understanding the viability of existing equipment is key to determining whether the project is appropriate for NMEC and for selecting the appropriate measure application types (MATs)[22]. Documentation should demonstrate that existing equipment was still serving the requirements of the building to demonstrate that replacement will, in fact, lead to consumption reduction. For one project, after talking to the customer and acquiring the assessment report, the evaluation team noticed that the pre-existing equipment condition for two measures was poor, with, for example, ducts patched with cardboard.

MAT documentation: Most projects reviewed during the evaluation did not provide adequate explanation of selected MATs. One project selected a MAT of Normal Replacement (NR) but provided no supporting evidence showing the condition of pre-

[21] A measure, in this context, is a specific customer action that reduces or otherwise modifies energy end-use patterns or a product whose installation and operation at a customer's premises reduces the customer's on-site energy use.

[22] For more detailed definitions of each MAT, see: https://www.caltf.org/measure-application-types-1

existing equipment. Through the customer interview, DNV determined that the pre-existing equipment had been operating properly. As a result, evaluators reconsidered the MAT and ultimately assigned Accelerated Replacement (AR) instead of NR. Another project designated all replacement measures as AR but also provided information showing that the pre-existing equipment had been broken. Therefore, the evaluated MAT for the two measures became NR instead of AR.

Measures scoped but not installed: Two projects planned to include lighting upgrades, which would have significantly contributed to total anticipated savings. However, the final savings reports indicated that neither project completed a lighting upgrade. If a participant decides not to install some measures with high forecasted savings, the PAs should confirm that the change does not reduce savings to the point NMEC methods are no longer feasible. Additionally, projects need to update the forecasted savings and EUL after installation so that the initial tracking data claims accurately reflect the completed project, whether savings are lower or higher than originally forecasted.

Code compliance: One of the three projects evaluated provided no documentation regarding how the proposed equipment performance would meet to-code compliance.[23]

Recommendation

Even though the savings for site-level NMEC projects are meter-based, it is essential to clearly document the viability of the existing equipment, the details of the measures installed, and code compliance.

Regression-based modeling is the core of NMEC methods, and projects do not consistently provide transparent, well-documented models following standard practices.

Model replication issues: DNV attempted to reproduce projects' model savings results using the code, spreadsheets, and/or other tools provided in the project documentation. Documented and replicated values tended to align, though exact replication was not possible in any of the three cases. One project required changes to the provided code to achieve similar results, suggesting that the code provided was not the version that produced the outputs. For the other two projects, the differences were less meaningful, but some results in the final savings report did not align with the results in the provided output files.

Data processing transparency: None of the projects included the scripts used to process and clean the data used in the models. Strictly speaking, an analyst can replicate the model using the prepared data and the provided modeling code. However, when projects do not include the code used to transform raw data to model-ready data files, analysts cannot verify the methods and assumptions used to prepare the data.

More complete model narratives: While documentation usually explained the relatively novel aspects of models (for example, use of airflow data as an occupancy proxy), the full model narrative was frequently incomplete. Documentation did generally stipulate the model used, the variables used, and whether non-routine events were addressed, but included very little information explaining decisions or interpreting the results. Without clear explanations for modeling decisions, evaluators are left guessing.

Model parameters and summaries: All evaluated M&V plans included a description of baseline model variables; descriptions of non-routine events and related adjustments; and basic goodness of fit statistics. However, none of the M&V plans provided basic baseline model output and specification details, such as parameter estimates, p values, temperature

[23] Assembly Bill 802 requires that utility-incented measures at existing buildings "bring them into conformity with, or exceed, the requirements of Title 24 of the California Code of Regulations."

bin values, and occupied/unoccupied mode settings. These values are essential to ensure successful replication efforts and to fully assess whether the model reflects accurate engineering conditions at the site.

Output errors: For all three projects, the model documentation included temperature and consumption values with timestamps that were off by a day for daily models and an hour for hourly models. While this issue did not make a large difference in savings, it could dramatically affect demand savings, and it was a recurring issue that should be addressed for future projects.

Recommendations

- PAs should provide the data processing scripts or files used to prepare the data for modeling, along with the complete, final modeling scripts and files needed to exactly reproduce the savings in the project documentation and the CEDARs tracking database.
- PAs should provide complete model output as well as a model narrative, including an explanation of any modeling decisions made, changes from the M&V plan, and an analysis of any substantial deviations in savings.

1.5.3 Process findings and recommendations

Participants continued to indicate high levels of satisfaction with the site-level NMEC programs, driven by the programs' technical support and incentives.

When asked to rate program satisfaction on a scale of zero to 10, where zero is completely dissatisfied and 10 is completely satisfied, respondents gave an average rating of 8.4, indicating a high level of satisfaction. This is an increase even over last year's high average rating of 8.1. Respondents highlighted programs' technical support, which boosted their staff's capacity and provided essential information for their decision-making processes. No respondent provided a satisfaction rating lower than 7 and therefore there were no detractors (i.e., respondents who provided a rating of 3 or less.)

Recommendation

Programs should continue all levels of technical support provided to customers, especially pro-active efforts to build relationships.

Customers want a more streamlined process and better communication from program administrators and implementers about program expectations and timeline limitations.

When asked for program improvement suggestions, respondents most frequently mentioned streamlining the participation process and improving communication. One respondent said, "One of the difficulties was the amount of back and forth with technical reviewers. In hindsight, it was a lot of work on our end and we would have preferred to let a consultant or vendor deal with that." Another said, "Communicating and understanding timeline requirements from the onset would be beneficial for both sides. What timeline does the program require and what can we accomplish it in that timeline based on our own restrictions?" Both issues could potentially be improved with clear communication about expectations and timeline requirements.

Recommendations

- In some cases, PAs should consider requesting an exception to the 18-month installation period when there are extenuating circumstances and other reasonable solutions have not worked.

- PAs and implementers should clearly explain timeline needs early in the process, providing, for example, clear project documentation and regular check-in meetings.

2 INTRODUCTION

This report presents key findings from our impact and net-to-gross evaluation of site-level normalized metered energy consumption (NMEC) programs, performed on behalf of the California Public Utilities Commission (CPUC). NMEC uses a set of statistical tools and approaches to estimate the impact of energy efficiency programs on energy consumption, based on pre- and post-intervention meter data. While most other energy efficiency programs claim final savings based on deemed[24] or calculated results, NMEC programs calculate and claim final savings based on measured impacts at the meter.

Each NMEC project makes two claims in the California Energy Data and Reporting System (CEDARS) tracking database. First, at the time of project installation, an NMEC project makes an engineering-based, forecasted savings claim. Second, a year later, after the performance period, the project calculates its meter-based normalized savings and makes a true-up claim representing the difference between the initial claim and the results of the performance-based measurement.

Site-level NMEC projects offer unique advantages compared to other energy efficiency programs. They can help unlock the potential savings stranded when customers maintain equipment beyond its expected useful life. Typically, energy efficiency programs which claim savings based on calculated results assume a hypothetical, code baseline. Customers operating below that baseline are not incentivized to upgrade to that baseline and may not be able to afford to upgrade beyond it. Because NMEC savings are calculated against a pre-intervention, metered baseline, customers can realize savings on any improvement in efficiency. This can be especially valuable to customers who struggle to raise capital for site improvements.

Site-level NMEC projects can also shift risk away from the ratepayer and onto the program administrator (PA), implementer, and customer. Post-installation savings do not remain hypothetical; they are calculated based on actual shifts in metered consumption. If, for any reason, a project does not realize savings, that failure will be discovered and reflected in the true-up claim.

This evaluation estimates energy savings for NMEC projects at the individual commercial site level. The gross evaluation provides savings estimates[25] for site-level NMEC projects with initial claims[26] in program year (PY) 2021 or PY2022 that were trued-up in PY2023. The net-to-gross (NTG)[27] evaluation provides program attribution for site-level NMEC projects with initial claims in PY2023, accounting for those participants who would have installed the program measure or equipment (or some part of it) even in the absence of an incentive program and whose savings, therefore, cannot be fully attributed to the program.

2.1 Background

Over the last decade, the CPUC and the California Program Administrators[28] (PAs) have worked to develop whole-building measurement and verification (M&V) program pathways that can achieve deep savings in commercial buildings.

- 2012: The CPUC requested that its regulated investor-owned utilities (IOUs) develop energy efficiency programs to encourage more comprehensive commercial building retrofits (Decision 12-05-015, 2012).[29]

[24] Deemed refers to researched, vetted, and predictable savings for EE technologies and services with well-established properties. This contrasts with custom savings for EE technologies and services that require unique calculations and do not use predefined values.

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

[26] Site-level NMEC projects typically have two claims, with the initial claim occurring at the time of installation and the true-up claim occurring following the 12-month performance period.

[27] Net savings are changes in energy use attributable to a particular energy efficiency program and consider savings from participants who would not have purchased energy-efficient technologies without the program. Savings attributable to participants who would have purchased energy-efficient technologies with or without the program influence are excluded from net savings. These participants whom the program did not influence are considered free riders.

[28] A program administrator is an entity tasked with the functions of portfolio management of energy efficiency programs and program choice.

[29] CPUC, *Decision 12-05-015*, May 18, 2012, https://www.calmac.org/events/Decision_12-05-15.pdf.

- 2015: The governor signed California Assembly Bill 802, which directed the CPUC to allow savings claims using an NMEC methodology (AB 802 Williams 2015).[30]
- May 2019: The CPUC released the Commercial Whole Building Demonstration Joint Study. This study evaluated a 12-building demonstration program and developed recommendations for future NMEC programs.[31]
- December 2019: The Lawrence Berkeley National Laboratory (LBNL) published its Option C Technical Guidelines, which showed how to use NMEC methods to calculate energy and demand savings for site-level NMEC projects.[32]
- 2020: The CPUC released an updated "Rulebook for Programs and Projects Based on Normalized Metered Energy Consumption" (NMEC rulebook).[33]
- 2023: The CPUC released a draft revised NMEC rulebook that is currently going through the comment process.[34]
- 2023: DNV completed the Site-level NMEC Evaluability study on behalf of the CPUC.[35] The Evaluability Study investigated project characteristics and identified those projects ready for evaluation.
- 2024: The first comprehensive Site-level NMEC impact evaluation was published.[36]

The Evaluability Study and last year's PY2020 – 2022 impact and net-to-gross evaluation were the first comprehensive site-level NMEC evaluations since the NMEC pathway expanded beyond the pilot phase.[37] This evaluation builds on those two previous evaluations. Guided by the site-level NMEC workplan,[38] this evaluation will produce two reports:

1. **PY2023 Site-level NMEC Impact Evaluation Report**: This report covers the gross and net evaluation of fully claimed projects from PY2023. DNV will deliver a draft report to the public in April 2025.

2. **PY2023 Site-level NMEC Additional Research Report**: DNV will provide a second report after the Impact Evaluation Report, which will cover the early gross evaluation projects and additional research questions. The early gross evaluation will include projects with initial claims in PY2022 and PY2023 that have not yet been trued up. When possible the early evaluation will include most of the same activities as the typical full gross evaluation. The additional research questions seek to fill information gaps regarding how site-level NMEC programs are functioning. DNV will deliver this second report in the summer of 2025.

2.2 Evaluation objectives

For this evaluation we estimated the gross and net savings of site-level NMEC programs and assessed the application of NMEC program requirements, referring to the Rulebook for Programs and Projects Based on NMEC,[39] which includes the CPUC's specific requirements for NMEC programs and measurement and verification (M&V) plans.

The objectives of this evaluation were to:

---

[30] California Legislative Information, *AB-802 Energy efficiency, Assembly Bill No. 802, Chapter 50*, October 8, 2015, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160AB802.

[31] CPUC, *Commercial Whole Building Demonstration Joint Study Report*, May 1, 2019, https://www.calmac.org/publications/Commercial_Whole_Building_Joint_Study_ID_PGE0431.01.pdf.

[32] Ibid.

[33] CPUC, *Rulebook, version 2.0*, January 7, 2020, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.

[34] CPUC, *Rulebook for Programs and Projects Based on Normalized Metered Energy Consumption, version 2.1*, Filed November 17, 2023, https://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M520/K881/520881077.PDF.

[35] CPUC, *Site-Level NMEC Evaluability Study,* December 7, 2023, https://www.calmac.org/publications/Site-Specific_NMEC_Evaluability_Study_Report_-_Final.pdf.

[36] CPUC, *Site-Level Normalized Metered Energy Consumption (NMEC) Impact and Net-to-Gross Evaluation, Program Years 2020–2022*, May 23, 2024, https://www.calmac.org/publications/Site-level_NMEC_Evaluation_Final_Report_PY2020-2022.pdf.

[37] The only other evaluation to-date that touched on site-level NMEC was *PY2018–2019 California Statewide On-Bill Financing Impact Evaluation*, written by Opinion Dynamics and published in 2022. That report focused only on the On-Bill Financing (OBF) Program, which was primarily a population NMEC program, but did assess some projects via site-level NMEC. DNV considered the findings in that report as we assessed the wider site-level NMEC programs.

[38] CPUC, *Workplan, Program Year 2023*, December 12, 2024, https://pda.energydataweb.com/api/view/4101/PY2023%20CPUC%20Site-level%20NMEC%20Evaluation%20Workplan%20-%20Final.pdf.

[39] CPUC, *Rulebook, version 2.0*, January 7, 2020, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.

- Estimate gross kWh, peak kW, and therm savings for site-level NMEC projects with true-up claims in PY2023.
- Estimate net kWh, peak kW, and therm savings for site-level NMEC projects with initial claims in PY2023.
- Assess the methods used by implementers to estimate meter-based savings.
- Provide timely feedback to the CPUC, program administrators (PAs), and other stakeholders, facilitating program improvements and supporting future program design efforts.
- Provide meaningful and actionable recommendations to improve program performance in delivering energy efficiency savings.

2.3 Evaluated programs

The CPUC offers the site-level NMEC pathway as part of multiple programs that serve commercial or commercial-like buildings. Our site-level NMEC gross savings evaluation, described in Table 2-1, included two programs with site-level NMEC claims trued-up in PY2023. Our net-to-gross evaluation, described in Table 2-2, included an additional nine programs with initial claims in PY2023.

Table 2-1. Programs included in the gross evaluation

PA	Program ID	Program name	Description
SCE	SCE-13-L-003I	Public Sector Performance-Based Retrofit HOPPs	Targets public sector buildings with stranded savings resulting from improvement delays or indefinite equipment repairs
SoCalREN	SCR-PUBL-B3	Public Agency Metered Savings Program	Targets public sector buildings with stranded savings

Table 2-2. Additional programs included in the net-to-gross evaluation

PA	Program ID	Program name	Description
PG&E	PGE_Com_001	Grocery Efficiency Program (CoolSave)	An NMEC-specific program targeting grocery stores, offering comprehensive retrofits and retro-commissioning
	PGE_Com_002	Laboratory Performance Efficiency Program (Smart Labs)	An NMEC-specific program targeting laboratories for ventilation, other retrofits, and BRO measures
	PGE_Com_003	Commercial Efficiency Program	Open to the entire commercial segment, offering site-level NMEC, population NMEC, custom, and deemed delivery channels
	PGE21011	Commercial Calculated Incentives	Open to the entire commercial segment, offering incentives for EE opportunities identified through utility-sponsored audits, facility/process assessments, or retro-commissioning studies.
	PGE210911	On-Bill Financing Alternative Pathway	Open to the entire commercial segment, offering on-bill financing without participation in another program
	PGE2110012	University of California/California State University	An institutional partnership between University of California/California State University and IOUs, providing performance-based incentives for lighting, controls, HVAC, new construction, and commissioning measures
SCE	SCE-13-L-003I	Public Sector Performance-Based Retrofit HOPPs	Targets public sector buildings with stranded savings resulting from improvement delays or indefinite equipment repairs
	SCE-13-SW-002B	Commercial Calculated Program	Open to the entire commercial segment, NMEC projects must include at least two distinct EE measures (including capital, BRO, to-code, and above code measures) affecting at least two distinct building systems
SoCalREN	SCR-PUBL-B3	Public Agency NMEC Program	Targets public sector buildings with stranded savings
SDG&E	SDGE4004	Commercial Large Customer Services (>20KW) Program	Open to large commercial customers on qualifying rates schedules with a monthly demand greater than 20 kW, providing end-to-end services, including marketing, outreach, engineering, operations, customer service, and data management and reporting
	SDGE4012	Federal Customer Services Program	Open to federal buildings, US Postal Service, military bases, and tribal nations on qualifying rates schedules, providing end-to-end services, including marketing, outreach, engineering, operations, customer service, and data management and reporting

3 METHODOLOGY

In determining and/or evaluating sample design, achieved sample sizes, gross savings, measurement and verification (M&V) activities, net savings approach, and final results expansion procedures, DNV followed International Performance Measurement and Verification Protocol (IPMVP) and the California Evaluation Protocol.

3.1 Sample designs

The gross population includes all projects for which savings have been finalized and the project has been trued-up in the PY2023 tracking data. The gross population is quite small compared to the number of sites that were expected to be trued-up in PY2023. DNV will review the sites that were expected to be trued-up in the PY2023 Site-level NMEC Additional Research Report later this year. Please see the PY2023 Site-level NMEC Evaluation, Measurement, and Verification Workplan for more information.[40]

Table 3-1. Gross sample coverage by PA

PA	Population (N)	Sample design quota	Final sample (n)	Percentage of sample complete
SCE	1	1	1	100%
SoCalREN	2	2	2	100%
Total	**3**	**3**	**3**	**100%**

The NTG population is larger and distinct from the gross population, as it includes projects with initial claims in PY2023. The NTG population is the expected gross population for the next evaluation. NTG evaluations hinge on project decision making, and by the time a project makes initial savings claims those decisions have been made.[41] It benefits our evaluation to complete NTG interviews as close to the time of decision making as possible, increasing the likelihood of accurate recall and reducing the likelihood of decision maker turnover.

We attempted a census of the NTG population. Table 3-2 shows the net sample design population, sample target, and sampled sites by PA. Overall, we interviewed participants covering 89% of projects.

Table 3-2. Net sample coverage by PA

PA	Population (N)	Sample design quota	Final sample (n)	Percentage of sample complete
PG&E	17	17	13	76%
SCE	25	25	24	96%
SDG&E	2	2	2	100%
SoCalREN	2	2	2	100%
Total	**46**	**46**	**41**	**89%**

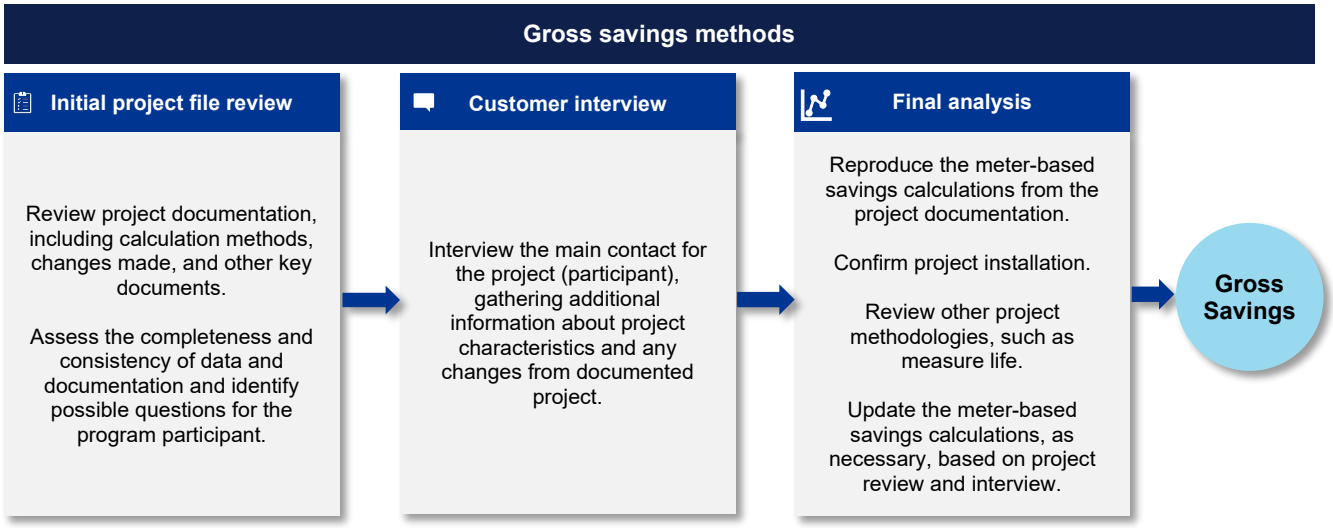
3.2 Gross savings methods

Figure 3-1 summarizes the key steps of the gross savings evaluation, which are further described in the following sections.

[40] CPUC, *Workplan, Program Year 2023*, December 12, 2024, https://pda.energydataweb.com/api/view/4101/PY2023%20CPUC%20Site-level%20NMEC%20Evaluation%20Workplan%20-%20Final.pdf.

[41] CPUC, *Energy Division Staff Guidance: NMEC Reporting*, April 24, 2020, https://www.cpuc.ca.gov/-/media/cpuc-website/divisions/energy-division/documents/energy-efficiency/rolling-portfolio-program-guidance/nmec-reporting-guidance_04242020.pdf.

Figure 3-1. Gross savings methodology



3.2.1 Initial project file review

During the initial review of project files, DNV used a modified version of the Custom Core Template (CCT) to validate project eligibility and several key project details, described below.

- **Installed measures**: We reviewed the measure documentation for completeness and determined which planned measures projects had, in fact, implemented and when they had installed those measures. Projects did not install all planned measures. In other cases, the scope of a project changed between planning and implementation. These kinds of discrepancies could affect savings and required further validation during participant interviews.
- **Measure-level measure application type (MAT)**: Because MATs are important for determining effective useful life (EUL), we reviewed PAs' MAT assignments, identified documentation that supported these assignments, and prepared questions to confirm them, particularly when attempting to verify whether a measure was accelerated by the program.
- **Measure-level EUL**: Measure-level EUL is the basis of the savings-weighted project EUL used to calculate a project's lifetime savings. We reviewed the measure-level EULs provided in the documentation and investigated the sources of those EULs.
- **Engineering-based savings estimates**: We confirmed the presence and general reasonableness of the provided engineering-based savings estimates. In cases where projects had multiple EULs, we examined the engineering-based savings estimates more closely, as they are used to calculate the savings-weighted EUL.
- **Project dates:** We determined key project dates such as project implementation start and end and the dates of identified non-routine events (NREs). These dates are important for identifying any overlap between installation and the baseline or performance period models and for addressing any NREs. We flagged dates to collect or confirm with participants during interviews.
- **Non-IOU fuel sources:** We reviewed project documentation for onsite generation, as it could impact NMEC model results if the generation source is on the same meter as the participating building. Additionally, savings must be less than the energy imported from the grid.
- **Non-routine events (NREs):** We reviewed identified NREs and identified potential additional NREs by examining other site activities (e.g., space repurpose) and energy use plots.
- **Project Review dispositions:** For projects that went through CPUC Project Review, we reviewed the relevant dispositions.
- **Other:** We also looked for other less common situations, such as fuel switching and Early Opinions.

3.2.2 Customer interview

Our customer interviews aimed to confirm the installation of proposed measures, the operation of those measures, key project dates, the existence of onsite generation, NREs, facility operation changes, occupancy patterns, and the effects of COVID-19. For replacement measures, we collected information about the condition of pre-existing equipment and program influence in order to evaluate the MAT. When necessary, we followed up with additional data requests of the participant and the PA.

3.2.3 Final analysis

Our final analysis included an engineering and policy review and a model review. During the engineering and policy review, we modified the projects based on the customer interview and additional data provided.

Our evaluation model review included both model replication and model validation. During model replication, we reproduced the models and savings results using the code, spreadsheets, or other tools provided in the project documentation. During model validation, on the other hand, we independently reproduced the models and savings results outside the provided documentation, modifying the models as necessary to more closely align with best practices and CPUC guidance.

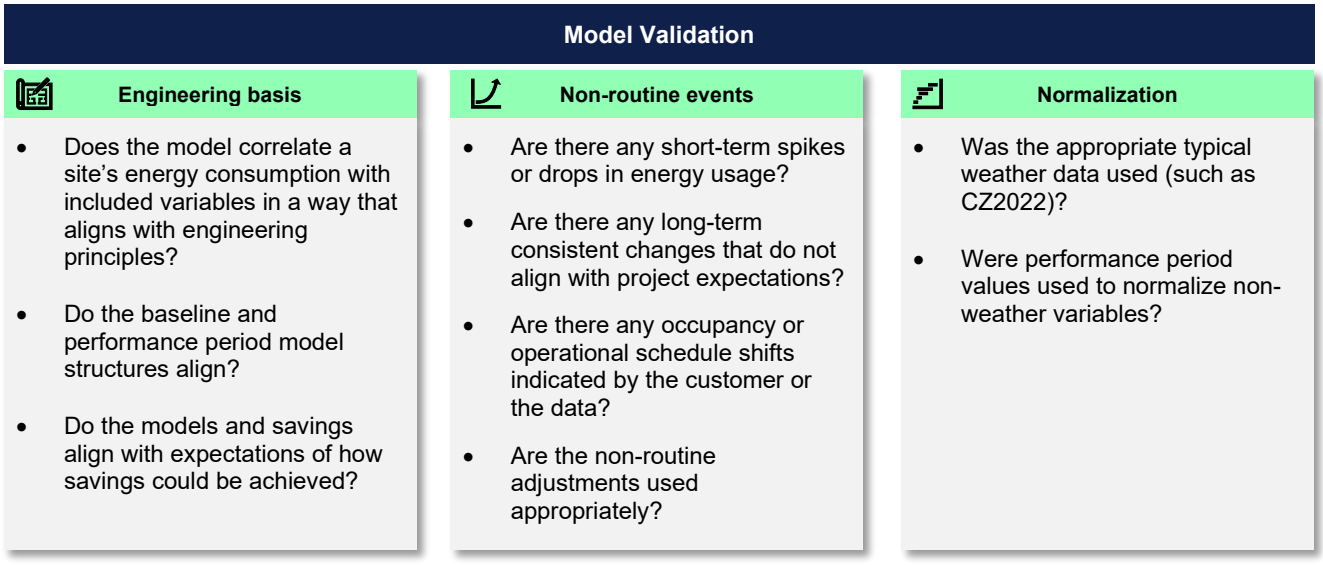
3.2.3.1 Model replication

For each project we replicated the baseline model, the performance model, the normalization of baseline and performance consumption, and the calculation of normalized savings. We considered a replication to be successful if it achieved the same goodness of fit statistics for both baseline and performance models, fractional savings uncertainty (FSU) for the baseline model, total normalized savings, normalized savings fraction, and normalized FSU.

3.2.3.2 Model validation

The purpose of validation was to identify discrepancies between the way the PAs modeled their data and standard modeling practices and CPUC guidelines. Such discrepancies could produce biased estimates of savings. For each change we made during validation, we documented the discrepancy and estimated the savings impact.

Figure 3-2. Model validation



Engineering basis: We reviewed for appropriateness all dependent and independent variables in the model. At minimum we expected each model to use consumption as the dependent variable and actual outdoor temperature as an independent variable. For any other independent variables, we considered whether the PA provided sufficient justification for including

those variables and whether, according to engineering principles, the site's energy consumption would be expected to correlate with those additional variables. We also checked that both baseline and performance models included the same dependent and independent variables, as this is crucial for valid model comparison and savings estimation. Finally, we considered if any essential variables or additional data were not included in the models. Where possible, we requested missing data from the participant.

We assessed how the model parameter estimates changed from baseline to performance periods. Parameter estimates represent how building consumption changes with the associated independent variable and should be consistent with engineering principles. For example, if an installed measure would reduce the effect outdoor temperature has on consumption, we would expect the performance model's temperature parameter estimates to be smaller and of less statistical significance than those of the baseline model. Any lack of alignment between parameter estimates and engineering-based outcomes suggests that those parameters are capturing some unknown influence on energy consumption.

Non-routine events: We assessed error and model fit plots to investigate the presence of NREs. NREs represent abnormal changes in building consumption that can severely bias models that do not properly account for them. We looked for short-term spikes or drops, long-term consistent changes, and other trends in energy consumption. For those NREs we did find, we confirmed that the PA properly accounted for them, which is generally done by removing from the model any data that coincides with short-term NREs or including in the model indicator variables during long-term NREs. We also checked whether NREs were appropriately accounted for during normalization.

Normalization: We assessed whether projects followed standard guidance during normalization. For temperature normalization, we confirmed that projects used an appropriate typical meteorological year (TMY) data set. We also checked that the bounds of temperature values in the chosen TMY data did not exceed 10% of the temperature bounds from either the baseline or performance period model. For other variables, standard practice is to use performance period values if the bounds of the baseline model's values do not exceed 10% of the bounds of the performance period. In cases where this does not hold, it is acceptable to use the baseline period's data. Any other set of values is considered non-standard and requires justification.

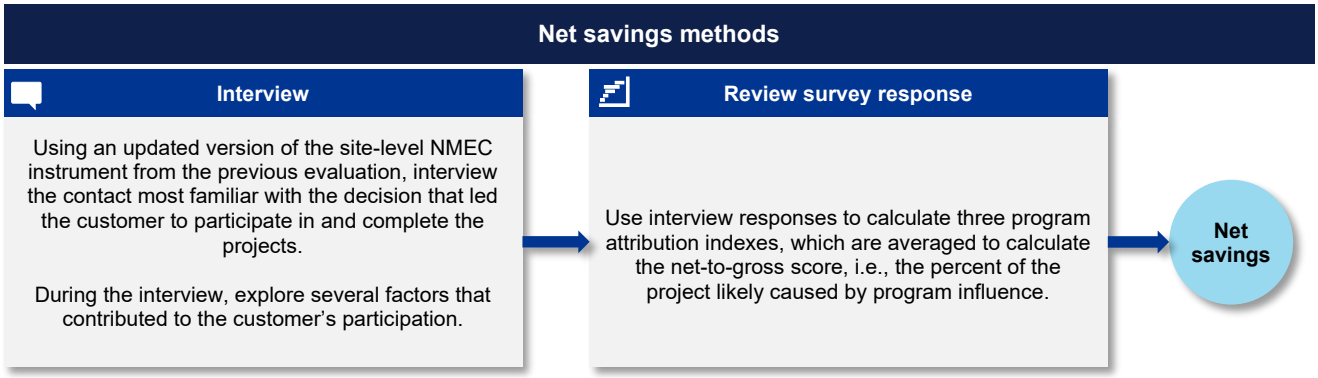
3.2.4 COVID-19 impacts

All three of the projects included in the gross evaluation were impacted by COVID-19. In the wake of the pandemic a library's schedule changed, a performing arts center was completely shut down for a few months and then opened with reduced hours, and an office building was closed for a few months and then reopened with reduced hours. To address these issues, one project used an occupancy variable and two used indicator variables to produce reasonable normalized savings estimates. Recognizing the difficulties faced by meter-based programs during a pandemic, our evaluation focused on the reasonableness of efforts to account for COVID impacts. We expect, in future PYs, that these COVID-related adjustments to NMEC models will continue to decline as the impacts of COVID recede.

3.3 Net savings methods

Figure 3-3 summarizes the key steps of the net savings evaluation, which are further described in the following sections.

Figure 3-3. Net savings methods



3.3.1 Customer interview

DNV interviewers completed NTG in-depth interviews (IDIs) with participant decision makers for projects with initial claims in PY2023. Projects with initial claims in PY2023 were only included in the NTG sample.

We used the IDIs to calculate three program attribution indexes (PAI_1, PAI_2, and PAI_3), capturing program and non-program influences, participants' prior plans, and project timing. IDIs also included questions on project scope, program processes, program satisfaction, and firmographics.

3.3.2 PAI methods

Table 3-3 presents the methodology used to calculate each PAI, which reflects an adjustment to the NMEC evaluation approach of PY2020 – 2022.[42] While building on a well-established approach that has been used for nearly a decade to evaluate commercial programs, the adjustments to the NTGR scoring algorithm better reflect NMEC program influences and reduce overlap among the PAIs, countering the previous methodology's tendency to push NTGRs to 0.5.

[42] For a full description of the NTG methodology update, see the *Group D NTG Methodology Update* memorandum provided to the CPUC by DNV on January 7, 2025.

Table 3-3. NTGR scoring methodology

Score	Description	Calculation
Program attribution index 1 (PAI_1)	• PAI_1 reflects two different ways of measuring the relative influence of program and non-program factors on project decision making. • Respondents rate each program influence factor and nonprogram influence factor using a 0-to-10 scale, where 0 means "not at all important" and 10 means "extremely important." PIF is the maximum score given to any program influence. NPIF is the maximum score given to any non-program influence. • Respondents also divvy up ten points between their collective program factors and their collective non-program factors. • PIP is the number of points given to the collective program factors. NPIP is the number of points given to the collective non-program factors.	$PAI_1 = 10 * \frac{PIF * PIP}{((PIF * PIP) + (NPIF * NPIP))}$
Program attribution index 2 (PAI_2)	• PAI_2 reflects the prior plans of a participant, namely if a site had plans in place and budget set aside for the capital project prior to interacting with the NMEC program. • If the respondent says they made the decision to do a project after interacting with the program, PAI_2 is set to 10. • If they say they made the decision before interacting with the program, PAI_2 is set to 0. • If they give a mixed response, PAI_2 is set to 5.	Decision made **after** program contact: $PAI_2 = 10$ Decision made **before** program contact: $PAI_2 = 0$ Mixed response: $PAI_2 = 5$
Program attribution index 3 (PAI_3)	• PAI_3 reflects program influence on project scope and timing. • The score is calculated by subtracting from 10 the respondent's ranked likelihood that they would have installed a project of the same scope in the same time frame even if the program had not been available. • S is the respondent's ranked likelihood that they would have installed a project *of the same scope* even if the program had not been available. • T is the respondent's ranked likelihood that they would have installed a project *in the same time frame* even if the program had not been available. • PAI_3 is calculated with the greater of S or T.	$PAI_3 = 10 - Max(S, T)$
Customer-level net-to-gross ratio (NTGR)	• The NTGR is calculated as the average of the three program attribution index scores.	$NTGR_{cust} = \frac{Average(PAI_1, PAI_2, PAI_3)}{10}$

Table 3-4 lists the four influences we asked respondents to rate, with each categorized as either a program or a non-program influence. These ratings are used to calculate PAI_1. If respondents rated "Company policies or mandates" five or greater, we asked a follow-up question to identify which specific policies influenced their decision. If respondents indicated that corporate sustainability or environmental policy was the main influence, this influence was discarded so as not to penalize companies for having environmental goals.

Table 3-4. PAI_1 influences

Influence Type	Influence Factor
Program	**Financial incentives** (e.g., financing, rebates, or performance payments)
Program	**Technical assistance** (e.g., program-provided feasibility study or facility or system energy audit)
Non-program	**Company policies or mandates** (e.g., maintenance/replacement schedules or regulatory requirements; note that this excludes corporate sustainability goals)
Non-program	**Other payback/return on the project** (excluding program incentives, e.g., utility bill savings)

3.3.3 Example PAI calculations

Figure 3-4 shows three hypothetical examples of PAI_1 calculations for illustrative purposes. The first example shows a case where a hypothetical respondent considered at least one program influence and one non-program influence as extremely important in their decision-making process, providing a PIF and NPIF of 10. The respondent viewed all program influences collectively as equally influential as all non-program influences collectively on their decision-making process (providing a PIP of 5 and an NPIP of 5). This results in a PAI_1 of 5, indicating equal program and non-program influence. In Example 2 the respondent viewed at least one program influence and one non-program influence as extremely important in their decision-making process, just as in Example 1. However, in Example 2 the respondent viewed the importance of all program influences collectively as more important than the importance of all non-program influences collectively, providing a PIP of 8 and NPIP of 2. The result is a PAI_1 of 8, indicating high program influence. In Example 3, the respondent views one non-program influence as the most important factor in their decision-making process, providing an NPIF of 10, while giving their most important program influence only a 7. This respondent also views the importance of all non-program influences collectively as more important than all program influences collectively, providing an NPIP of 6 and PIP of 4. Example 3 results in a PAI_1 of 3.1, indicating low program influence.

Figure 3-4. PAI_1 calculation examples

Example 1: Equal importance

PIF = 10
NPIF = 10
PIP = 5
NPIP = 5

$$10 * \frac{10 * 5}{((10 * 5) + (10 * 5))}$$

PAI_1 = 5.0

Example 2: Higher collective program importance

PIF = 10
NPIF = 10
PIP = 8
NPIP = 2

$$10 * \frac{10 * 8}{((10 * 8) + (10 * 2))}$$

PAI_1 = 8.0

Example 3: Higher individual and collective non-program importance

PIF = 7
NPIF = 10
PIP = 4
NPIP = 6

$$10 * \frac{7 * 4}{((7 * 4) + (10 * 6))}$$

PAI_1 = 3.1

3.3.4 Adjustments to NTGR scoring

As mentioned above, at the end of 2024, as part of a larger effort to update all evaluation NTG methodologies, the CPUC adopted updates to the site-level NMEC NTG methodology to better align with NMEC program goals and delivery. The updates revised both the NTG interview instrument and the scoring algorithm. The following key changes were made to the instrument.

- **Improved flow:** The revised NMEC instrument asks respondents all questions relating to a specific topic (e.g., decision making timing) at once rather than mixed throughout the interview. The new instrument flows better and feels more cohesive.
- **Edited list of PAI-1 influences that better aligns with program offerings and goals:** Program Attribution Index 1 (PAI-1) of the NTG instrument was previously based on a list of about 20 influences for which respondents provided individual importance ratings. The list was originally developed for CIAC and thus some influences were not applicable to or aligned with program goals and offerings or NMEC. For example, the CIAC instrument treated "the age or condition of the old equipment" as a non-program influence, which would reduce the NTGR if a respondent indicated that was an important influence in their decision to do the project. The NMEC program, however, is designed to capture savings from old equipment that customers would have kept repairing rather than replacing; it is not appropriate to penalize an NMEC program just because updated equipment was old. "The age or condition of the old equipment" was removed from the NMEC instrument, along with several other influences that similarly did not apply. Streamlining the list of influences also reduces respondent burden, making it easier for them to remain engaged throughout the survey.
- **Added follow up questions to not discourage corporate sustainability policies:** The revised instrument includes a new follow-up question to probe customer's corporate sustainability policies, ensuring that NTGRs are not reduced because of a beneficial customer policy. Specifically, when respondents provide a rating of five or more to the non-program influence "Corporate policies or guidelines," they are then asked, "What specific corporate policy influenced your decision to [adopt/install/implement] [PROJECT]?" If they indicate the policy is a corporate sustainability policy, the influence rating will not reduce the NTGR.

One key update to the scoring algorithm reduced overlap between PAIs. In the original NTG algorithm PAI-1 and PAI-2 both incorporated customer's ratings of program and non-program influences on the decision-making process for the project, preventing customers' perceptions on individual influences and on overall influences from being considered together. PAI-2 also included an input relating to the timing of the decision to implement a project relative to program interaction. The updated method used in this study revised PAI-1 and PAI-2 to reduce the tendency of the previous algorithm to push NTGRs to 0.5. The update removes overlap between the two indexes so that PAI-1 is only based on customer reported influence ratings and PAI-2 is only based on the timing of the decision to implement a project relative to program interaction.

4 RESULTS

This section presents findings related to gross and net savings by key reporting dimensions. It includes a discussion of the differences between gross savings claims and evaluated results. In addition, we provide an examination of the drivers of the NTGR, which measures the program's influence on decisions to implement efficiency measures.

4.1 Gross electricity savings and realization rates

Table 4-1 presents the gross electricity realization rates and savings results by PA. Claimed savings are the savings claimed by the PAs in the CEDARS tracking database. Documented savings are the savings tabulated in project final savings reports. As the claimed savings often diverged from the documented savings due to data entry errors, the table compares savings from both the tracking database and project documentation. Verified savings are the savings resulting from this evaluation. The gross realization rate (GRR) compares the verified savings with the savings claimed in the tracking data (Claimed). The documented realization rate (DRR) compares the savings verified through the evaluation with the savings provided in the project documentation.

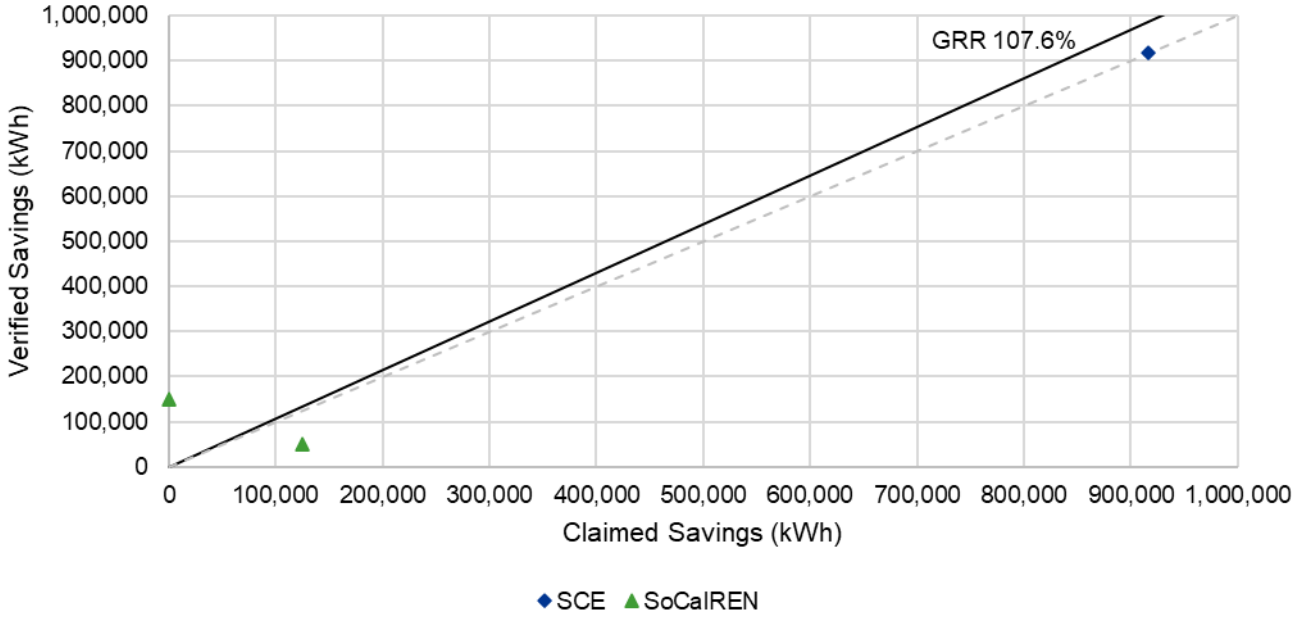
Table 4-1. Gross electricity savings

Program administrator	Projects	Claimed (kWh)	Documented (kWh)	Verified (kWh)	GRR	DRR
First-year savings						
SCE	1	915,788	915,788	916,970	100.1%	100.1%
SoCalREN	2	124,473	280,888	202,349	162.6%	72.0%
Statewide	**3**	**1,040,261**	**1,196,676**	**1,119,319**	**107.6%**	**93.5%**
Lifecycle savings						
SCE	1	10,806,303	10,806,298	11,095,336	102.7%	102.7%
SoCalREN	2	1,730,175	3,763,570	3,301,463	190.8%	87.7%
Statewide	**3**	**12,536,478**	**14,569,868**	**14,396,799**	**114.8%**	**98.8%**

It is important to note that these results reflect only the three site-level NMEC projects trued-up in PY2023. While these projects provide some important insight as case studies into how projects are documented and claimed, the small sample size limits how far the results can be generalized. SCE's GRR and DRR are relatively similar because the final savings claims (sum of initial and true-up claims) mostly matched the documented savings. SoCalREN's GRR and DRR are further apart because one of their two projects submitted two true-up claims, over-adjusting the initial claim and zeroing out the first-year kWh savings. Overall, both the GRR and DRR are higher than in the PY2020 – 2022 evaluation, which found a first-year savings GRR of 70.9% and a DRR of 81.2% for kWh. In the PY2020 – 2022 evaluation, 15 out of 22 projects (68%) had tracking data errors, whereas one out of three (33%) had tracking data issues in this evaluation. As with the PY2020 – 2022 projects, the three projects evaluated in this report were also disrupted by the COVID-19 pandemic, which continued to make modeling more challenging.

Figure 4-1 compares the claimed savings and verified savings. The diagonal dashed line indicates where each sample point would be plotted if the project realized 100% of the claimed savings. The points below the dashed line achieved less verified savings than claimed savings, while the points above the dashed line achieved greater verified savings than claimed savings. Even with the small sample size of three projects, none of the projects had 100% realization rates. One project mistakenly claimed zero savings when summing up the initial claim and true-up claims, resulting in a large increase in savings. Another project had a realization rate of 41.2%, and the last project had a realization rate of 100.1%.

Figure 4-1. First-year electric energy savings scatterplot

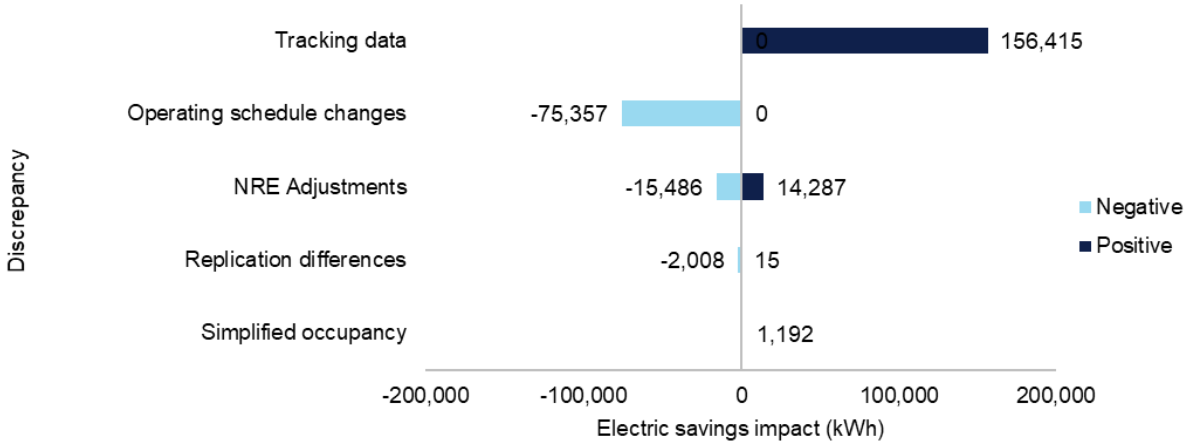


The following sub-sections present an analysis of the discrepancies between claimed and verified first year gross savings for the sampled projects. Table 4-2 summarizes the types of discrepancies identified. Figure 4-5 shows the savings impacts of each type of discrepancy.

Table 4-2. Savings discrepancy factors

Discrepancy factor	Description
Tracking data	Differences attributed to incorrect adjustments or unexplained changes to savings that occurred between completion of the analysis and entry into the PA tracking system
Replication differences	The difference between the savings replicated using the provided code and project data and those reported in the final reports or output files
Operating schedule changes	Changes related to aligning baseline and performance period operating hours
Non-routine event (NRE) adjustments	Changes from excluding periods where the data may be skewed due to abnormal facility operations, disruptions, or non-operation, ensuring accurate estimates of project savings
Simplified occupancy temperature relationship	Changes due to removing a separate temperature relationship when an occupancy proxy variable is already included in the model or to make baseline and performance models consistent

Figure 4-2. Summary of first year kWh savings discrepancy factors by sum of savings impact



4.1.1 Tracking data

Consistent with the PY2020 – 2022 evaluation, the largest savings discrepancy in PY2023 is due to differences between tracking data and project-specific final savings reports. While two out of the three projects we evaluated had tracking data claims that matched the project documentation, the remaining project had a large discrepancy between tracked and documented savings. This project was trued-up twice, which ultimately zeroed out the initial claimed savings. The adjustment from zero to the project's documented savings resulted in the largest savings discrepancy identified for the three evaluated projects.

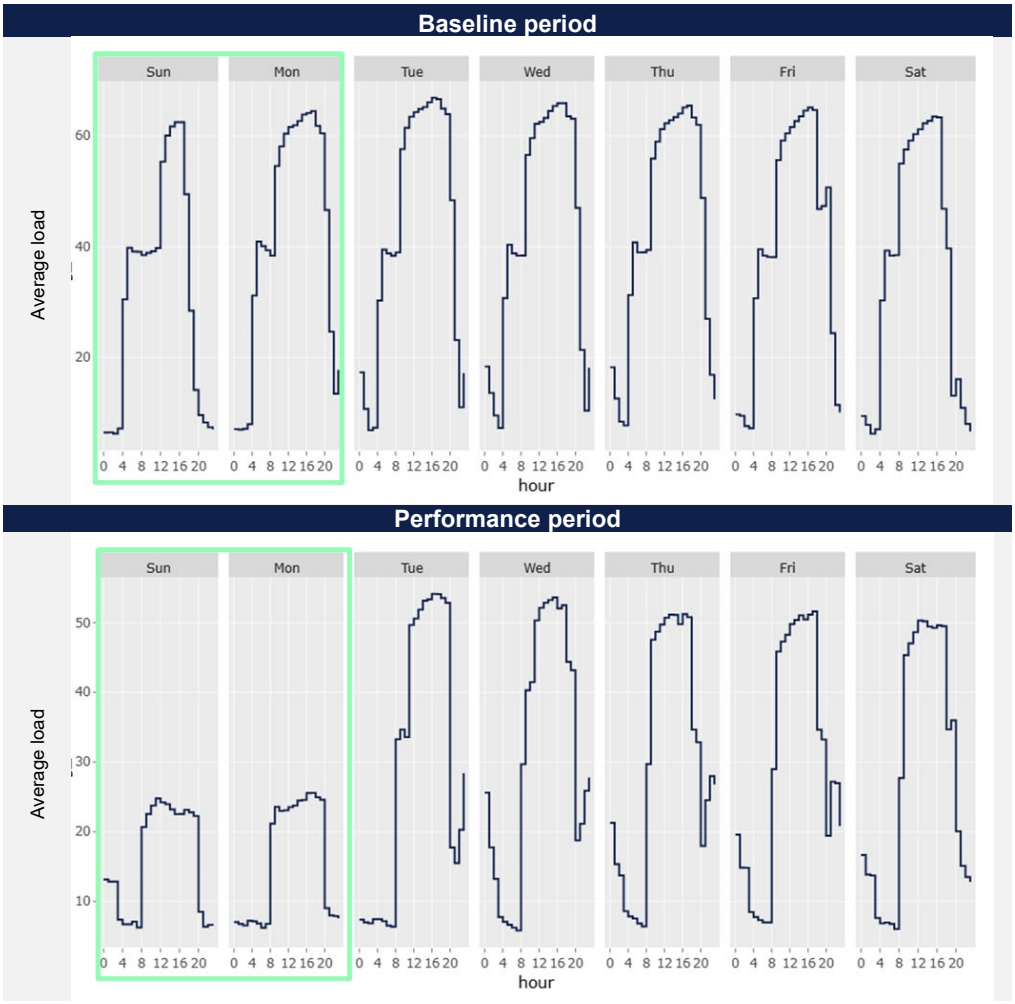
4.1.2 Operating schedule changes

A building's operating schedule can have a substantial impact on its energy use. In some cases, a model may normalize building schedules that differ from the baseline period to the performance period. However, for two of the evaluated projects, the operating schedule shifted in a way that the models could not address.

A library that had been open seven days a week prior to COVID-19 (during the project's baseline period) opened with reduced hours just five days a week following lockdowns. Since the library had been open every day of the week during the baseline period, there were no comparable days for the two days a week in the performance period during which the library was closed. In this case the PA did not account for changes in the operating schedule, causing the model to overestimate savings by incorrectly attributing energy reductions from reduced operating hours to the program.

Figure 4-3 shows the baseline and performance period average daily load shapes. For performance period weekdays, we reduced the baseline period's non-firm load (energy consumption that varies based on occupancy, weather, and equipment use) by 20% to account for reduced consumption, reflecting the 20% fewer hours the facility was open in the performance period compared to the baseline. The two green boxes highlight the marked change in shape between the baseline and performance periods for weekends. Given the available information, DNV had two options for addressing changes in weekend consumption: either set savings to zero, as no realistic baseline consumption for the closed site existed or use baseline period Sunday morning consumption levels as a proxy for typical closed weekend consumption. This second approach compared baseline levels (approximately 40 kW) with performance period levels (approximately 28 kW) across the 12 hours of weekend partial consumption. The difference represents 12% of the claimed savings. Recognizing that energy savings occurred during the weekend, the evaluation assigned non-zero savings only for the 12 hours with partial load in the performance period.

Figure 4-3. Library average daily load shapes



4.1.3 Replication differences

While we were not able to exactly replicate the savings results for any of the three evaluated projects, two out of the three projects were within 15 kWh (less than 0.01% of normalized baseline consumption), which is relatively close. One project had a larger discrepancy of 2,000 kWh (approximately 0.2% of normalized baseline consumption). For this project, the provided script did not address COVID-19 or NRE-related independent model variables when normalizing baseline and performance period energy consumption, which accounted for the discrepancy. While most of these discrepancies are relatively small compared to the total project savings, NMEC project documentation should exactly reproduce the savings in the project documentation and the CEDARs tracking database. Replication is the necessary first step of any subsequent validation process.

4.1.4 Non-routine event adjustments

We adjusted the way in which NREs were handled for two of the three projects.

The library had three NRE periods across the baseline and performance periods. In addition to the COVID-related shutdown for approximately five weeks during the baseline period, there were two unexplained 10-week periods of evident energy use increases in the spring and fall of the performance period. The project documentation did not explain what caused these NREs, and we were unable to identify a cause when talking with the customer. Due to the presence of COVID, we accepted

the identification of these NREs despite those 20 weeks representing roughly 40% of the performance period data. Given the extensive length of these periods, we applied a different non-routine adjustment approach, excluding data from the NRE periods altogether when modeling the impact, rather than capturing only a baseload effect with a dummy variable. These changes led to a minimal increase of 1% in the estimated impact.

The office building had a COVID-related NRE when the office was shut down for the first five months of the performance period. The PA addressed this NRE with the use of an indicator variable. When we reviewed the energy usage data for the site, we determined that the NRE period lasted for two additional weeks beyond the original non-routine adjustment (NRA). This discrepancy resulted in an 11% decrease relative to reported savings, but savings were higher than the zero tracking claimed savings.

4.1.5 Simplified occupancy temperature relationship

Models for all three projects used a daily time of week and temperature (TOWT) approach to quantify electric savings. The hourly TOWT model can distinguish between occupied and unoccupied hours by identifying hours with usage significantly below the expected mean for the hour. The occupied and unoccupied assignments are used to create two distinct models to allow for different levels of weather dependence. For instance, a building might set back the thermostat setpoint in the evenings and weekends. At the hourly level, we would see different responses during these unoccupied periods than during occupied periods. At the daily level, we might expect different temperature correlation on the weekends relative to weekdays, for example, but not for a single weekday relative to other weekdays.

For the three evaluated projects we observed inaccurate assignments when the algorithm tried to classify whole days, rather than hours, as occupied or unoccupied. For one project in which the building was in use five days of the week, the model classified only Wednesdays as having a different weather dynamic. For the three projects we evaluated removing the additional occupancy-related temperature spline made relatively little difference to the overall savings, but we did it to make sure that the model aligned with the way the building was being used.

4.2 Gross demand savings and realization rates

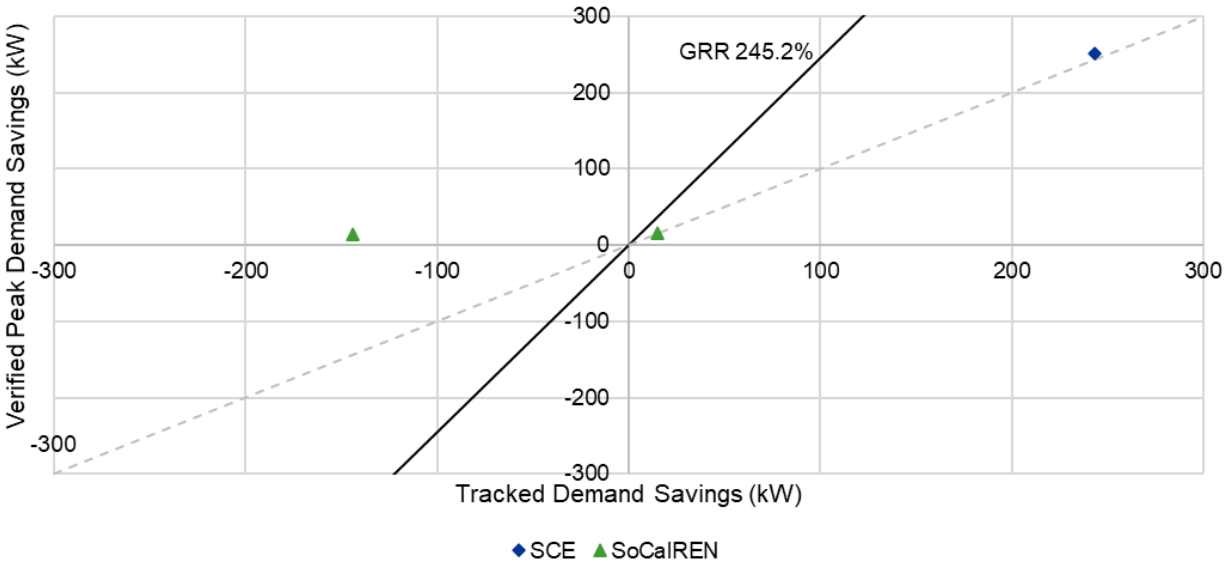
In contrast to the PY2020 – 2022 evaluation, all of the PY2023 projects that claimed electric savings also claimed peak demand savings. Table 4-3 presents the gross demand realization rates and savings results by PA. As with the electric energy savings, we have provided GRRs and DRRs, as the claimed savings diverged from the documented savings due to data entry errors for two of the three projects.

Table 4-3. Gross peak demand (kW) savings

Program administrator	Projects	Claimed	Documented	Verified	GRR	DRR
SCE	1	243.2	243.0	250.8	103.1%	103.2%
SoCalREN	2	-128.9	35.8	29.3	-22.7%	82.0%
Statewide	**3**	**114.2**	**278.8**	**280.1**	**245.2%**	**100.5%**

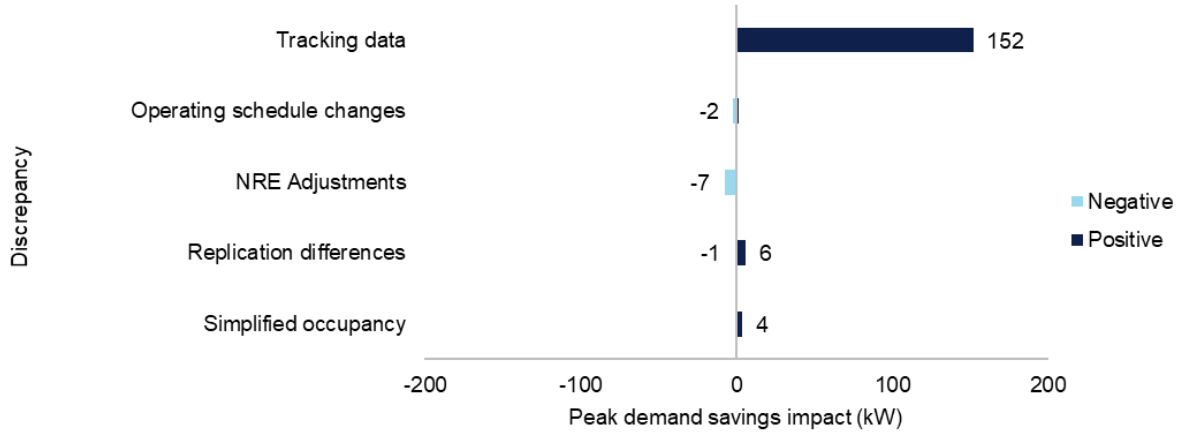
Figure 4-4 compares the claimed savings to weighted verified savings. The diagonal dashed line indicates where each sample point would be plotted if the project realized 100% of the claimed savings. The points below the dashed line achieved less verified savings than claimed savings, while the points above the dashed line achieved greater verified savings than claimed savings. Two of the three projects had realization rates of 103%, very close to the dashed line. The other project claimed negative demand savings due to a data entry error in which the project was accidentally trued-up twice but achieved positive verified savings.

Figure 4-4. Peak demand savings scatterplot



All identified demand discrepancies fell into the same categories as the electric energy discrepancies (see Table 4-2). Figure 4-5 shows the savings impacts of each type of demand discrepancy.

Figure 4-5. Summary of first-year peak demand savings discrepancy factors by project count and savings



The largest source of demand discrepancy is tracking data, mirroring our electric energy savings evaluation and driven largely by the accidental reporting described above wherein one project trued-up twice. The other discrepancies were due to the same issues described in the electricity savings section above.

4.3 Net savings results and ratios

Table 4-4 shows the net-to-gross ratios (NTGRs) for electric energy savings (76.6%), demand savings (75.2%), and natural gas savings (76.0%). In this table, “Claimed” means tracked savings. The NTGRs are higher than in last year’s evaluation, which found NTGRs of 45.9%, 41.7%, and 46.5% respectively. The increase in NTGRs from the previous evaluation to this one is the result of both actual changes in program influence as reported by this year’s sample compared to last year’s sample and updates to the scoring methodology.

Increases in reported program influence: In last year's evaluation respondents gave an average of 5.1 out of ten points to program influences, indicating that program and non-program influences were equally important in their organizations' decisions to do the EE projects. This year respondents gave an average of 7.5 points to the program influences, indicating that program influences were more important than non-program influences. One respondent said of program staff, "They offer an extension of our staffing. I don't have time to go and do research on what program are available and what aligns most with our school district's needs. So having ongoing meetings with them to touch base gives me an extension of my capacity."

Decision timing: Additionally, no respondent in the current evaluation indicated that their organization had decided to do the project before interacting with the program, compared to 40% of respondents last year. Similarly, in response to a new question in this year's survey instrument, 80% of respondents indicated that their organization first learned about the opportunities included in their EE projects from program or utility staff. One respondent said, "[The biggest strength of the program was the] awareness that those audits brought to us. The audits bring a lot of issues to our attention which are hard for us to recognize ourselves."

Methodology update: To assess the impact of our updated methodology on NTGR estimates, we also calculated NTGRs for this year's sample using the previous methodology. Table 4-5 compares the NTGRs using the current and the previous methodologies. Using the previous methodology, NTGRs for this year's sample range from 60.0% to 62.8%, accounting for some, but not all, of the increase in NTGRs from the previous evaluation. The previous methodology was based on that used for CIAC, and the changes better aligned the NTG methodology with NMEC program goals and delivery. See section 3.3 for more details on the changes made to the NTG methodology.

Table 4-4. Net savings results by PA

Program administrator	Projects	First-year net savings				Lifecycle net savings	
		Claimed	Net	NTGR	RP%*	Net	RP%*
Energy (kWh)							
PG&E	12	12,834,707	9,673,519	75.4%	±7.0%	78,347,649	±4.0%
SCE	24	2,070,845	2,058,834	99.4%	±0.0%	25,976,314	±0.0%
SoCalREN	2	874,782	558,986	63.9%	±0.0%	1,931,105	±0.0%
SDG&E	2	816,100	413,600	50.7%	±0.0%	4,963,196	±0.0%
Statewide	**40**	**15,780,335**	**12,291,339**	**76.6%**	**±5.0%**	**106,255,068**	**±3.0%**
Demand (kW)							
PG&E	12	1,517	1,167	76.9%	±10.0%		
SCE	18	147	147	100.0%	±0.0%		
SoCalREN	1	53	45	86.1%	±0.0%	NA	
SDG&E	2	265	134	50.7%	±0.0%		
Statewide	**33**	**1,982**	**1,494**	**75.2%**	**±8.0%**		
Natural gas (therms)							
PG&E	5	321,168	244,120	76.0%	±7.0%	1,626,209	±4.0%

Table 4-5. NTGR methodology and sample comparison

NTGR	PY2020-2022 evaluation previous methodology	PY2023 evaluation previous methodology	PY2023 evaluation current methodology
Statewide electricity	45.9%	61.3%	76.6%
Statewide demand	41.7%	61.0%	75.2%
Statewide natural gas	46.5%	60.2%	76.0%

4.3.1 Program attribution index results

As discussed in section 3.3, we calculated NTGRs using an approach that differs slightly from the NMEC evaluation of PY2020 – 2022, better reflecting NMEC program influences, reducing overlap among the program attribution indexes, and countering the previous methodology's tendency to push NTGRs to 0.5.

Table 4-6 describes the three program attribution indexes (PAI_1, PAI_2, and PAI_3) and shows the average score for each indicator.

Table 4-6. Program attribution index (PAI) results

Program attribution index	Basis	Average
PAI_1	Respondents' ratings on the importance of individual program and non-program influences in their decision to implement a project	7.7
PAI_2	Respondents' rating on the timing of project implementation relative to program interaction	9.7
PAI_3	Respondents' ratings for the likelihood they would have implemented a similar project scope on a similar timeline in the absence of the program	7.9

We calculated the NTGR by averaging the three PAIs, resulting in a NTG score primarily based on influence ratings rather than the more direct timing and scope change measurements. Detailed PAI results are included in APPENDIX A.

4.4 Participant satisfaction and program feedback

In addition to net-to-gross questions, the survey asked questions about program satisfaction, program strengths, and suggested areas for improvement.

Interviewers asked respondents, "Who first brought this project to your organization for consideration?" Table 4-7 summarizes respondents' answers. Eighty percent of respondents, representing 95% of sites, indicated that either program or utility staff first made them aware of the opportunities included in the project. Respondents reported learning about projects from audits, presentations by program staff, and being approached by program vendors with recommendations. One respondent said they did not know who brought the specific project to their organization's consideration first, but indicated they rely on program support to research all efficiency upgrades. They said, "We have an ongoing relationship with [the program staff]. They propose things and other times if we're curious about things we might ask about it. For [this specific project] I'm not sure what happened first. But even if we have an idea [on our own] we look for their research on the project first before we can move forward."

Table 4-7. Who first brought project to respondents' organizations' consideration

Who first brought this project to your company's consideration?	Percent of respondents	Percent of sites
Program staff	60%	78%
Utility staff	20%	17%
Internal staff	10%	2%
Don't know	10%	2%

Respondents indicated a high level of satisfaction with the program, driven by the programs' technical support and incentives. When asked, "On a scale of 0 – 10, where 0 is 'completely dissatisfied' and 10 is 'completely satisfied', how would you rate your overall satisfaction with the [program]?" respondents gave an average rating above 8 (Table 4-8). Eighty percent of respondents, representing 95% of sites, were "promoters," providing a rating of 8 or above. In an open-ended question about the strengths of the program, respondents indicated that the technical support and incentives provided by the program drove their satisfaction (Table 4-9). These results are nearly identical to the previous evaluation. No respondents provided a satisfaction rating lower than 7 and thus there were no detractors.

Three respondents elaborated on the value of technical support provided by the program:

- "Some of the stuff I wouldn't be able to tell my board what the benefits are without the technical support."
- "[The strength is that the program allows us] to upgrade equipment and make our facilities more energy efficient and sustainable without having to put up much money upfront to accomplish the work."
- "[The strength was the] awareness that those audits brought to us. The audits bring a lot of issues to our attention which is hard for us to recognize because we have so many stores."
- "They offer an extension of our staffing. I don't have time to go and do research on what program are available and what aligns most with our school district's needs. So having ongoing meetings with them to touch base gives me an extension of my capacity."

Table 4-8. Program satisfaction

Metric	By participants	By sites
Average satisfaction	8.4	8.7
% promoters (≥8)	80%	95%
% detractors (≤3)	0%	0%

Table 4-9. Program strengths

Strength	Percent of respondents	Percent of sites
Technical support	50%	70%
Incentives	70%	27%
Energy audits	10%	19%
Helps achieve energy efficiency	10%	9%
Capacity building	20%	5%

While almost one-third of respondents had no suggestions for improvements to the program, those that did have suggestions most frequently wanted a more streamlined process. Twenty percent of respondents, representing 47% of sites, suggested streamlining program participation (Table 4-10). One said, "For the NMEC program, one of the difficulties was the amount of back and forth with technical reviewers. In hindsight, it was a lot of work on our end and we would have preferred to let a consultant or vendor deal with that." Twenty percent of participants, representing 5% of sites, also suggested improving communication about expectations and requirements. One said, "Communicating and understanding timelines

from the onset would be beneficial for both sides. What timelines does the program require and what can we accomplish it in that timeline based on our own restrictions?"

Table 4-10. Suggestions for program improvements

Improvement	Percent of participants	Percent of sites
Streamline process	20%	47%
Improve communication of expectations and timelines	20%	5%
Allow for flexibility in changes in building processes over time	10%	19%
Reduce need for costs upfront	10%	2%
Do not blame subcontractors for confusion	10%	2%
Keep programs running despite new federal administration	10%	2%
No suggestions	30%	22%

4.5 Measure application type (MAT) for NMEC

Measure application types (MATs)[43] are an energy efficiency categorization of the installed equipment and the condition of pre-existing systems. For custom projects MATs are required for all measures, and projects use them to calculate effective useful life (EUL) and the baseline against which savings are estimated. Site-level NMEC projects also require MATs, but, in contrast with custom projects, only use MATs to determine the appropriate measure life—not the appropriate baseline. While custom projects use different baselines depending on the MAT, all site-level NMEC projects use an existing condition baseline to estimate savings due to the performance-based approach.

Site-level NMEC projects assign MATs to individual measures to inform the appropriate measure-level EUL. Using measure-level engineering savings forecasts, projects then average measure-level EULs on a weighted basis to calculate the expected project EUL. The different uses of MATs for custom and NMEC projects has caused some confusion in the assignment of MATs for site-specific NMEC projects.

The following custom project MATs are allowable in site-level NMEC projects and inform the EUL in different ways:

- **Accelerated Replacement (AR):** "The replacement of existing equipment that could and would remain operational without program intervention."[44] Replacement of "'operating equipment that when broken, non-functional, or unable to provide the intended service is typically repaired' can be classified as AR."[45] AR measures are assigned a measure-specific EUL.
- **Add-On Equipment (AOE):** The installation of "new equipment onto existing host equipment, improving the nominal efficiency of the host system."[46] AOE measures are assigned a default EUL equal to the shorter of the remaining useful life of the host equipment or the EUL of the measure.
- **Behavioral, Retro-commissioning, and Operational (BRO):** Information or education programs that influence energy-related practices (behavioral), activities and installations that restore equipment performance (retro-commissioning), as well as measures that improve the efficient operation of installed equipment (operational). BRO measures are assigned a three-year EUL.
- **Building Weatherization (BW):** The installation of "non-mechanical building efficiency improvements such as windows, insulation, air sealing, and duct sealing."[47] BW measures are assigned a measure-specific EUL.

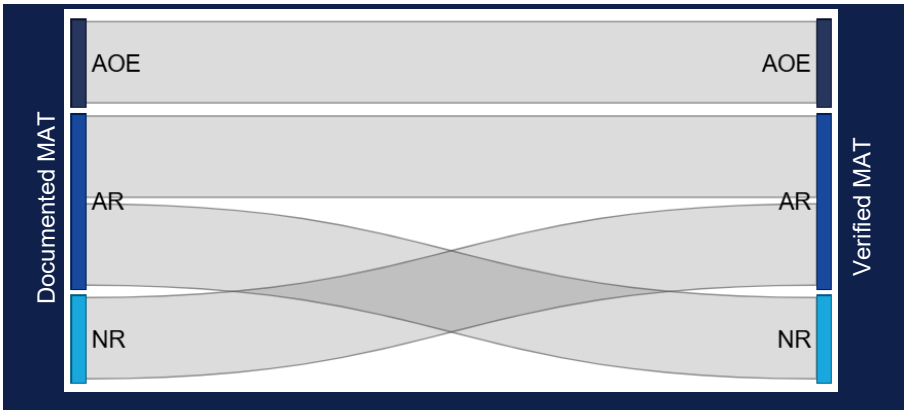
[43] CPUC, *Statewide Custom Project Guidance Document, version 1.4*, June 2, 2021, 5, https://file.ac/OEr-2p-bk3A/.
[44] Ibid., 6.
[45] Ibid., 6.
[46] Ibid., 6.
[47] Ibid., 7.

- **Normal Replacement (NR):** The replacement of existing equipment that has failed, no longer meets needs, or is scheduled to be replaced for reasons unrelated to the program. Treatment of NR measures in NMEC projects is unique, as addressed in Section 4.5.2. NR measures are assigned a measure-specific EUL.

4.5.1 MAT assessment

Figure 4-6 shows the original, PA-claimed PY2023 MATs on the left and the verified, evaluated MATs on the right, weighted by measure-level forecasted savings. While the evaluation re-assigned MATs for some measures, the MAT assessment did not result in any changes in savings this year, since, irrespective of NR or AR MAT assessment, NMEC evaluations use existing conditions as the baseline for savings estimates. (We verified that all of the AOE measures were appropriately classified.) This is a marked improvement over the projects in the PY2020 – 2022 evaluation in which all of the SCE and SoCalREN measures were classified as NR. Though it did not impact savings this year, the distinction between AR and NR continues to be an issue.

Figure 4-6. Measure application type classification, weighted by forecasted savings



4.5.2 Accelerated vs. Normal Replacement

As discussed, MATs were originally developed for custom projects, for which they are used to select the appropriate baseline and EULs. For NMEC, however, the baseline is, by definition, the existing conditions, given the meter-based savings measurement approach. The NR MAT, which uses code or standard practice as the baseline, clashes with NMEC's purpose and approach. Consequently, the distinction between NR and AR does not change the baseline (existing conditions) or the EUL (full measure life). Table 4-11 summarizes the baseline differences for AR and NR measures for NMEC and Custom projects.

Table 4-11. AR and NR baseline comparison between custom and NMEC

	NMEC	Custom
Accelerated replacement (AR)	Existing conditions	Dual (combination of existing conditions baseline and code or standard practice)
Normal replacement (NR)	Existing conditions	Code or standard practice

For custom projects, NR implies that without the program the participant would have installed standard-efficiency replacement equipment. The custom NR MAT limits savings to the incremental difference between the standard-efficiency and high-efficiency options. This requires an estimate of savings relative to a standard-efficiency baseline—an estimate that is not possible in the performance-based framework where the baseline may well be below standard efficiency. This highlights why NMEC has to treat NR measures differently.

One of NMEC's objectives is to unlock the stranded savings in buildings that could maintain and repair below-code systems. Those buildings may not be able to afford high-efficiency measures, but simply bringing systems up to code can drive savings. With that in mind, the existing conditions baseline is appropriate for NMEC projects regardless of MAT. NMEC measures claim the savings between the pre-existing and the installed equipment conditions for the full EUL of the measure—even if the installed conditions are only standard-efficiency.

Because the NR MAT is an incongruous fit with NMEC's performance-based framework, its use in NMEC projects is only appropriate in limited scenarios where an otherwise sound set of non-NR measures can justify the NMEC classification, it makes sense to install the NR measures at the same time as the larger NMEC project, and the up-to-code savings of those NR measures can be addressed by some form of adjustment. The current NMEC rulebook states, "The Project M&V Plan must account for any normal replacement measures within the scope of the project."[48] The draft revised NMEC rulebook calls for the adjustment of NR measure savings to remove below code savings.[49] The NR MAT suggests that the existing equipment required replacement regardless of program intervention, and this adjustment would remove the to-code savings that would have occurred in the absence of the program.

Custom and NMEC projects also treat the AR MAT differently, which covers repair-eligible systems for which replacement is *accelerated* by program incentives. For custom projects, an AR measure corresponds to a dual baseline to account for both the accelerated replacement and the hypothetical eventual replacement. During the *accelerated* period of time between installation and the moment when equipment would have been replaced even without the program, custom projects calculate savings using the existing conditions as a baseline (similar to NMEC). After the moment when equipment would have been hypothetically replaced even without the program and until the end of the measure EUL, custom projects calculate savings as the incremental difference above standard-efficiency (similar to NR).

NMEC projects, on the other hand, are more generous to AR measures, assigning the existing condition as the baseline for the measure's full EUL. Under NMEC, repair-eligible systems that have aged beyond the EUL also continue to be eligible for AR and thus eligible for savings calculated from an existing conditions baseline for the proposed equipment's full EUL.

4.5.3 MATs by project

Given the small population size of the PY2023 gross evaluation, we are including details for these three projects to illustrate some of the issues encountered related to the AR and NR distinction and how the MAT relates to the NTGR.

Performing arts center

The performing arts center project classified two measures as NR in its documentation: a chiller replacement and a glycol/water-chilled water loop replacement. The documentation did not indicate why the NR MATs were selected. The gaps in the MAT and viability documentation are further discussed in section 4.7.1. Based on the customer interview and the project documentation, the existing equipment had been meeting the building's needs and was fully functioning. Based on this information, we re-classified these measures as AR.

Library

The library project involved replacement of a chiller and air-handling unit (AHU). Prior to the project, the existing chiller was working but required frequent repairs. The AHU was providing adequate service but had repairable issues such as corrosion. The project classified both measures as AR in the project documentation, and we maintained that classification based on limited details available in project documentation and the interview with the facility representative.

---

[48] CPUC, *Rulebook, version 2.0*, January 7, 2020, 14, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.
[49] CPUC, *Rulebook, version 2.1*, November 17, 2023, 7, https://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M520/K881/520881077.PDF.

Office building

The office building project classified all five installed measures as AR. We reclassified two measures as NR during the evaluation: AHU and HVAC control system upgrades. The project documentation showed that the existing AHU was broken and needed to be replaced. The documentation included pictures of the old AHU with cardboard patches—a clearly temporary, makeshift fix to a not-fully-operable system. Additionally, photos showed that the existing HVAC control system was disconnected. We estimated that only about 30% of the AHU replacement savings and 20% of the new HVAC control system savings were above code. The AHU and HVAC control system measures had the highest forecasted savings within the project. We saw some indications that the other three measures installed could have also been NR, but found insufficient evidence to support the change from AR to NR.

Gross and net savings for replacement measures

Each of the three PY2023 projects was part of the NTG evaluation in the PY2020-2022 evaluation. Given that the distinction between AR and NR—namely, the possibility that the equipment required replacement anyway—is related to NTG assessment, we closely examined the NTG scores for these projects to make sure that the gross and net evaluation results do not over-adjust for free-ridership.

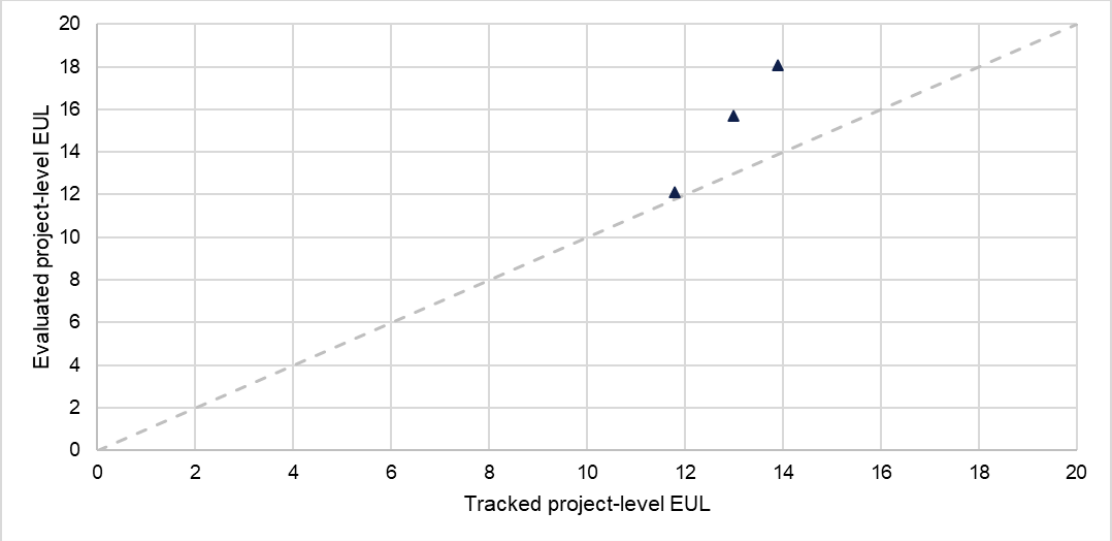
Documentation for the three projects generally did not include evidence that the equipment required replacement because it could not meet the essential services of the facilities. As a result, we determined that, more often than not, the replacement measures were most appropriately classified as AR. Analysis of NTG survey results for these three projects show high levels of free-ridership, suggesting that the customer would have replaced the equipment even without program intervention—and that the equipment did not necessarily *require* replacement. In these three cases, the gross savings reflect the full savings relative to existing conditions, whereas net savings are substantially reduced to reflect the minority of savings influenced by the programs. As the NMEC rulebook's treatment of NR measures may evolve, as addressed in Section 4.5.1, special care should be given to the distinction between gross and net savings in future evaluations.

4.6 Effective useful life (EUL)

The project effective useful life (EUL) is used to calculate the project lifecycle savings. The final savings for NMEC projects rely on a meter-based approach, but the metering results cannot inform the estimation of project EUL. To estimate a project-level EUL, measure-level EULs are assigned according to MAT and are weighted by the forecasted savings.

Figure 4-7 compares tracked and evaluated project-level EULs. All three projects underestimated project-level EULs. None of the three projects provided complete supporting documentation such as a DEER EUL ID or other justification for the measure-level EULs used.

Figure 4-7. Project-level effective useful life (EUL) scatterplot



All three projects had EUL discrepancies which are shown in and described below.

Table 4-12. Claimed vs. verified EUL by project

Project	Claimed EUL	Verified EUL
Performing arts center	11.8	12.1
Library	13.9	18.1
Office	13	15.7

Performing arts center

The performing arts center project included the replacement of a pre-existing glycol/water primary/secondary chilled water loop with a water-only primary chilled water loop. Project documentation claimed a measure EUL of 15 years, even though the documented DEER ID corresponded to an EUL of 20 years. Our evaluated measure EUL is 20 years. Thus, the evaluated project EUL is 12.1 years compared to the claimed EUL of 11.8 years.

Library

The library project included three proposed measures, while only two were installed, but the project's claimed project EUL is the weighted average of the three proposed measures instead of the two installed measures. In addition, one of the two installed measures involved a chiller replacement measure for which project documentation claimed an EUL of 15 years with no EUL reference or other justification provided. The evaluated measure EUL for a water-cooled chiller is 20 years. Therefore, the evaluated project EUL is 18.1 years as compared with a claimed EUL of 13.9 years.

Office building

The office project included six proposed measures, though only five were installed, but the claimed project EUL reflects the weighted average of all six proposed measures instead of the five installed ones. The evaluated project is 15.7 years as compared with the claimed EUL of 13 years.

4.7 Project documentation discrepancies

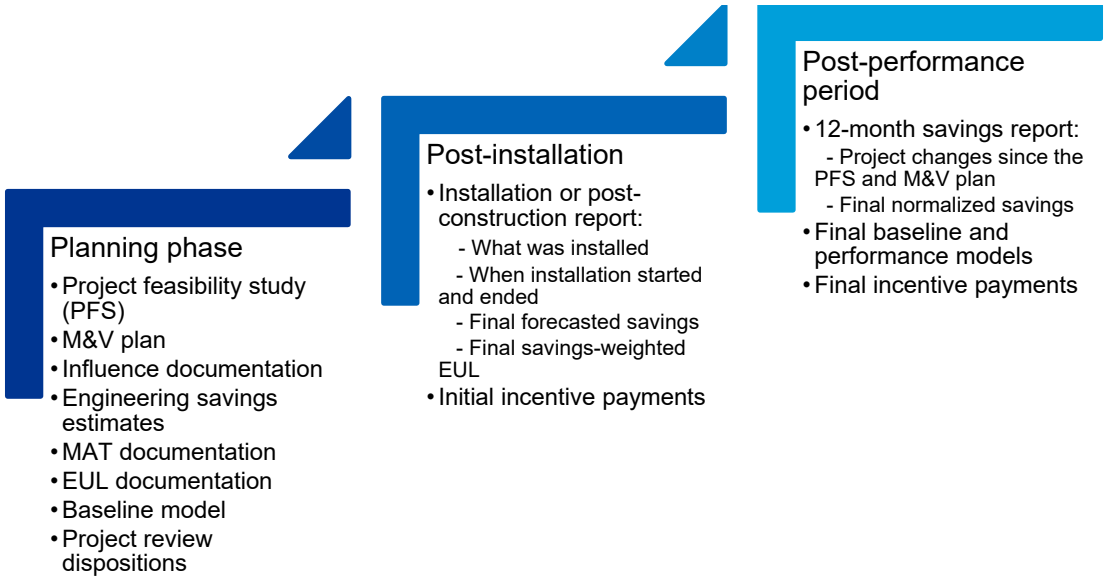
Efficient and thorough project evaluation relies on clear and concise project documentation. The documentation supports the fundamental NMEC claim that the change in pre- to post-implementation consumption is caused by program-related efforts.

While traditional program influence (e.g., NTG) remains an important focus, empirical, performance-based savings estimates require evidence that the program is responsible for, and deserves credit for, the full measured impact. Furthermore, establishing clear documentation expectations may reduce some of the administrative burden and duplicated efforts identified by participants during the interviews.

In this section we walk through the documentation required at each phase of an NMEC project and discuss the types of documentation actually provided as well as gaps that should be addressed to improve clarity and transparency. Separate from this evaluation, the Custom Project Review (CPR) Continuous Improvement initiative is working to comprehensively assess and update the CPR process, including the way in which projects are documented, to increase the CPR process's value to ratepayers and better position it to provide continued value as programs evolve.[50]

The project documentation for NMEC projects typically represents information from three distinct periods within the project: the planning phase, post-installation, and the end of the performance period. Figure 4-8 summarizes some of the key types of documentation expected at each phase of the project. We also expect that the documentation for each phase builds off the documentation from the preceding phase and identifies any changes made. Sections 4.7.1 through 4.7.3 walk through each phase of the project documentation and address the quality of the documentation provided for the three evaluated projects. Section 4.7.4 takes a closer look at the analytical and modeling documentation provided across the three project phases.

Figure 4-8. Site-level NMEC documentation phases



4.7.1 Planning phase documentation

Participants prepare the project feasibility study (PFS) and measurement and verification (M&V) plan as part of the initial project scoping. The PFS and M&V plan can be the same document or two different documents. Either way, this documentation typically includes information about the site; a program influence narrative; the planned energy efficient activities and measures; measure-level MATs; the expected measure-level savings; and EULs.

The NMEC rulebook says that project documentation during the application phase should include:

- **Savings and incentive estimates:** "Estimates of energy savings and incentive payments"

[50] CPUC, *Decision 11-07-030*, calmac.org, July 14, 2011, https://edcentralserver.files.com/f/6XsdKwUyhqo.

- **M&V Plan and NMEC feasibility:** A "project M&V plan and demonstration of [the] feasibility of [the] NMEC analytical approach"
- **EUL documentation:** Documentation of "methods and values used to develop [the] project EUL"
- **GRR and NTGR adjustments:** "Planned adjustments for gross-realization rate (GRR) and net-to-gross (NTG) factors"[51]

In this section we address key aspects of NMEC projects that should be documented during the planning phase and discuss the documentation provided for the three evaluated projects.

Viability of existing equipment

Understanding the viability of existing equipment is important for both determining whether the project is appropriate for NMEC and for selecting the appropriate MAT. Documentation should demonstrate that existing equipment was still serving the requirements of the building. This is essential to demonstrate that replacement will, in fact, lead to consumption reduction. It is also a prerequisite to establish that the project is not a forced replacement. In the PY2020 – 2022 evaluation, we found an instance where the existing equipment was in such poor repair it did not actually use energy. Its replacement with a new, functioning system, expected did not achieve savings but, in fact, increased energy use. There was no documentation demonstrating the viability of that existing, non-functional equipment. In the current evaluation (PY2023), in two instances it was only through direct communication with the participant that it became clear the existing equipment had been disconnected or was in such poor condition that it was patched with cardboard and duct tape. This information was not included in the project documentation from the PA. In these cases, the projects still produced savings due to the types of measures (controls and AHU), but documenting and considering this information is still important, ensuring both that project savings estimates are reasonable and that an NMEC approach is feasible for the site.

Code compliance

Documenting code compliance is necessary to ensure that a project is eligible for NMEC. NMEC is unique among energy efficient programs in allowing projects that only meet code rather than requiring projects exceed existing code. This makes it even more essential to document that this minimal threshold is met. One of the three projects evaluated provided no documentation regarding how the proposed equipment's performance would compare to code.[52] For the two projects that did provide code compliance information, the project feasibility study documented whether the proposed equipment performance would be above code. In addition, more detailed information about equipment energy performance was available in the post-construction inspection form.

MATs

MAT documentation is needed as part of EUL documentation and as part of the documentation needed to demonstrate that a project is eligible for NMEC. We continue to see evidence of confusion regarding MATs in NMEC projects, as well as insufficient documentation to support claims. (See Section 4.5 for a detailed discussion of this issue.) The PFSs for two of the three projects reviewed during this evaluation did not provide adequate documentation to explain why particular MATs were assigned. One project, for example, classified a measure as Normal Replacement (NR), but the project documentation provided no supporting evidence why the project was suitable for NMEC, not documenting the condition of the pre-existing equipment. After the customer interview, we requested the project design narrative from the customer—which suggested that the pre-existing equipment was maintaining the essential services of the facility and did not require replacement. Based

[51] CPUC, *Rulebook, version 2.0*, January 7, 2020, 14, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.

[52] Assembly Bill 802 requires that utility-incented measures at existing buildings "bring them into conformity with, or exceed, the requirements of Title 24 of the California Code of Regulations."

on the limited information available from documentation and the interview with the customer, evaluators determined an MAT of Accelerated Replacement (AR) instead of NR. Another project documented all replacement measures as AR. After talking to the customer and acquiring the assessment report, the evaluation team noticed that the pre-existing equipment condition for two measures was poor, with, for example, ducts patched with cardboard. However, despite their condition, they were continuing to service the space indicating an appropriate history of repair.

Effective Useful Life

The rulebook specifically requires that NMEC projects provide EUL documentation. Evaluators must carefully review EUL, as the resulting lifetime savings are important for calculating cost-effectiveness and total system benefit. Two out of three projects reviewed during the gross evaluation had insufficient documentation to explain how they justified certain measure-level EUL values. As with the current evaluation, the prior evaluation also found that a significant number of measures were assigned shorter EULs than they deserved.

Baseline model

A clear description of the baseline model demonstrates NMEC feasibility and sets the modeling structure for the performance period. It is essential that project documentation provide all necessary information regarding this process because modeled consumption forms the basis of savings claims for NMEC. For this evaluation, all evaluated M&V plans included a description of baseline model variables; descriptions of non-routine events and related adjustments; and basic goodness of fit statistics. However, none of the M&V plans provided basic baseline model output and specification details, such as parameter estimates, p values, temperature bin values, and occupied/unoccupied mode settings. Evaluators need these values to successfully replicate efforts and to fully assess whether the model reflects accurate engineering conditions at the site.

The documentation provided in the planning phase serves as the starting point for the documentation for future project phases, which should refer to the original plan, document any changes made subsequent to that plan, and explain those changes. Given that NMEC projects necessarily span years, it is unreasonable to expect evaluators (or even customers) to fill in documentation gaps at the end of the project. Additionally, all of the documentation discussed here is required prior to project implementation and may be reviewed if the project is selected for ex ante NMEC project review. The NMEC project review process offers an early opportunity to review project documentation for both completeness and reasonableness. These reviews assess whether project documentation is sufficient and then highlight places where the project will require additional attention as it moves toward the ex post evaluation process. In addition, resolving project issues and questions during the review phase should clarify expectations going into the ex post evaluation.

4.7.2 Post-installation documentation

Projects typically include an installation report, post-construction inspection report, or some sort of documentation confirming installation. While the NMEC rulebook does not specifically require an installation report, it does stipulate that “the reporting period stage begins once the measures are implemented and/or installed and confirmed to be working and producing savings.”[53] The rulebook also states that “NMEC project savings forecast estimates may be based either on approved deemed-measure workpapers or may be calculated using engineering or modeling methods consistent with Commission adopted custom project savings-calculations…”[54] Additionally, “Project savings estimates must reflect measure-level savings to inform expected useful life (EUL)…[and] project lifecycle savings must be based on a weighted average EUL method.”[55]

[53] CPUC, *Rulebook, version 2.0*, January 7, 2020, 15, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.
[54] Ibid., 15-16
[55] Ibid., 16

Ultimately, the information required for the PFS is not correct if it is not updated to reflect the final scope of the installed project.

We have found that the measures installed are often different from the measures originally scoped. Some changes are relatively minor, installing a similar measure for example, but others are substantial, such as deciding not to install a large saving measure. For a project to make an initial claim using the appropriate forecasted savings and EUL, the PA must know what was installed and when it was installed and update savings forecasts and EULs as necessary before making the claim. Additionally, if a participant decides not to install some measures with high forecasted savings, the PAs should confirm that the project's fractional savings and the fractional savings uncertainty for the baseline model still make the project a good fit for NMEC.

One of the three projects evaluated did not provide sufficient documentation to determine what was installed. For this project, we determined during the customer interview that one planned measure was not installed, even though the documentation did not mention any changes to this measure. In addition, the installation inspection form did not provide clear information about whether all planned measures were installed or provide the measure specifications, such as tonnage or efficiency.

Because NMEC savings claims are ultimately established based on meter-based performance, some may see the importance of the accompanying documentation to be secondary to the meter-based results. However, project documentation plays an essential role, providing an engineering basis to establish the validity of the NMEC project and claim. The lack of documentation regarding changes made to savings forecasts and EULs appears to indicate that the projects were not assessed following project installation. This is particularly important to note, as this likely means that the initial claims, which are made after project installation, did not reflect the final installed project savings and EUL.

4.7.3 Post-performance period documentation

The final savings report is a key document that summarizes the final project, including any changes to the project throughout implementation, the performance period, and savings normalization. The NMEC rulebook states that the final savings report, which it refers to as the "final M&V report," should include the following:

- **M&V documentation:** Documentation of "the activities carried out per the M&V Plan" [56]
- **NRE findings and adjustments:** "Data collection (pre- and post- installation) adjustment models and all findings related to routine and non-routine events" [57]
- **Final savings:** "First year and lifecycle savings claims, final avoided energy use and final normalized energy savings"[58]
- **Deviations from M&V plan:** "Any deviations from the proposed M&V Plan should be documented and substantiated"[59]
- **Commission recommendations**: The report "should reflect Commission staff review recommendations, if the project was selected for review"[60]

While all projects provided a final savings report with normalized savings estimates (final savings), very few provided a sufficient discussion of project implementation, deviations from the plans, changes at the site (NREs), model specifications, or savings results—the last particularly important when a project does not achieve expected savings.

[56] CPUC, *Rulebook, version 2.0*, January 7, 2020, 15, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.
[57] Ibid., 15.
[58] Ibid., 15.
[59] Ibid., 16.
[60] Ibid., 16.

Changes made following CPUC staff reviews

Project documentation should identify any changes made as a result of CPUC staff review, clearly documenting how the project is following review recommendations. One of the three projects evaluated went through the CPUC staff review process (CPR). However, this project's documentation did not specify any changes made due to the review or explain why changes were not made. For instance, the CPR disposition pointed out that the baseline model's R^2 value was lower than LBNL goodness of fit guidance and suggested some updates to the engineering calculations, but project documentation did not discuss or address either issue.

Baseline model changes

Baseline models should be set during the project feasibility phase, as part of the screening process that determines whether the project is well suited for NMEC. Occasionally, if installation is delayed, projects may need to adjust baseline models, as required by the NMEC rulebook. Similarly, unforeseen events such as COVID-19 may also require modeling adjustments. It is important that participants explain and justify any changes to the baseline model in documentation. For example, moving the baseline period closer to the start of the implementation period would generally be a reasonable adjustment, whereas moving the baseline period to show the greatest savings level would not be a reasonable adjustment. One project changed the modelling approach for both the baseline and performance period due to COVID-19 impacts and explained the additions of occupancy related variables.

Site changes

The pre- to post-implementation difference in consumption must represent only reductions related to program-related changes. Easily available information and load data made clear that two of the three sites reduced operating hours between the baseline and the performance periods, but this was not discussed in the final savings report. Just as a partial or complete shutdown of a site would appear as performance savings but would not qualify as NMEC savings, a reduction in active hours must be documented and addressed.

4.7.4 Analytical method documentation

The NMEC rulebook requires that all NMEC projects be transparent, well documented, and replicable: "Data, methods and calculations must be made available to the PAs [as] well as the Commission and its impact evaluators….The methods used to calculate savings for NMEC programs must be documented in the program-level M&V Plan sufficiently such that savings calculations are able to be replicated by the PAs as well as the Commission and its impact evaluators."[61]

Data preparation transparency

The project files included the raw, unmodified consumption data and other inputs to the model, the processed data that goes into the model, and the code or other tools used to run the model. However, none of the project files included the code used to transform the raw data into prepared data or a narrative explaining modeling decisions. While evaluators can replicate models using the prepared data and the provided modeling code, they cannot verify the methods and assumptions used to prepare the raw data if they don't know what those methods and assumptions are. Across both the PY2020 – 2022 and PY2023 evaluations we have seen the data preparation approach in only a handful of cases—and in one of those (from PY2020-2022) we identified a key problematic assumption.

[61] CPUC, *Rulebook, version 2.0*, January 7, 2020, 18, https://www.cpuc.ca.gov/-/media/cpuc-website/files/legacyfiles/n/6442463694-nmec-rulebook2-0.pdf.

Final model

All three of the projects provided modeling code. However, in all three cases the code provided did not exactly replicate the model fit statistics and normalized savings reported in the final savings report. For one of the three models the provided code produced a substantially different result than either the final savings report or the tracking data. When normalizing consumption, the provided code did not address COVID-19 or other NRE-related independent variables included in the baseline and performance period models. It is not the role of the evaluator to infer information from output files and modify the code in order to replicate savings. In this case, the discrepancy between the normalized savings in the final savings report and the savings produced by the provided code was approximately 92,000 kWh. The evidence suggests version control may be an issue, as models are understandably being updated as the project progresses.

Additionally, the provided code did not always produce outputs that clarified the final model specification. Across both the PY2020 – 2022 and PY2023 evaluations output typically included goodness of fit statistics and normalized savings but often did not include the full set of parameter estimates. These are the key outputs of the model that directly produce normalized consumption estimates and final savings estimates. It is very difficult to replicate and validate the project modeling process without these model outputs.

Model narrative

The documentation for the three evaluated projects explained which variables they used in the models and whether the models addressed NREs, but documentation did not explain modeling decisions and did not provide any interpretation of the results. When projects use anything other than the most basic models, a discussion of specific modeling decisions increases analytical transparency and clarity.

5 KEY FINDINGS AND RECOMMENDATIONS

5.1.1 Gross and net savings findings and recommendations

NTGRs have increased from last year's evaluation, driven by both increases in reported program influence and by methodological changes.

Increases in reported program influence: The NTG survey asked respondents to divide a total of 10 points across two types of influential factors, those that are program-related and those that are non-program-related. The relative allocation of these points indicates the importance of all program influences relative to all non-program influences in the decisions to implement projects. In last year's evaluation, respondents gave an average of 5.1 points to program influences, indicating that program and non-program influences were equally important in their organizations' decisions to do the EE projects. This year, respondents gave an average of 7.5 points to the program influences, indicating that program influences were more important than non-program influences. Respondents particularly highlighted programs' technical support with one saying, "They offer an extension of our staffing. I don't have time to go and do research on what programs are available and what aligns most with our school district's needs. So having ongoing meetings with them to touch base gives me an extension of my capacity." Better program involvement with the customers improved program influence.

Decision timing: Additionally, no respondent in the current evaluation indicated that their organization had decided to do the project before interacting with the program, compared to 40% of respondents last year. Similarly, in response to a new question in this year's survey instrument, 80% of respondents indicated that their organization first learned about the opportunities included in their EE projects from program or utility staff. One respondent said, "[The biggest strength of the program was the] awareness that those audits brought to us. The audits bring a lot of issues to our attention which are hard for us to recognize ourselves."

Methodology update: To assess the impact of the methodology update on NTGR estimates, DNV calculated NTGRs for this year's sample using the previous methodology. Table 1-5 compares the NTGRs using the current and the old methodology. Using the previous methodology, NTGRs for this year's sample range from 60.0% to 62.8%. Approximately half of the increase in NTGRs from PY2020-22 to PY2023 is attributable to the changes in scoring methodology.[62]

Table 5-1. NTGR methodology and sample comparison

NTGR	PY2020-2022 evaluation Old methodology	PY2023 evaluation Old methodology	PY2023 evaluation New methodology
Statewide electricity	45.9%	61.3%	76.6%
Statewide demand	41.7%	61.0%	75.2%
Statewide natural gas	46.5%	60.2%	76.0%

The accuracy of savings claims in the tracking database system improved compared to PY2020-2022. However, incorrectly entered savings claims remained an issue in PY2023.

The NMEC savings claim process is more complicated than the typical custom claim process as it must accommodate the final, meter-based savings estimate, which projects calculate a year after implementation. Projects claim engineering-based, forecasted savings during implementation. A year later, after the performance period, projects calculate the meter-based normalized savings and enter into tracking a true-up claim that represents the difference between the two values. The two claims should sum to the final, meter-based savings estimate. The novel claims process for NMEC led to a reporting inaccuracy for one out of three projects during PY2023. This was an improvement compared to PY2020-2022, when

[62] Note it is not possible to run the PY2020-2022 evaluation results through the new methodology because questions were asked differently to respondents in that evaluation. However, the Group D NTG Methodology Update_FinalMem-01072025-CLEAN memorandum provides a sensitivity analysis (Section 3.3.1) to the extent possible. The sensitivity analysis found for NMEC that the new methodology increased NTGRs by about 9%. The new methodology tends to make high NTGRs higher and low NTGRs lower when compared to the previous methodology, which pushed NTGRs towards 0.5.

incorrectly entered savings claims in the tracking database system were the largest source of savings discrepancies. In the PY2020-2022 evaluation, seven out of 22 sampled projects incorrectly entered savings.

The PY2023 project with the reporting inaccuracy (mentioned above) submitted two true-up claims, which over-adjusted the initial claimed savings. Together, the three claims for this project resulted in zero first-year kWh savings.

During the Response to Recommendations process following the PY2020-PY2022 evaluation report, the PAs indicated that they were working to improve the claims process. Future evaluations will continue monitoring this issue.

Recommendation

PAs should continue to focus on improving the timeliness and accuracy of site-level NMEC claims, following existing NMEC reporting guidance:[63] make the initial claim in the quarter measures are installed, use a ProjectID that can be tracked across project years, and make the true-up claim in the quarter in which the performance period is completed.

Models should reflect empirical conditions within pre- and post-installation periods and be normalized across the two periods.

NMEC savings claims rely on statistical models that characterize site-level consumption and allow for the comparison of pre- and post-installation consumption at the site on an apples-to-apples basis. To be eligible for NMEC, the pre-installation model needs to demonstrate the ability to explain site-level consumption well enough that expected savings will be discernible. The process also needs to encompass all changes in site-level consumption that are not program related.

Reliance on modeling algorithms rather than empirical schedule data: While the NMEC Rulebook is not prescriptive regarding the details of modeling decisions, it does recognize the central role of quantitative empirical data in the NMEC modeling process. All three evaluated sites used a popular site-level modeling package that decides how to characterize the site solely on statistical fit. While this approach has merit, NMEC implementers should also consider the wider context of empirical data. For example, implementers should consider known site schedules and override model-based outcomes where indicated by empirical and contextual information. Similarly, the modeler needs to decide if schedules affect only non-weather-correlated, baseload consumption or also affect weather correlated (HVAC) consumption. Empirical data has a role to play in those decisions as well.

Accounting for known pre-to post-installation changes: Meter-based savings methods must be informed by any empirical data relevant to energy consumption during the pre- and post-installation periods. For example, a project with different operating schedules before and after installation, a change that is unrelated to program participation, should explicitly address the implications for energy consumption. In two instances, a building shortened its operating hours in the period after project installation. With fewer open hours, energy consumption will likely drop even without any other intervention. Neither of the two projects with this issue attempted to address it, in effect, claiming as savings consumption reductions actually due to a decrease in operating hours. The evaluator's adjustment to reflect the schedule changes resulted in lower verified savings.

Recommendations

- PAs should ensure that models align with empirical data regarding how the building operates and how savings will be achieved. This will help enhance model accuracy and PA savings estimates.

[63] This is explained in reporting guidance published by Energy Division as NMEC Reporting Guidance 04242020.pdf that was distributed to the PAs.

- The PAs should ensure that all projects address any changes to site operations so that projects claim only savings due to energy efficiency improvements rather than, for example, a reduction in hours.

Projects used Effective Useful Life (EUL) values without providing justification.

Two out of three projects reviewed during the gross evaluation had insufficient documentation to justify certain measure-level EUL values.[64] A measure's EUL indicates how long the first-year savings will persist and must be based on documentation, just as with non-meter-based custom projects. EULs cannot be measured using one-year of post-installation consumption data because EULs are generally longer than one year. Evaluators must carefully review EUL, as the resulting lifetime savings are important for calculating cost-effectiveness and total system benefit. For two replacement measures, projects used an EUL of 15 years, without providing justification, while the DEER-based EUL for both measures is 20 years. For these three projects, the evaluated EULs were higher than the claimed EULs, which resulted in an increase in lifetime savings.

Recommendation

Measure documentation should include a description of the measure, its EUL, and its respective DEER EUL ID (or other similar citation) to justify each measure's EUL.

5.1.2 Documentation findings and recommendations

Project documentation did not sufficiently catalogue either existing equipment or installed equipment.

DNV expects that project documentation will explain the project's initial plans, including pre-existing equipment types and condition, as well as any changes that occurred during project development, implementation, or the performance period. Some projects made a good effort to provide clear documentation. Documentation for all projects, however, could have been improved. The following areas present the greatest room for improvement.

Viability documentation: Understanding the viability of existing equipment is key to determining whether the project is appropriate for NMEC and for selecting the appropriate measure application types (MATs)[65]. Documentation should demonstrate that existing equipment was still serving the requirements of the building to demonstrate that replacement will, in fact, lead to consumption reduction. For one project, after talking to the customer and acquiring the assessment report, the evaluation team noticed that the pre-existing equipment condition for two measures was poor, with, for example, ducts patched with cardboard.

MAT documentation: Most projects reviewed during the evaluation did not provide adequate explanation of selected MATs. One project selected a MAT of Normal Replacement (NR) but provided no supporting evidence showing the condition of pre-existing equipment. Through the customer interview, DNV determined that the pre-existing equipment had been operating properly. As a result, evaluators reconsidered the MAT and ultimately assigned Accelerated Replacement (AR) instead of NR. Another project designated all replacement measures as AR but also provided information showing that the pre-existing equipment had been broken. Therefore, the evaluated MAT for the two measures became NR instead of AR.

Measures scoped but not installed: Two projects planned to include lighting upgrades, which would have significantly contributed to total anticipated savings. However, the final savings reports indicated that neither project completed a lighting

[64] A measure, in this context, is a specific customer action that reduces or otherwise modifies energy end-use patterns or a product whose installation and operation at a customer's premises reduces the customer's on-site energy use.

[65] For more detailed definitions of each MAT, see: https://www.caltf.org/measure-application-types-1

upgrade. If a participant decides not to install some measures with high forecasted savings, the PAs should confirm that the change does not reduce savings to the point NMEC methods are no longer feasible. Additionally, projects need to update the forecasted savings and EUL after installation so that the initial tracking data claims accurately reflect the completed project, whether savings are lower or higher than originally forecasted.

Code compliance: One of the three projects evaluated provided no documentation regarding how the proposed equipment performance would meet to-code compliance.[66]

Recommendation

Even though the savings for site-level NMEC projects are meter-based, it is essential to clearly document the viability of the existing equipment, the details of the measures installed, and code compliance.

Regression-based modeling is the core of NMEC methods, and projects do not consistently provide transparent, well-documented models following standard practices.

Model replication issues: DNV attempted to reproduce projects' model savings results using the code, spreadsheets, and/or other tools provided in the project documentation. Documented and replicated values tended to align, though exact replication was not possible in any of the three cases. One project required changes to the provided code to achieve similar results, suggesting that the code provided was not the version that produced the outputs. For the other two projects, the differences were less meaningful, but some results in the final savings report did not align with the results in the provided output files.

Data processing transparency: None of the projects included the scripts used to process and clean the data used in the models. Strictly speaking, an analyst can replicate the model using the prepared data and the provided modeling code. However, when projects do not include the code used to transform raw data to model-ready data files, analysts cannot verify the methods and assumptions used to prepare the data.

More complete model narratives: While documentation usually explained the relatively novel aspects of models (for example, use of airflow data as an occupancy proxy), the full model narrative was frequently incomplete. Documentation did generally stipulate the model used, the variables used, and whether non-routine events were addressed, but included very little information explaining decisions or interpreting the results. Without clear explanations for modeling decisions, evaluators are left guessing.

Model parameters and summaries: All evaluated M&V plans included a description of baseline model variables; descriptions of non-routine events and related adjustments; and basic goodness of fit statistics. However, none of the M&V plans provided basic baseline model output and specification details, such as parameter estimates, p values, temperature bin values, and occupied/unoccupied mode settings. These values are essential to ensure successful replication efforts and to fully assess whether the model reflects accurate engineering conditions at the site.

Output errors: For all three projects, the model documentation included temperature and consumption values with timestamps that were off by a day for daily models and an hour for hourly models. While this issue did not make a large difference in savings, it could dramatically affect demand savings, and it was a recurring issue that should be addressed for future projects.

[66] Assembly Bill 802 requires that utility-incented measures at existing buildings "bring them into conformity with, or exceed, the requirements of Title 24 of the California Code of Regulations."

Recommendations

- PAs should provide the data processing scripts or files used to prepare the data for modeling, along with the complete, final modeling scripts and files needed to exactly reproduce the savings in the project documentation and the CEDARs tracking database.
- PAs should provide complete model output as well as a model narrative, including an explanation of any modeling decisions made, changes from the M&V plan, and an analysis of any substantial deviations in savings.

5.1.3 Process findings and recommendations

Participants continued to indicate high levels of satisfaction with the site-level NMEC programs, driven by the programs' technical support and incentives.

When asked to rate program satisfaction on a scale of zero to 10, where zero is completely dissatisfied and 10 is completely satisfied, respondents gave an average rating of 8.4, indicating a high level of satisfaction. This is an increase even over last year's high average rating of 8.1. Respondents highlighted programs' technical support, which boosted their staff's capacity and provided essential information for their decision-making processes. No respondent provided a satisfaction rating lower than 7 and therefore there were no detractors (i.e., respondents who provided a rating of 3 or less.)

Recommendation

Programs should continue all levels of technical support provided to customers, especially pro-active efforts to build relationships.

Customers want a more streamlined process and better communication from program administrators and implementers about program expectations and timeline limitations.

When asked for program improvement suggestions, respondents most frequently mentioned streamlining the participation process and improving communication. One respondent said, "One of the difficulties was the amount of back and forth with technical reviewers. In hindsight, it was a lot of work on our end and we would have preferred to let a consultant or vendor deal with that." Another said, "Communicating and understanding timeline requirements from the onset would be beneficial for both sides. What timeline does the program require and what can we accomplish it in that timeline based on our own restrictions?" Both issues could potentially be improved with clear communication about expectations and timeline requirements.

Recommendations

- In some cases, PAs should consider requesting an exception to the 18-month installation period when there are extenuating circumstances and other reasonable solutions have not worked.
- PAs and implementers should clearly explain timeline needs early in the process, providing, for example, clear project documentation and regular check-in meetings.

APPENDIX A. DETAILED NET-TO-GROSS RESULTS

PAI details

We calculated NTGRs using an approach that differs slightly from the NMEC evaluation of PY2020-2022. While building on a well-established approach that has been used for nearly a decade to evaluate commercial programs, the adjustments to the NTGR scoring algorithm better reflect NMEC program influences and reduce overlap among the program attribution indexes, countering the previous methodology's tendency to push NTGRs to 0.5.

Table A-1 describes the three program attribution indexes (PAI_1, PAI_2, and PAI_3) and shows the average score for each indicator.

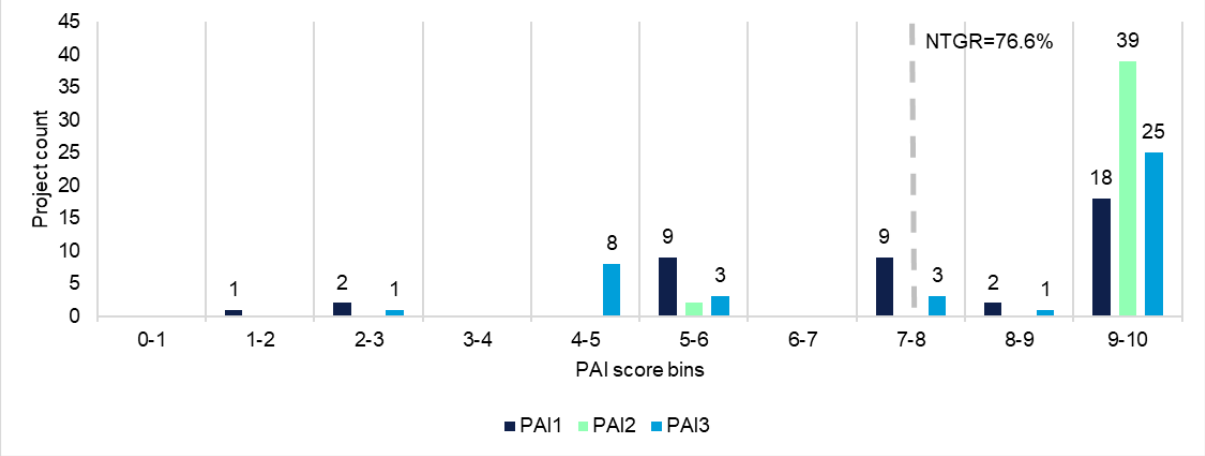
The NTGR is calculated by averaging the three PAIs, resulting in a NTG score primarily based on influence ratings rather than the more direct timing and scope change measurements.

Table A-1. Program attribution index (PAI) results

Program attribution index	Basis	Average
PAI_1	Respondents' ratings on the importance of individual program and non-program influences in their decision to implement a project	7.7
PAI_2	Respondents' rating on the timing of project implementation relative to program interaction	9.7
PAI_3	Respondents' ratings for the likelihood they would have implemented a similar project scope on a similar timeline in the absence of the program	7.9

We calculated the NTGR by averaging the three PAIs, resulting in a NTG score primarily based on influence ratings rather than the more direct timing and scope change measurements. Detailed PAI results are included in APPENDIX A.Figure A-1 shows the distribution of the three scores. For each PAI, most sites have a score of 7 or above.

Figure A-1. Program attribution score distribution



Program attribution index score 1 (PAI_1) individual influence ratings

As shown in Table 3-3, PAI_1 captures the importance of program and non-program influences. DNV asked respondents to rate how important various potential influences were on their decision to implement their project when they did. Respondents provided a rating on 0-to-10 scale, where 0 means "Not at all important" and 10 means "Extremely important." We also

asked respondents to divide ten "points" between the collective program influences they identified as material and the collective non-program influences they identified as material. We calculated PAI_1 using both of these influence ratings as shown in Table 3-3.

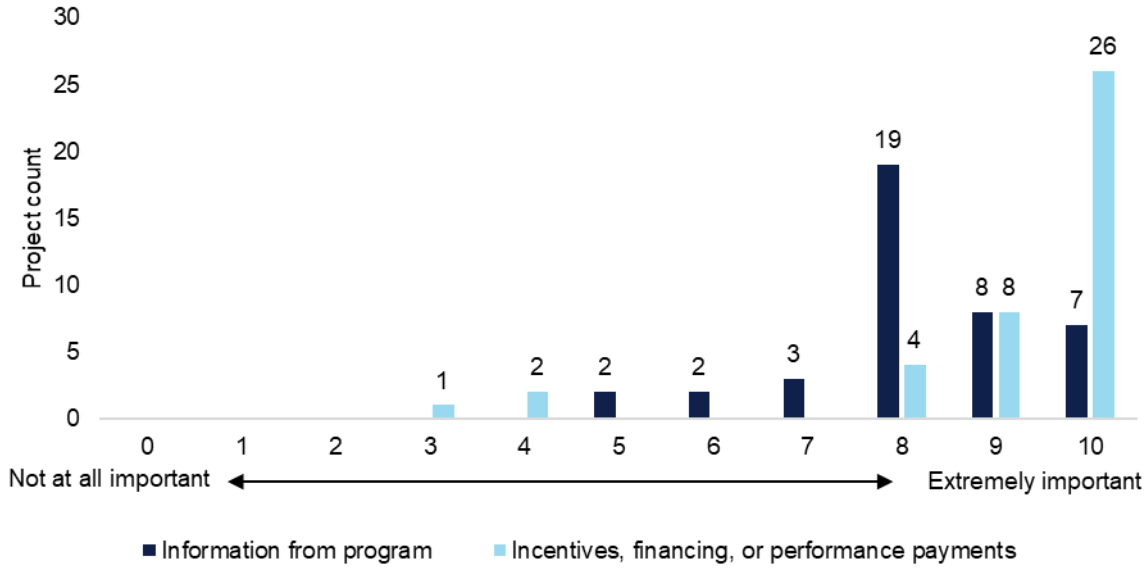
Table A-2 shows the average rating provided by respondents for each potential influence weighted by number of projects per respondent. Across all four influences, program and non-program, the program's technical assistance and rebates were the only influences to receive an average rating of 8 or above.[67]

Table A-2. Influence ratings for program and non-program influences

Type	Influence	Average
Program	Incentives, financing, or performance payments	9.1
Program	Information provided by the program	8.2
Non-program	Payback without incentives	6.7
Non-program	Company practices	6.3

Figure A-2 shows the distribution of respondents' ratings of the program influences. Counts are shown by number of sites per respondent. Respondents indicated that program influences were generally more important than non-program influences by providing more scores of 8 or above program influences.

Figure A-2. Program influence distribution.



[67] If respondents said an influence was "not applicable" their response rate treated as a rating of "0."

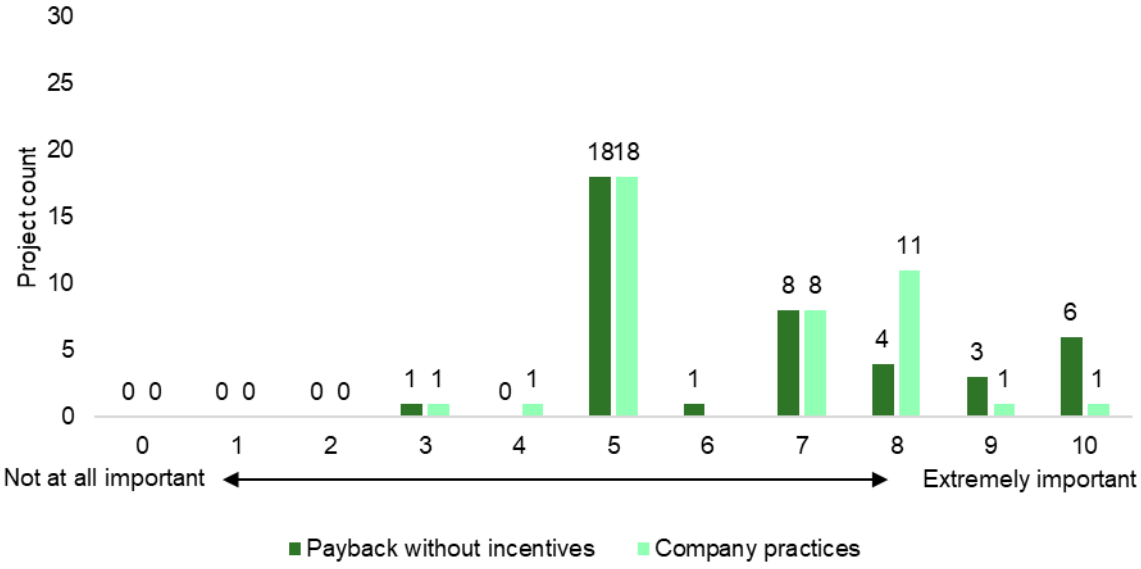
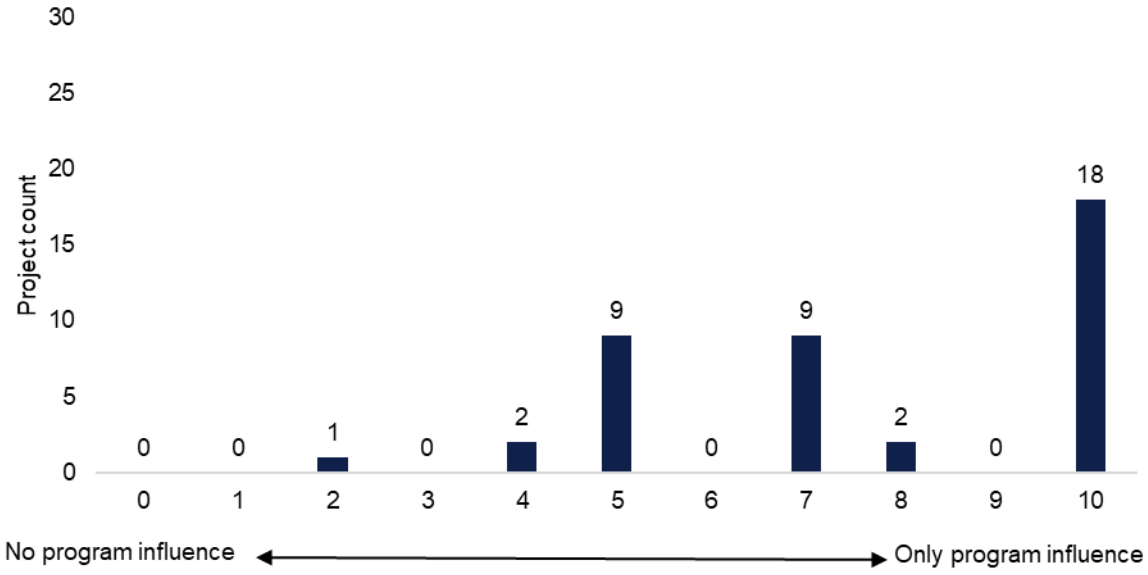


Figure A-3 shows the points awarded to program influences collectively when respondents were asked to divide ten "points" between the collective program influences and the collective non-program influences. Respondent could give 10 "points" to the collective program influences indicating that the program influences were the only influences that mattered in their decision making, zero points to the collective program influences indicating that the program influences had no impact on their decision making, or some amount of points in between zero and 10. Values are identical across sites for a single participant if that participant indicated that their organization used one decision-making process across sites. Respondents representing a majority of the sites provided ratings of 7 or above indicating that the collective program influences were more important in their decision making than the collective non-program influences. Respondents representing 7% of sites provided a rating less than 5, indicating that the non-program influences were more important than the program influences in their decision making.

Figure A-3. Collective program influence relative to collective non-program influence.



Program attribution index score 2 (PAI_2) individual influence ratings

PAI_2 reflects the prior plans of a participant, namely if a site had plans in place and budget set aside for the project prior to interacting with the NMEC program.

If the respondent reported that they made the decision to do a project **after** interacting with the program, PAI_2 was set to 10.

If they reported that they made the decision **before** interacting with the program, PAI_2 was set to 0.

If they gave a mixed response—for example if some there were prior plans for some measures in the project but not others—PAI_2 was set to 5.

Table A-3 details how respondents' answered the survey question underlying PAI_2. Ninety percent of respondents, representing 95% of sites said the decision to do a project was made after interacting with the program. No respondents said the decision to do a project was made before interacting with the program.

Table A-3. Decision making timing compared with incentive and technical assistance timing

Was the decision to do this project made before or after you began discussions with [implementer] regarding the availability of incentives or technical assistance for this measure?	Percentage of respondents	Percentage of sites
Before	0%	0%
After	90%	95%
Mixed	10%	5%

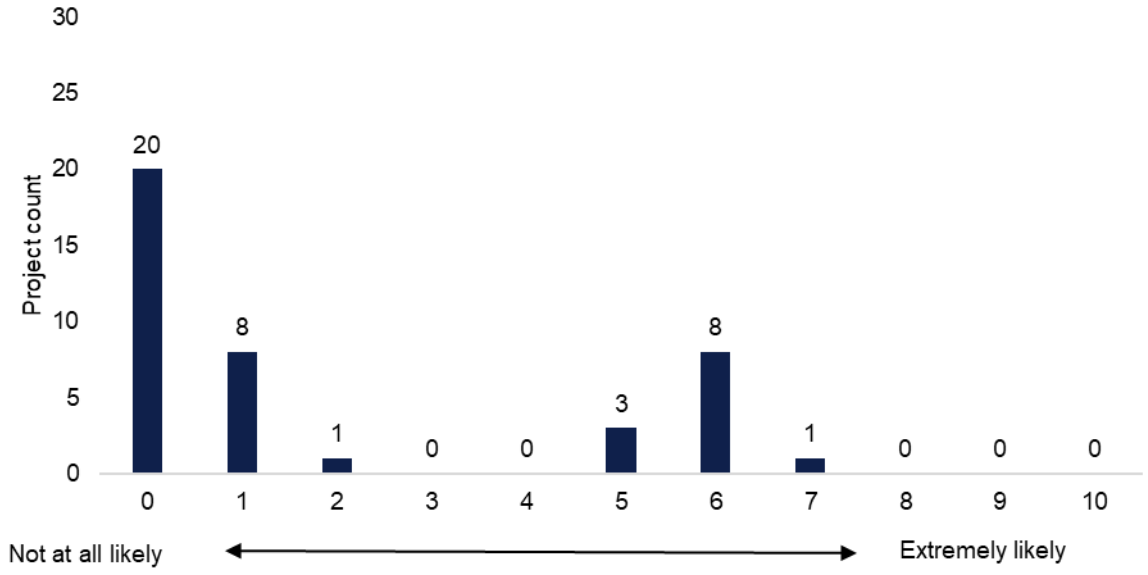
Program attribution index score 3 (PAI_3) individual influence ratings

PAI_3 captures respondents' estimates of the program's influence on project scope and timing. (Again, see Table 3-3 for more details on NTG methodology.)

Respondents were asked to rate the likelihood their project would have taken the same scope without the NMEC program, using a scale of 0 to 10, where 0 is "Not at all likely" and 10 is "Extremely likely." Using the same scale, they were also asked to rate the likelihood that, even without the NMEC program, they would have implemented their project at the same time.

In response to the scope question, respondents representing 71% of projects provided a rating of 3 or below—their project scope would have been different without the program. Respondents were not specifically asked to elaborate on how the scope would have differed, but those who did provide an explanation indicated the scope of their project would have been smaller, involving less extensive energy efficiency improvements. Figure A-4 shows respondents' ratings by site of the likelihood their projects would have taken the same scope in the absence of the program.

Figure A-4. Likelihood project would have had the same scope without the program



Note: "0" = "Not at all likely" and "10" = "Extremely likely"

Table A-4 shows respondent's responses to the question asking what the likelihood was that they would have conducted the project at the same time as they did without the program. When asked about project timing, 70% of respondents, representing 90% of sites, said it was, "very unlikely" that they would have implemented their projects when they did without the program. Only 10% of respondents, representing 3% of sites, said it was "somewhat likely" they would have implemented their project at the same time and none said that was "very likely."

Table A-4. Likelihood of implementing project at the same time without the program

If the program had not been available, what is the likelihood that you would have conducted the project at the same time as you did?	Percent of respondents	Percent of sites
Very likely	0%	0%
Somewhat likely	10%	3%
Neither likely or unlikely	0%	0%
Somewhat unlikely	20%	7%
Very unlikely	70%	90%

When asked how much later they would have implemented their projects without the program, 20% of respondents, representing 17% of sites, said they would never implemented the project. Another 60% of respondents, representing 79% of sites, said they would have implemented their projects two or three years later than they did. Only 10% of respondents, representing 2% of sites, said they would have done that project at the same time or earlier. These responses, summarized in Table A-5, were not part of the PAI_3 score.

Table A-5. Whole project timing

Without the assistance received from the [program] (including any incentive funds, program information, energy audits, technical assistance, and any other support) would your organization have completed the whole project…	Percent of respondents	Percent of sites
About the same time or earlier than you did	10%	2%
At least a year later than you did	10%	2%
At least two years later than you did	10%	20%
At least 3 years later than you did	50%	59%
Or never	20%	17%

NTGR results by PA

Table A-6. Electricity NTGR by PA

PA	Projects	Customers	First year		Lifecycle	
			NTGR	Relative precision	NTGR	Relative precision
PG&E	12	4	75.4%	±8.0%	75.5%	±4.0%
SCE	24	2	99.4%	±0.0%	99.3%	±0.0%
SoCalREN	2	2	63.9%	±69.0%	66.2%	±68.0%
SDG&E	2	1	50.7%	±0.0%	50.7%	±0.0%
Total	**40**	**9**	**76.6%**	**±8.0%**	**78.1%**	**±4.0%**

Table A-7. Peak demand NTGR by PA

PA	Projects	Customers	First year	
			NTGR	Relative precision
PG&E	12	4	76.9%	±12.0%
SCE	18	1	100.0%	±0.0%
SoCalREN	1	1	86.1%	±0.0%
SDG&E	2	1	50.7%	±0.0%
Total	**33**	**7**	**75.2%**	**±8.0%**

Table A-8. Natural gas NTGR by PA

PA	Projects	Customers	First year		Lifecycle	
			NTGR	Relative precision	NTGR	Relative precision
PG&E	5	4	76.0%	±7.0%	76.1%	±4.0%

NTGR results by project review status

Table A-9. Electricity NTGR by project review status

Project Review	Projects	Customers	First year		Lifecycle	
			NTGR	Relative precision	NTGR	Relative precision
Yes	17	7	76.3%	±8.0%	76.7%	±4.0%
No	23	5	77.5%	±19.0%	84.8%	±15.0%
Total	**40**	**9**	**76.6%**	**±8.0%**	**78.1%**	**±4.0%**

Table A-10. Peak demand NTGR by project review status

Project Review	Projects	Customers	First year	
			NTGR	Relative precision
Yes	13	6	74.0%	±9.0%
No	20	3	80.5%	±20.0%
Total	**33**	**7**	**75.2%**	**±8.0%**

Table A-11. Natural gas NTGR by project review status

Project Review	Projects	Customers	First year		Lifecycle	
			NTGR	Relative precision	NTGR	Relative precision
Yes	3	2	74.7%	±7.0%	75.8%	±4.0%
No	2	2	77.6%	±6.0%	77.2%	±8.0%
Total	**5**	**4**	**76.0%**	**±7.0%**	**76.1%**	**±4.0%**

APPENDIX B. NET-TO-GROSS SURVEY INSTRUMENT

NTG INTERVIEW GUIDE FOR SITE-LEVEL NMEC PROGRAM PARTICIPANTS BASIC RIGOR LEVEL

Interview Information

Interviewer		Survey Length (min)	
Completion Date			

Contact Information

Phone	
Email	

Call Tracking

Date/Time	Notes

This guide for the site-level NMEC Net-to-Gross evaluation follows the same methodology as previously used for the Group D Custom programs NTG evaluation. Changes have been made to better align with site-level NMEC program delivery. While the Custom programs evaluation addressed questions at an individual measure level, for NMEC site-level projects, the questions have to be adjusted to refer to each project as a whole. Still, the guide contains questions asking if the scope of individual measures in the project would have been different without the NMEC program. The results from this interview guide will be used to calculate NTG using the updated methodology developed approved in conjunction with this guide.

Introduction

[NOTE: THE QUESTIONS IN THIS INTERVIEW GUIDE WILL NOT NECESSARILY BE READ VERBATIM BUT MAY BE MODIFIED TO SUIT THE INTERVIEW]

1. [IF END USER CONTACT OTHER THAN LEAD END USER CONTACT IS ANSWERING THE PHONE] Hi, my name is X OF DNV. We are calling on behalf of the California Public Utilities Commission (CPUC) and [RELEVANT PA]. According to our records, the facility at [ADDRESS] participated in [PA]'s [PROGRAM] in the last three years. This program offered financial incentives for energy efficiency upgrades. Can we please speak to [LEAD END USER CONTACT] about this project?

   [IF THEY ASK WHY WE WANT TO TALK TO LEAD END USER CONTACT] We are interviewing customers that participated in [PA's PROGRAM] to gain a better understanding of how and why they decided to install energy efficiency measures through this program. By receiving financial incentives through this program, your organization agreed to participate in this follow-up study on your experiences with this program.

2. [IF LEAD END USER CONTACT IS ANSWERING THE PHONE] Hi, my name is X OF DNV. We are calling on behalf of the California Public Utilities Commission (CPUC) and [RELEVANT PA]. According to our records, the facility at [ADDRESS] made energy efficiency upgrades which received financial incentives through [PA's PROGRAM]. You are listed as the lead contact for this project. Are you able to answer questions about the decision-making process and motivations for doing this project? Any information that you provide will remain strictly confidential. We will not identify or attribute any of your comments or organization information.

    a. [IF YES, SKIP TO Q3]

    b. [IF YES, BUT THEY CAN'T DO THE INTERVIEW AT THAT TIME, SCHEDULE ANOTHER TIME, BUT ALSO REMIND THEM THAT COMPLETING THE INTERVIEW IS MANDATORY AS A CONDITION OF THEIR COMPANY RECEIVING THE FINANCIAL INCENTIVES FROM THE PROGRAM]

    c. [IF NO, AND IF CONTACT NAME WAS OBTAINED BY PA ACCOUNT REP] I was told by your account rep <ACCT REP NAME> that you were the most knowledgeable and the most involved with the decision to implement the project I just mentioned. Is that correct?

        i. [IF YES] So, we will need to interview you. Can we go ahead with this interview?

            1. [IF YES BUT IT'S NOT A CONVENIENT TIME, RESCHEDULE THE INTERVIEW AND THEN BEGIN WITH Q3]

            2. [IF NO] Just to be clear, by receiving financial incentives through this program, your organization agreed to participate in this follow-up study on your experiences with this program. Considering that information, can we go ahead with this interview?

                a. [IF YES. GO TO Q3]

                b. [IF NO] Okay, you may be contacted directly by the CPUC to complete this process.

        ii. [IF NO, OBTAIN ALTERNATE CONTACT INFO AND SCHEDULE INTERVIEW WITH NEW PERSON]

d. [IF THEY ASK HOW LONG THE INTERVIEW WILL TAKE] I will do my best to keep the survey under 30 minutes

e. [IF THEY ASK TO VERIFY WITH THE PA's THAT THE RESEARCH IS LEGITIMATE, GIVE THEM ONE OF THE FOLLOWNG PA CONTACT NAMES DEPENDING ON WHO THEIR PA IS:

PG&E {CONTACT INFO]

SCE [CONTACT INFO]

SCG [CONTACT INFO]

SoCalREN [CONTACT INFO]

Confirmation Of Correct Respondent, Project Background

3. [A1] According to our records, your company implemented energy efficiency improvements at <%ADDRESS> in the last three years, correct?

    a. [IF YES, SKIP TO Q4]

b. [IF NO, MARK ANY CORRECTED INFORMATION IN THE FOLLOWING TABLE]

Project Information	Information from Tracking Data (pre-entered)	Corrected information (if relevant)
Measure		
Address		
Install Date		

4. [C3] What was your specific role and involvement in the project?
5. [C3a] Which other contacts were involved in moving this project forward?
    a. [OBTAIN NAME[S] AND CONTACT INFO]
6. [A1b] [IF RELEVANT] Our records show that your organization implemented or installed more than one energy efficiency measure as part of a project through the <%PA>'s <%PROGRAM> Program. Was the decision-making process for the implementation and installation of these measures made at the whole project level? Or was there a separate decision-making process for each measure?
    a. [DECISION-MAKING DONE AT PROJECT LEVEL]
    b. [SEPARATE DECISION-MAKING PROCESS FOR EACH MEASURE]
    c. [Don't know]
    d. [Refused]

Timing Questions

For the rest of this interview, we will refer to all the upgrades we just discussed, collectively, as "the project."

7. [A5b_2] Who first brought this project to your company's consideration? [DON'T READ LIST BELOW, GET COMPANY NAMES AND THEN POST CODE]
    a. Program staff
    b. Utility staff
    c. Non-program vendor
    d. Customer staff
    e. Other
8. [A5b] About when was this project first put forward for your company's consideration?
    a. [RECORD YEAR AT MINIMUM, QUARTER/MONTH GRANULARITY IF THEY CAN] ___
    b. [Don't know]
    c. [Refused]

9. [A5d] About when did your organization first begin discussions with [RELEVANT PA] regarding funding/incentives and/or possible technical assistance for this project?

    a. [RECORD YEAR AT MINIMUM, QUARTER/MONTH GRANULARITY IF THEY CAN]

    b. [Don't know]

    c. [Refused]

10. [N2] Was the decision to do this project made before or after you began discussions with [RELEVENT IMPLEMENTER] regarding the availability of incentives or technical assistance for this project? [IF NEEDED: technical assistance may include identifying the project, providing project financials, technical support during contractor selection, installation and/or project commissioning, or verification of energy savings]

    a. Before

    b. After

    c. Don't know

    d. Refused

11. **CONSISTENCY CHECK** [IF INDICATED PROJECT WAS FIRST PUT FORWARD BY THE PROGRAM BUT THAT THEY MADE THE DECISION TO DO THE PROJECT BEFORE LEARNING ABOUT INCENTIVES OR TECHNICAL ASSISTANCE FOR THIS PROJECT] Earlier you indicated that the project was first put forward to your organization by [A5b_2 response]. However, in your response to the previous question, you indicated the decision to go ahead with the project was made before learning about the incentives or technical assistance for [RELEVENT IMPLEMENTER]. To clarify, technical assistance from [RELEVENT IMPLEMENTER] can include presenting you with a project that your organization previously had not considered. Should I revise your answer about the timing of your decision to go ahead with the project before or after learning about incentives or technical assistance from [RELEVENT IMPLEMENTER]?

12. [N5b] If the program elements had not been available, including incentives or other types of program assistance, what is the likelihood that you would have conducted the project at the same time as you did? [GIVE RESPONDENT OPTIONS OF: VERY LIKELY, SOMEWHAT LIKELY, NEITHER LIKELY NOR UNLIKELY, SOMEWHAT UNLIKELY, VERY UNLIKELY]

13. Without the assistance received from the [PROGRAM] (including any incentive funds, program information, energy audits, technical assistance, and any other support) would your organization have completed the whole project…:

    a. About the same time or earlier than you did

    b. About a year later than you did

    c. About two years later than you did

    d. 3 or more years later than you did

    e. Or never

    f. Don't know

14. [IF A5b_2 = a or b] Were you aware of the specific opportunities included in this project before [A5b_2 response] suggested the project?

    a. [IF YES] Which parts of the project did you already know were potential opportunities?

Scope Questions

15. Did the project increase energy efficiency beyond what was required by code?

    a. Yes

        i. [IF YES] Was any of the increase in efficiency beyond code due to the program?

    b. No [SKIP QUESTIONS 22 and 23]

    c. Don't know [SKIP QUESTIONS 22 and 23]

    d. Not applicable [SKIP QUESTIONS 22 and 23]

16. [N5] Using a likelihood scale from 0 to 10, where 0 is "Not at all likely" and 10 is "Extremely likely", if elements of this PROGRAM had not been available, including incentives or other types of program assistance, what is the likelihood that you would have implemented the exact same project that you did? Record 0 to 5 score (_____)

Actions taken without the program

17. [IF PROJECT HAD CAPITAL MEASURES] [Revised N6] I would like you to think about what action you would have taken if the program had not been available in regards to the capital improvements that were part of this project [IF NEEDED: LIST CAPITAL MEASURES]. Which of the following alternatives would you have been MOST likely to do:

    a. Install the same equipment

    b. Continue to operate the existing equipment as-is [IF NEEDED: done nothing]

    c. Install fewer units

    d. Repair existing equipment

    e. Installed equipment with a different efficiency level

    f. Something else (specify what _____)

18. On a 0-to-10 scale of likelihood, how likely it is that you would have implemented alternative action if you had not installed the program qualifying equipment?

19. [IF HAD BRO MEASURES] I would like you to think about what action you would have taken if the program had not been available in regards to [LIST PROJECT'S BRO MEASURES] improvements that were part of this project. Which of the following alternatives would you have been MOST likely to do:

    a. Implement the same measures

    b. Continue operating as it was without implementing the measures

    c. Something else (specify what _____)

20. [A3a] Has your organization done similar types of projects at this or other California locations in the past?

    a. [IF YES] Which programs?

    b. [A3a] [IF YES] What measures were included in those projects?

    c. [A3aa] [IF YES] Did your experience participating in [RELEVANT PA PROGRAM] have any impact on your installation decision for these similar projects?

        i. [IF YES] What impacts did they have?

Influence Questions

Now I'm going to ask you a few questions about possible factors that may have influenced your decision to go ahead with this project.

I'm going to ask you to rate the importance of the program as well as other factors that might have influenced your decision to conduct this project. Think of the degree of importance as being shown on a scale with equally spaced units from 0 to 10, where 0 means not at all important and 10 means extremely important, so that an importance rating of 8 shows twice as much influence as a rating of 4.

21. [N3] Now, using this 0-to-10 rating scale, where 0 means "Not at all important" and 10 means "Extremely important," please rate the importance of each of the following **in your decision to implement this project at the time you did. [ASK IN RANDOM ORDER]**

22. [IF WENT ABOVE CODE, N3a] Now I'd like to ask you the same question but thinking only about the portions of your project up to code requirements, using this 0-to-10 rating scale, where 0 means "Not at all important" and 10 means "Extremely important," please rate the importance of each of the following in your decision to implement the portion of the project to bring measures up to code **at the time you did. [ASK IN RANDOM ORDER]**

23. [IF WENT ABOVE CODE, N3a] Lastly, I'd like to ask you the same question but thinking only about the portions of your project that went above code requirements, using this 0-to-10 rating scale, where 0 means "Not at all important" and 10 means "Extremely important," please rate the importance of each of the following in your decision to implement the portions of the project that went above code **at the time you did. [ASK IN RANDOM ORDER]**

Factor	Scoring (Overall)	Scoring (To Code)	Scoring (Above Code)
b. [N3b] The availability of the [RELEVANT PA PROGRAM] incentives, financing, or performance payments.	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused
c**. [N3c] [IF RELEVANT] Information provided through program-provided feasibility study, the facility or system energy audit, or the technical assistance provided**	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused
j. [N3j] Company practices such as to comply with company policies or goals, maintenance/replacement schedules, and regulatory requirements	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused IF > 5 What specific company policies or practices influenced your decision to do the project?	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused
n. [N3n] Payback or return on the PROJECT without factoring in program incentives	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused	# Record 0 to 10 score (_____). -97. Don't know -98. Not applicable -99. Refused

Comparing Program Influences to Non-Program Influences

Next, I would like you to rate the importance of the PROGRAM in your decision to do this project as opposed to other non-program factors that may have influenced your decision such as...

Non-Program Factor	Non-Program Factors for Which They Gave an Influence Score of >=8
j. [N3j] Standard practice at your organization such as to comply with company policies or goals, maintenance/replacement schedules, and regulatory requirements	
n. [N3n] Payback or return on the PROJECT without factoring in program incentives	

As another reminder, earlier you mentioned some program-related factors that said were important including: [ONLY CITE EXAMPLES WHERE THE INFLUENCE SCORE WAS >= 8 (MIDDLE COLUMN OF FOLLOWING TABLE)]

Program Factors	Program Factors for Which They Gave an Influence Score of >=8
15b. [N3b] The availability of the [RELEVANT PA PROGRAM] incentives	
15c. [N3c] Information provided through PA-provided feasibility study, the facility or system energy audit, or the technical assistance provided	

24. [N41] If you were given 10 points to award in total, how many points would give to the importance of the program and how many points would you give to these other factors?

    a. [N41] How many of the 10 points would you give to the importance of the PROGRAM in your decision? # _____rating of the importance of PROGRAM

    b. [N42] And how many points would you give to these other factors? # _____rating of the importance of all Other Factors

Process Evaluation Battery

25. [PP1] What do you believe the PROGRAM'S primary strengths are?

26. [PP2] What concerns do you have about the PROGRAM, if any? (IF NEEDED: What do you view as the primary features that need to be improved?)

27. [PP4] On a scale of 0 - 10, where 0 is completely dissatisfied and 10 is completely satisfied, how would you rate your OVERALL satisfaction with the PROGRAM?

    a. [IF SATISFACTION RATING LESS THAN 7] Why do you say that?

Firmographics

And finally, I have a few questions about the characteristics of your business.

28. [CC12A, CC12B] In about what year was this [business/organization/entity] established at this location? Would you say it was…

    a. After 2020
    b. In the 2010s
    c. In the 2000s
    d. In the 1990s
    e. In the 1980s
    f. In the 1970s
    g. In the 1960s or
    h. Before 1960
    i. Don't know
    j. Refused

29. [CO] About what percentage of your operating costs does energy account for? PAUSE....Would you say...

    a. Less than 1 percent
    b. 1 to 2 percent
    c. 3 to 5 percent
    d. 6 to 10 percent
    e. 11 to 15 percent
    f. 16 to 20 percent
    g. 21 to 50 percent
    h. Over 51 percent
    i. Don't know
    j. Refused

30. [CCC1, CCC3] How many square feet of heated or cooled floor area is this facility? PAUSE.... Would you say that the heated or cooled floor area is…

    a. 1,500 sq feet or less
    b. 1,500 to 5,000 sq ft
    c. 5,001 to 10,000 sq ft
    d. 10,001 to 25,000 sq ft
    e. 25,001 to 50,000 sq ft
    f. 50,001 to 75,000 sq ft
    g. 75,001 to 100,000 sq ft
    h. Over 100,000 sq ft
    i. Don't know
    j. Refused

31. [C1] What is the main activity conducted at this facility?

32. [C3] Approximately how many people are currently working at the facility where the project was conducted, including both full and part time?

    a. Ten or less
    b. Between 11 and 25
    c. 26 to 50
    d. 51 to 75
    e. 76 to 100
    f. 101 to 250
    g. 251 to 500
    h. 501 to 1000
    i. 1001 to 2500
    j. 2501 to 5000 or
    k. 5000 or more
    l. Don't know
    m. Refused

33. [C4] Does your [business/organization/entity] own, lease or manage this facility?

    a. Own

    b. Lease/Rent

    c. Manage

    d. Don't know

    e. Refused

34. [C5] How many locations does your organization have? Is it....

    a. 1

    b. 2 to 4

    c. 5 to 10

    d. 11 to 25

    e. Over 25

    f. Don't know

    g. Refused

That's all the questions I had. Thank you so much for your time.

APPENDIX C. STANDARD HIGH-LEVEL SAVINGS TABLES

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	93,590	93,590	1.00	100.0%	
PY23_NMEC_site	PGE	Total	93,590	93,590	1.00	100.0%	
PY23_NMEC_site	SCE	SLNMEC Gross-only	9,977	11,095	1.11	0.0%	1.11
PY23_NMEC_site	SCE	SLNMEC Net-only	26,129	26,129	1.00	100.0%	
PY23_NMEC_site	SCE	Total	36,106	37,225	1.03	72.4%	1.11
PY23_NMEC_site	SCR	SLNMEC Gross-only	1,557	3,301	2.12	0.0%	2.12
PY23_NMEC_site	SCR	SLNMEC Net-only	2,627	2,627	1.00	100.0%	
PY23_NMEC_site	SCR	Total	4,185	5,929	1.42	62.8%	2.12
PY23_NMEC_site	SDGE	SLNMEC Net-only	4,081	4,081	1.00	100.0%	
PY23_NMEC_site	SDGE	Total	4,081	4,081	1.00	100.0%	
PY23_NMEC_site		*Statewide*	*137,961*	*140,824*	*1.02*	*91.6%*	*1.25*

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	89,593	82,570	0.92	0.0%	0.96	0.88	0.96	0.88
PY23_NMEC_site	**PGE**	**Total**	**89,593**	**82,570**	**0.92**	**0.0%**	**0.96**	**0.88**	**0.96**	**0.88**
PY23_NMEC_site	SCE	SLNMEC Gross-only	9,977	11,095	1.11	100.0%	1.00	1.00		
PY23_NMEC_site	SCE	SLNMEC Net-only	26,129	27,311	1.05	0.0%	1.00	1.05	1.00	1.05
PY23_NMEC_site	**SCE**	**Total**	**36,106**	**38,406**	**1.06**	**27.6%**	**1.00**	**1.03**	**1.00**	**1.05**
PY23_NMEC_site	SCR	SLNMEC Gross-only	1,229	3,301	2.69	100.0%	0.79	1.00		
PY23_NMEC_site	SCR	SLNMEC Net-only	2,627	1,997	0.76	0.0%	1.00	0.76	1.00	0.76
PY23_NMEC_site	**SCR**	**Total**	**3,856**	**5,298**	**1.37**	**31.9%**	**0.92**	**0.89**	**1.00**	**0.76**
PY23_NMEC_site	SDGE	SLNMEC Net-only	4,081	2,272	0.56	0.0%	1.00	0.56	1.00	0.56
PY23_NMEC_site	**SDGE**	**Total**	**4,081**	**2,272**	**0.56**	**0.0%**	**1.00**	**0.56**	**1.00**	**0.56**
PY23_NMEC_site		***Statewide***	***133,636***	***128,547***	***0.96***	***8.4%***	***0.97***	***0.91***	***0.97***	***0.90***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	6.8	6.8	1.00	100.0%	
PY23_NMEC_site	**PGE**	**Total**	**6.8**	**6.8**	**1.00**	**100.0%**	
PY23_NMEC_site	SCE	SLNMEC Gross-only	2.9	3.0	1.03	0.0%	1.03
PY23_NMEC_site	SCE	SLNMEC Net-only	0.0	0.0			
PY23_NMEC_site	**SCE**	**Total**	**2.9**	**3.0**	**1.03**	**0.0%**	**1.03**
PY23_NMEC_site	SCR	SLNMEC Gross-only	-1.7	0.4	-0.22	0.0%	-0.22
PY23_NMEC_site	SCR	SLNMEC Net-only	0.2	0.2	1.00	100.0%	
PY23_NMEC_site	**SCR**	**Total**	**-1.5**	**0.6**	**-0.38**	**-13.1%**	**-0.22**
PY23_NMEC_site	SDGE	SLNMEC Net-only	1.3	1.3	1.00	100.0%	
PY23_NMEC_site	**SDGE**	**Total**	**1.3**	**1.3**	**1.00**	**100.0%**	
PY23_NMEC_site		***Statewide***	***9.4***	***11.6***	***1.23***	***87.8%***	***2.91***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	6.4	6.1	0.96	0.0%	0.94	0.90	0.94	0.90
PY23_NMEC_site	**PGE**	**Total**	**6.4**	**6.1**	**0.96**	**0.0%**	**0.94**	**0.90**	**0.94**	**0.90**
PY23_NMEC_site	SCE	SLNMEC Gross-only	2.9	3.0	1.03	100.0%	1.00	1.00		
PY23_NMEC_site	SCE	SLNMEC Net-only	0.0	0.0						
PY23_NMEC_site	**SCE**	**Total**	**2.9**	**3.0**	**1.03**	**100.0%**	**1.00**	**1.00**		
PY23_NMEC_site	SCR	SLNMEC Gross-only	-1.7	0.4	-0.22	100.0%	1.02	1.00		
PY23_NMEC_site	SCR	SLNMEC Net-only	0.2	0.2	1.01	0.0%	1.00	1.01	1.00	1.01
PY23_NMEC_site	**SCR**	**Total**	**-1.5**	**0.6**	**-0.37**	**112.9%**	**1.02**	**1.00**	**1.00**	**1.01**
PY23_NMEC_site	SDGE	SLNMEC Net-only	1.3	0.7	0.56	0.0%	1.00	0.56	1.00	0.56
PY23_NMEC_site	**SDGE**	**Total**	**1.3**	**0.7**	**0.56**	**0.0%**	**1.00**	**0.56**	**1.00**	**0.56**
PY23_NMEC_site		***Statewide***	***9.0***	***10.4***	***1.15***	***12.5%***	***0.95***	***0.89***	***0.95***	***0.85***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	1,927	1,927	1.00	100.0%	
PY23_NMEC_site	**PGE**	**Total**	**1,927**	**1,927**	**1.00**	**100.0%**	
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0			
PY23_NMEC_site	SCE	SLNMEC Net-only	-1	-1	1.00	100.0%	
PY23_NMEC_site	**SCE**	**Total**	**-1**	**-1**	**1.00**	**100.0%**	
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0			
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0			
PY23_NMEC_site	**SCR**	**Total**	**0**	**0**			
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0			
PY23_NMEC_site	**SDGE**	**Total**	**0**	**0**			
PY23_NMEC_site		***Statewide***	***1,927***	***1,927***	***1.00***	***100.0%***	

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	1,844	1,716	0.93	0.0%	0.96	0.89	0.96	0.89
PY23_NMEC_site	**PGE**	**Total**	**1,844**	**1,716**	**0.93**	**0.0%**	**0.96**	**0.89**	**0.96**	**0.89**
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0						
PY23_NMEC_site	SCE	SLNMEC Net-only	-1	-1	1.00	0.0%	1.00	1.00	1.00	1.00
PY23_NMEC_site	**SCE**	**Total**	**-1**	**-1**	**1.00**	**0.0%**	**1.00**	**1.00**	**1.00**	**1.00**
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0						
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0						
PY23_NMEC_site	**SCR**	**Total**	**0**	**0**						
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0						
PY23_NMEC_site	**SDGE**	**Total**	**0**	**0**						
PY23_NMEC_site		***Statewide***	***1,844***	***1,715***	***0.93***	***0.0%***	***0.96***	***0.89***	***0.96***	***0.89***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	11,623	11,623	1.00	100.0%	
PY23_NMEC_site	**PGE**	**Total**	**11,623**	**11,623**	**1.00**	**100.0%**	
PY23_NMEC_site	SCE	SLNMEC Gross-only	846	917	1.08	0.0%	1.08
PY23_NMEC_site	SCE	SLNMEC Net-only	2,069	2,069	1.00	100.0%	
PY23_NMEC_site	**SCE**	**Total**	**2,914**	**2,986**	**1.02**	**71.0%**	**1.08**
PY23_NMEC_site	SCR	SLNMEC Gross-only	112	202	1.81	0.0%	1.81
PY23_NMEC_site	SCR	SLNMEC Net-only	787	787	1.00	100.0%	
PY23_NMEC_site	**SCR**	**Total**	**899**	**990**	**1.10**	**87.5%**	**1.81**
PY23_NMEC_site	SDGE	SLNMEC Net-only	816	816	1.00	100.0%	
PY23_NMEC_site	**SDGE**	**Total**	**816**	**816**	**1.00**	**100.0%**	
PY23_NMEC_site		***Statewide***	***16,253***	***16,415***	***1.01***	***94.1%***	***1.17***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	11,219	10,255	0.91	0.0%	0.97	0.88	0.97	0.88
PY23_NMEC_site	**PGE**	**Total**	**11,219**	**10,255**	**0.91**	**0.0%**	**0.97**	**0.88**	**0.97**	**0.88**
PY23_NMEC_site	SCE	SLNMEC Gross-only	846	917	1.08	100.0%	1.00	1.00		
PY23_NMEC_site	SCE	SLNMEC Net-only	2,069	2,162	1.05	0.0%	1.00	1.05	1.00	1.05
PY23_NMEC_site	**SCE**	**Total**	**2,914**	**3,079**	**1.06**	**29.0%**	**1.00**	**1.03**	**1.00**	**1.05**
PY23_NMEC_site	SCR	SLNMEC Gross-only	88	202	2.31	100.0%	0.78	1.00		
PY23_NMEC_site	SCR	SLNMEC Net-only	787	598	0.76	0.0%	1.00	0.76	1.00	0.76
PY23_NMEC_site	**SCR**	**Total**	**875**	**801**	**0.92**	**10.0%**	**0.97**	**0.81**	**1.00**	**0.76**
PY23_NMEC_site	SDGE	SLNMEC Net-only	816	454	0.56	0.0%	1.00	0.56	1.00	0.56
PY23_NMEC_site	**SDGE**	**Total**	**816**	**454**	**0.56**	**0.0%**	**1.00**	**0.56**	**1.00**	**0.56**
PY23_NMEC_site		***Statewide***	***15,825***	***14,589***	***0.92***	***5.9%***	***0.97***	***0.89***	***0.97***	***0.88***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	1.4	1.4	1.00	100.0%	
PY23_NMEC_site	**PGE**	**Total**	**1.4**	**1.4**	**1.00**	**100.0%**	
PY23_NMEC_site	SCE	SLNMEC Gross-only	0.2	0.3	1.03	0.0%	1.03
PY23_NMEC_site	SCE	SLNMEC Net-only	0.1	0.1	1.00	100.0%	
PY23_NMEC_site	**SCE**	**Total**	**0.4**	**0.4**	**1.02**	**37.7%**	**1.03**
PY23_NMEC_site	SCR	SLNMEC Gross-only	-0.1	0.0	-0.25	0.0%	-0.25
PY23_NMEC_site	SCR	SLNMEC Net-only	0.0	0.0	1.00	100.0%	
PY23_NMEC_site	**SCR**	**Total**	**-0.1**	**0.1**	**-1.11**	**-68.7%**	**-0.25**
PY23_NMEC_site	SDGE	SLNMEC Net-only	0.3	0.3	1.00	100.0%	
PY23_NMEC_site	**SDGE**	**Total**	**0.3**	**0.3**	**1.00**	**100.0%**	
PY23_NMEC_site		***Statewide***	***2.0***	***2.1***	***1.08***	***93.5%***	***2.20***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	1.3	1.2	0.93	0.0%	0.97	0.90	0.97	0.90
PY23_NMEC_site	**PGE**	**Total**	**1.3**	**1.2**	**0.93**	**0.0%**	**0.97**	**0.90**	**0.97**	**0.90**
PY23_NMEC_site	SCE	SLNMEC Gross-only	0.2	0.3	1.03	100.0%	1.00	1.00		
PY23_NMEC_site	SCE	SLNMEC Net-only	0.1	0.2	1.05	0.0%	1.00	1.05	1.00	1.05
PY23_NMEC_site	**SCE**	**Total**	**0.4**	**0.4**	**1.04**	**62.3%**	**1.00**	**1.02**	**1.00**	**1.05**
PY23_NMEC_site	SCR	SLNMEC Gross-only	-0.1	0.0	-0.25	100.0%	1.02	1.00		
PY23_NMEC_site	SCR	SLNMEC Net-only	0.0	0.0	1.01	0.0%	1.00	1.01	1.00	1.01
PY23_NMEC_site	**SCR**	**Total**	**-0.1**	**0.1**	**-1.09**	**166.8%**	**1.03**	**1.00**	**1.00**	**1.01**
PY23_NMEC_site	SDGE	SLNMEC Net-only	0.3	0.1	0.56	0.0%	1.00	0.56	1.00	0.56
PY23_NMEC_site	**SDGE**	**Total**	**0.3**	**0.1**	**0.56**	**0.0%**	**1.00**	**0.56**	**1.00**	**0.56**
PY23_NMEC_site		***Statewide***	***1.9***	***1.9***	***0.98***	***6.5%***	***0.98***	***0.88***	***0.98***	***0.87***

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	291	291	1.00	100.0%	
PY23_NMEC_site	**PGE**	**Total**	**291**	**291**	**1.00**	**100.0%**	
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0			
PY23_NMEC_site	SCE	SLNMEC Net-only	0	0	1.00	100.0%	
PY23_NMEC_site	**SCE**	**Total**	**0**	**0**	**1.00**	**100.0%**	
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0			
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0			
PY23_NMEC_site	**SCR**	**Total**	**0**	**0**			
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0			
PY23_NMEC_site	**SDGE**	**Total**	**0**	**0**			
PY23_NMEC_site		***Statewide***	***291***	***291***	***1.00***	***100.0%***	

Table C-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
PY23_NMEC_site	PGE	SLNMEC Net-only	282	259	0.92	0.0%	0.97	0.89	0.97	0.89
PY23_NMEC_site	**PGE**	**Total**	**282**	**259**	**0.92**	**0.0%**	**0.97**	**0.89**	**0.97**	**0.89**
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0						
PY23_NMEC_site	SCE	SLNMEC Net-only	0	0	1.00	0.0%	1.00	1.00	1.00	1.00
PY23_NMEC_site	**SCE**	**Total**	**0**	**0**	**1.00**	**0.0%**	**1.00**	**1.00**	**1.00**	**1.00**
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0						
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0						
PY23_NMEC_site	**SCR**	**Total**	**0**	**0**						
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0						
PY23_NMEC_site	**SDGE**	**Total**	**0**	**0**						
PY23_NMEC_site		***Statewide***	***282***	***259***	***0.92***	***0.0%***	***0.97***	***0.89***	***0.97***	***0.89***

APPENDIX D. STANDARD PER-UNIT SAVINGS TABLES

Table D-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
PY23_NMEC_site	PGE	SLNMEC Net-only	1	0.0%		5.5	1,799,803.0	223,523.0	223,523.0
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0.0%	0.0%	11.8	15.8	1.3	1.3
PY23_NMEC_site	SCE	SLNMEC Net-only	1	0.0%		13.1	1,045,168.7	82,749.1	82,749.1
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0.0%	0.0%	13.3	4.7	0.3	0.3
PY23_NMEC_site	SCR	SLNMEC Net-only	1	0.0%		3.4	1,313,676.7	393,651.9	393,651.9
PY23_NMEC_site	SDGE	SLNMEC Net-only	1	0.0%		12.0	2,040,250.7	408,050.1	170,020.9

Table D-2. 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
PY23_NMEC_site	PGE	SLNMEC Net-only	1	0.0%		5.5	37,065.9	5,588.1	5,588.1
PY23_NMEC_site	SCE	SLNMEC Gross-only	0	0.0%	0.0%	11.8	0.0	0.0	0.0
PY23_NMEC_site	SCE	SLNMEC Net-only	1	0.0%		13.1	-28.8	-2.5	-2.5
PY23_NMEC_site	SCR	SLNMEC Gross-only	0	0.0%	0.0%	13.3	0.0	0.0	0.0
PY23_NMEC_site	SCR	SLNMEC Net-only	1	0.0%		3.4	0.0	0.0	0.0
PY23_NMEC_site	SDGE	SLNMEC Net-only	1	0.0%		12.0	0.0	0.0	0.0

Table D-3. 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
PY23_NMEC_site	PGE	SLNMEC Net-only	0	0.0%	0.0%	5.5	1,587,893.9	197,205.4	197,205.4
PY23_NMEC_site	SCE	SLNMEC Net-only	0	0.0%	0.0%	13.1	1,092,429.4	86,490.8	86,490.8
PY23_NMEC_site	SCE	SLNMEC Gross-only	1	0.0%		11.8	15.8	1.3	1.3
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0.0%	0.0%	3.4	998,394.3	299,175.4	299,175.4
PY23_NMEC_site	SCR	SLNMEC Gross-only	1	0.0%		13.3	4.7	0.3	0.3
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0.0%	0.0%	12.0	1,136,011.6	227,202.3	94,667.6

Table D-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
PY23_NMEC_site	PGE	SLNMEC Net-only	0	0.0%	0.0%	5.5	32,992.9	4,974.0	4,974.0
PY23_NMEC_site	SCE	SLNMEC Net-only	0	0.0%	0.0%	13.1	-28.8	-2.5	-2.5
PY23_NMEC_site	SCE	SLNMEC Gross-only	1	0.0%		11.8	0.0	0.0	0.0
PY23_NMEC_site	SCR	SLNMEC Net-only	0	0.0%	0.0%	3.4	0.0	0.0	0.0
PY23_NMEC_site	SCR	SLNMEC Gross-only	1	0.0%		13.3	0.0	0.0	0.0
PY23_NMEC_site	SDGE	SLNMEC Net-only	0	0.0%	0.0%	12.0	0.0	0.0	0.0

APPENDIX E. SITE-LEVEL NMEC EVALUATION REPORT COMPILED STAKEHOLDER COMMENTS

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

Table E-1. Responses to comments on draft report

Commenter	Comment	DNV Response
PG&E	[Page 16-23] Based on our experience with pay-for-performance (P4P) programming, we've found that traditional net-to-gross methodologies may not be the most suitable approach for SLNMEC as a performance-based program delivery model. Traditional evaluations typically occur after program completion, whereas P4P programs benefit from ongoing, real-time feedback. To better align with the goals of both policy and program implementation, we recommend that the CPUC convene an NMEC evaluation working group. This group should collaborate to develop a mutually acceptable alternative evaluation method that supports continuous improvement and evolution of the program.	Noted.
PG&E	[General] We recommend including a roadmap in the report that outlines the upcoming research, clearly identifying the gaps it aims to address and detailing a plan for integrating its findings in a timely and effective manner.	Thank you for your comment. The Executive summary (p. 1) describes the forthcoming additional research report that makes up the second part of the PY2023 evaluation and will explore several additional research questions. Those research questions are detailed more fully in the PY2023 site-level NMEC workplan, which is cited in this current report and available in the Energy Division's Public Document Area: https://pda.energydataweb.com/api/view/4101/PY2023%20CPUC%20Site-level%20NMEC%20Evaluation%20Workplan%20-%20Final.pdf

Commenter	Comment	DNV Response
PG&E	[Page 16: "It benefits our evaluation to complete NTG interviews as close to the time of decision making as possible, increasing the likelihood of accurate recall and reducing the likelihood of decision maker turnover."] Recommendation 1: Include details on sample attribution, for example, how many decisionmakers are no longer at the site? Include details on protocol for still attempting to achieve a valid sample. i.e., did the evaluator receive any assistance from PAs and implementer to engage with the customer? Expanded on in Section 3.2.2. Recommendation 2: The evaluator should consider recommending embedded evaluation approaches so that influence and NTG questions are included in customer onboarding and forwarded to the evaluation team at the point of project decision making, rather than relying on customer good-will and recall for a retrospective evaluation. Embedded evaluation combined with M&V 2.0 will improve the accuracy, speed, and cost-effectiveness of traditional evaluations of pay-for-performance programs.	NTG interviews are completed following the initial claim to ensure that interviews are conducted as soon as possible after the decision has been made. The evaluation was able to complete interviews for 90% of the projects (43 projects). Please see section 3.1 for more details on the sample, Appendix A for more NTG results, and Appendix B to see the NTG survey instrument, including the initial screener questions. We do follow a protocol for reaching out to the PAs for assistance when we are having difficulty recruiting a site. Additionally, where available, we do review and reference any provided program influence documentation. Please see the workplan for more information about the influence documentation work that is included in the forthcoming additional research question report.

Commenter	Comment	DNV Response
PG&E	[Page 17: “We reviewed the measure-level EULs provided in the documentation and investigated the sources of those EULs.”] Recommendation: Adding a guidance and/or sources for acceptable EUL sources will enable future consistency.	Thank you for your comment. This passage simply describes our assessment of the presence and reasonableness of documentation to support the EULs. The conclusions and recommendations section of the report speaks more directly to EUL sources. The associated finding was, "Projects used Effective Useful Life (EUL) values without providing justification." The recommendation is, "Measure documentation should include a description of the measure, its EUL, and its respective DEER EUL ID (or other similar citation) to justify each measure's EUL."

Commenter	Comment	DNV Response
PG&E	[Page 18: “For each project we replicated the baseline model, the performance model, the normalization of baseline and performance consumption, and the calculation of normalized savings.”] Depending on the modeling methodololgy (i.e. NMECR), the version of the modeling software may result in differing outputs. Is there an acceptable threshold for differing models? Recommendation: We recommend using an alternative calculation methodology to obtain model regression results if the original methodology cannot be carried out. We request clarification of the purpose of model replication as was done for model validation in section 3.2.3.2.	**Replication purpose** The NMEC rulebook specifically calls for model replication, whihc is the most basic test to ensure that the claim and documentation are aligned. Page 18 of the rulebook states that "upon request, the underlying participant consumption data and other data inputs must be made available to the PAs as well as the Commission and its impact evaluators such that savings calculations can be replicated to reach the same result." If the provided model cannot be used to replicate the claimed results and the documentation does not sufficiently explain the model approach to allow replication, it raises questions about whether the correct documentation was provided or whether there were errors in the claim. **Replication threshold** As we have generally been able to come close to the documented normalized savings, allowing us to advance to validation, we have not seen a need to apply a replication threshold. It is generally good practice to document what verision of modeling software is used to reduce the likelihood of these types of issues.

Commenter	Comment	DNV Response
PG&E	[Page 18: "Finally, we considered if any essential variables or additional data were not included in the models."] When the models pass the GOF criteria, the model developers are not obligated to introduce additional indepndent variables even if doing so would improve the model performance. Can the evaluators explain the purpose of considering additional data?	When read within the context of the rest of the paragraph, this sentence is specifically refering to the engineering basis for the model. If there is evidence of a dramatic change in occupancy or some other driver, and it is not included in the models, then there is no way to normalize those values across the two models. If this change happens during the installation period, there may not be negative effects on GOF, but there could be dramatic effects on savings estimation. Some examples of additional data we might expect in these cases: - Onsite generation production data - Variables consistent with the baseline model - Information regarding significant changes at the site, such as schedule changes (going from 7 open days to 5)

Commenter	Comment	DNV Response
PG&E	[Page 19-23] Readability Comment: this section is very detailed but is difficult to read and interpret. Recommendation: Break up this section more like the gross savings methodology before it to improve flow. For example: • Customer interview o PAI1 o PAI2 o PAI3 • Examples of calculations • Adjustments to NTGR scoring Alternate recommendation: Move the examples discussed on page 30 of this section to an appendix to simplify the flow but keep the content.	Thank you for this recommendation. We added sub-headings and moved a few sentences to clarify and emphasize the structure of the net savings methods section and to mirror the explicit structure of the gross savings methods section.
PG&E	[Page 27: "While we were not able to exactly replicate the savings results for any of the three evaluated projects, two out of the three projects were within 15 kWh (less than 0.01% of normalized baseline consumption), which is relatively close."] Comment: a 15 kWh difference for a regression is well within the industry accepted range of error and uncertainty in measurement. NMEC programs should be held to the same level of rigor as other programs, and not held to an unachievably high standard for perfection. 90/10 is the most common level of acceptable error, as per ASHRAE 14, where we expect the calculated value to be within 10% of the true value 90% of the time. Going deeper than that for NMEC programs is an inappropriate use of program funds. Recommendation: Use ASHRAE and IPMVP standards for error during evaluation.	Thank you for your comment. This is a general statement that the replicated savings are, indeed, very close and did not result in a specific finding, recommendation, or savings adjustment. We do not find that specific standards for error are necessary for replication. Replication is simply the first step to ensure that we are starting the evaluation from the same point.

Commenter	Comment	DNV Response
PG&E	[Page 39: "Understanding the viability of existing equipment is important for both determining whether the project is appropriate for NMEC and for selecting the appropriate MAT."] Comment: These programs are designed to target efficiency opportunities for equipment that is currently operational and meeting customer needs. We recommend that evaluators consider incorporating embedded evaluation approaches, where feasible, to identify equipment repairs and on-site adjustments in real time. For further guidance on appropriate NMEC impact evaluation methods, please refer to our comments under the Impact Evaluation Methodology.	While the evaluation makes every effort, including the early gross evaluation which will be published this summer, to provide earlier feedback, it is not the role of the evaluator to provide real time site monitoring. We do encourage the PAs and implementers to check in regularly with the site to identify any issues in a timely manner.

Commenter	Comment	DNV Response
PG&E	[Page 42] The report identifies persistent gaps in project documentation, including: • Insufficient detail on existing and installed equipment. • Lack of justification for measure application types (MATs) and effective useful life (EUL) values. • Incomplete or missing data processing scripts, making replication of results difficult. • Missing or unclear model parameter outputs and narratives. • The report states that none of the project files included the code used to transform the raw data into prepared data or a narrative explaining modeling decisions. • In the modern modeling era it is very likely that a third-party software firm is collecting and displaying the raw data, and that the modeler is making queries from that software. It is very unlikely that different extracts will result in exactly the same data, regardless of processing narrative. Rather, it would be more in line with modern techniques for NMEC evaluation to do an M&V plan review for reasonability, and then evaluate the design matrices for a sample of projects from each implementer to ensure that the M&V plan process is being followed and there are no errors. The NMEC rulebook requires that implementers submit the modeling code, which is, as mentioned in the report, often submitted and uses a common and popular modeling software that does not do the data cleaning. Comment: We believe that full replication is not required by the CPUC nor is cost-effective. However, we agree that Independent evaluators should be able to 1) Determine if the math is correct / M&V plans are being followed, and 2) Evaluate the reasonableness of an M&V approach. Recommendation: As mentioned above, we recommend that the CPUC convene an evaluation working group for NMEC.	Project M&V plans generally do not provide a sufficient description of the modeling approach or specification to determine reasonableness or whether the calculations are correct. In order to replicate, we usually try to run the provided code to confirm we get the same values reported by the PA. When sufficient data and code are provided, replication is a simple and quick check. The M&V and final savings reports reviewed so far have not included any description of data preparation steps, which can be key sources for calculation errors. This is particularly concerning when variables beyond energy consumption and weather are used, as the evaluator is unable to independently reproduce the variables. Additionally, data preparation is explicitly not something 3rd party software generally does. It is based on decisions made by the modeler and should be documented. It would not be unreasonable for evaluation to ask for raw, pre-data-prep data to track every aspect of the implementers actions that could affect the ultimate results.

Commenter	Comment	DNV Response
PG&E	[Page 47: "Model parameters and summaries: All evaluated M&V plans included a description of baseline model variables; descriptions of non-routine events and related adjustments; and basic goodness of fit statistics. However, none of the M&V plans provided basic baseline model output and specification details, such as parameter estimates, p values, temperature bin values, and occupied/unoccupied mode settings. These values are essential to ensure successful replication efforts and to fully assess whether the model reflects accurate engineering conditions at the site."] The report claims that M&V plans need to contain “parameter estimates, p values, temperature bin values, and occupied/unoccupied mode settings.” For evaluators need these values to successfully replicate efforts and to fully assess whether the model reflects accurate engineering conditions at the site.” Recommendation: The only output the evaluators need to replicate the exact results and ensure compliance with the M&V plan is the regression design matrix and an extract of the cleaned consumption data for the independent variable and the complete matrix of regressors that were the inputs to the regression solver (design matrix). We recommend including a request for implementers to request that the designers of the software package enable the function for the software to print and submit the design matrix in spreadsheet form to allow for exact replicability and lead to consistent outputs between implementers. We acknowledge that most users are not hand writing their design matrices and writing custom modeling code. If the CPUC is allowing the use of modeling software, the industry should seek to ensure that the software contains the necessary function for the implementation and evaluation of energy efficiency projects at the level of rigor required by the CPUC. In short: open-source is not enough if features are missing, and we request that the evaluator carefully document what features should be found in modeling software for the purposes of NMEC rulebook compliance and replicability.	Parameter estimates, p values, and occupied/unoccupied mode settings should be basic output from any package and should be easily copied into documentation provided to the evaluator. Temperature bins should, generally, be obvious in the provided independent variables, but if those independent variables are not provided, then the temperature bins should be otherwise provided.

Commenter	Comment	DNV Response
SDG&E	[Overarching] SDG&E suggests that evaluators create a separate table to summarize all conclusions and recommendations, similar to the format required by the Response to Recommendations (RTR) process. This table should include the conclusion, recommendation, and the party to whom the recommendation is directed (i.e. PA vs. CPUC).	We can provide a response to recommendations spreadsheet table with the final report. Thank you for this recommendation.
SDG&E	[Overarching] There are a couple of NMEC projects in SDG&E's service territory that have since been updated to Custom.	Thank you for reminding us of this change. Switching these projects from NMEC to Custom occurred after we had completed the NTGR analysis and we missed this revision. We have updated the report to reflect this change.
SDG&E	[Overarching] SDG&E requests the evaluators to provide a table that indicates how many projects per PA were selected to calculate the statewide NTGR and GRR results. Although there is a total number of projects within the document, it would be helpful to see the total number of projects specific for the NTGR results.	Please see tables 4.1 and 4.4 for this information. Both tables show the number of projects by PA. Appendix A includes additional NTGR information by PA.
SDG&E	[Page 10] Evaluator recommendation states "Programs should continue all levels of technical support provided to customers, especially pro-active efforts to build relationships." SDG&E seeks clarification on if this recommendation is geared towards the PA or the third-party implementer who administers the program.	This recommendation would apply to the PAs and potentially to third-party implementers, depending on the particular program structure.

Commenter	Comment	DNV Response
SCE	[Page 29,30: "The increase in NTGRs from the previous evaluation to this one is the result of both actual changes in program influence as reported by this year's sample compared to last year's sample and updates to the scoring methodology."] It is reasonable to conclude that the extensivework by the NMEC Project Coordination Group enabled this refresh of the NTG methods. The group produced good background material such as PG&E's SITE-LEVEL NMEC INFLUENCE GUIDELINES FOR NON-RESIDENTIAL PROGRAMS , April 11, 2023. Perhaps if the team agrees this could be a recommendation?	The changes to the NTG methodology were made independently from the NMEC PCG, but the updates were shared and discussed with the PCG.
SCE	[Page 29,30] An open question is whether other delivery methods would also show more reliable NTG estimates if they had similar coordination groups. One could argue that 3P delivery is a fundamentally new delivery methods and could have a tailored measure of NTG.	We understand this comment to imply that NTGRs should be third-party-implementer specific rather than PA-specific. However, the PA is responsible for the quality and influence of the project work.
SCE	[Page 7] PAs should continue to focus on improving the timeliness and accuracy of site-level NMEC claims, following existing NMEC reporting guidance: make the initial claim in the quarter measures are installed, use a ProjectID that can be tracked across project years, and make the true-up claim in the quarter in which the performance period is completed. SCE will discuss recommendations at the RTR forum	Noted.
SCE	[Page 7] Measure documentation should include a description of the measure, its EUL, and its respective DEER EUL ID (or other similar citation) to justify each measure's EUL SCE will discuss recommendations at the RTR forum	Noted.

Commenter	Comment	DNV Response
	[Page 9] Even though the savings for site-level NMEC projects are meter-based, it is essential to clearly document the viability of the existing equipment, the details of the measures installed, and code compliance	
SCE	SCE will discuss recommendations at the RTR forum	Noted.
	[Page 10: "1. PAs should provide the data processing scripts or files used to prepare the data for modeling, along with the complete, final modeling scripts and files needed to exactly reproduce the savings in the project documentation and the CEDARs tracking database. 2. PAs should provide complete model output as well as a model narrative, including an explanation of any modeling decisions made, changes from the M&V plan, and an analysis of any substantial deviations in savings."]	
SCE	SCE will discuss recommendations at the RTR forum	Noted.
	[Page 10: "Programs should continue all levels of technical support provided to customers, especially pro-active efforts to build relationships."]	
SCE	SCE will discuss recommendations at the RTR forum	Noted.
	[Page 10,11: "1. In some cases, PAs should consider requesting an exception to the 18-month installation period when there are extenuating circumstances and other reasonable solutions have not worked. 2. PAs and implementers should clearly explain timeline needs early in the process, providing, for example, clear project documentation and regular check-in meetings."]	
SCE	SCE will discuss recommendations at the RTR forum	Noted.

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.