Site-Level NMEC Technical Guidance: Program M&V Plans Utilizing Normalized Metered Energy Consumption Savings Estimation

Version 2.0
Date: December 15, 2019

Prepared for Energy Division by Lawrence Berkeley National Laboratory

Authors:
Jessica Granderson, Lawrence Berkeley National Laboratory
Paula Gruendling, Energy Division, California Public Utility Commission
Christie Torok, California Public Utility Commission
Peter C. Jacobs, BuildingMetrics Inc.
Nikhil Gandhi, Strategic Energy Technologies, Inc

Contents

Purpose of this guidance:

This guidance provides direction to Program Administrators/Implementers on considerations for Measurement and Verification Plans (M&V Plans) for "Site-Level" NMEC programs that intend to use normalized metered energy consumption (NMEC) to determine energy and demand savings (as opposed to "Population-Level" approaches[1]).

Programs are referred to as **"Site-Level" NMEC** programs where the following conditions hold:

- NMEC methods are used to determine energy savings claims;
- Programs and projects meet with the regulatory and filing requirements described in the CPUC Rulebook[1];
- NMEC methods used to determine savings are customized to the particular site and project to conform to site-specific conditions and adjust for the particular drivers of savings pertinent to the customer site and project;
- Energy Savings claims and project estimates of savings are submitted for a specific site or project;
- NMEC-determined energy savings rely on a project-specific M&V plan, customized to the specific characteristics of the site and project.

The guidance has the objective of informing the M&V Plan that will support the Implementation Plan for proposed programs targeting multiples measures, and whole building gross savings approaches in the commercial sector. This guidance should inform program planning, and the development of program designs and analytical methods for Site-Level NMEC approaches. This guidance does not specify requirements or rules related to gross or net energy savings claims resulting from individual projects completed through approved programs.

This guidance is designed to address key questions that the CPUC Energy Division has identified as critical for program implementers to consider when developing Site-Level NMEC program proposals. Key questions are as follows:

- How was the baseline model chosen, what is its form, and why is it a good fit for the program?
- Do the planned baseline models characterize baseline energy use adequately for the program's target population? That is, how can baseline *models* be screened and vetted[2]?

[1] For a fuller discussion of Site-Level and Population-Level definitions and regulatory requirements, please see the CPUC Rulebook. For the most up-to-date version of the Rulebook, please see the CPUC webpage on Energy Efficiency Rolling Portfolio Program Guidance, found here: http://www.cpuc.ca.gov/general.aspx?id=6442456320

- What metrics and targets for modeling precision should be used?
- Is the plan likely to lead to an accurate, quality savings result?
- Will there be sufficient coverage factor for independent variables in the baseline period?

- How will non-routine adjustments be identified, quantified, and reported?

Scope of this guidance:

The scope of this guidance includes:

- Overall program M&V Plan that must be submitted in the Implementation Plan[3] before a program is initiated, as contrasted to the site-specific M&V plans that are submitted with individual project applications.
- Gross savings determination, excluding considerations of net savings; i.e. adjusted avoided energy use or normalized savings as defined in the International Performance Measurement and Verification Protocol IPMVP[4].

This guidance aims at providing direction for program planning. This guidance does not cover site specific M&V plans, which are expected to apply the guidance in this document at a facility-tailored level.

Accordingly, the guidance includes both qualitative and quantitative content, organized into six primary sections:

1) baseline modeling narrative;
2) baseline model goodness of fit screening;
3) scenario analysis of uncertainty due to model error;
4) coverage factor for independent variables in the baseline period;
5) treatment of non-routine adjustments;
6) savings claims.

[2] Note that model screening and vetting is distinguished from screening and selection of specific buildings for recruitment into the program. It is expected however, that the model screening findings may indeed be leveraged in subsequent building screening and participant targeting activities.

[3] Decision 15-10-028, Appendix 4.

[4] Efficiency Valuation Organization (EVO). 2016. Core concepts: International performance measurement and verification protocol. EVO 10000-1:2016.

Basis of this guidance:

This guidance is based on existing industry best practices as defined in the IPMVP, ASHRAE Guideline 14[5], and Bonneville Power Administration Reference Guides[6]. The concepts in these references are extended and complemented with recent findings from the published literature to meet California's current NMEC needs and program plans.

Background information:

The IPMVP provides a comprehensive discussion of the circumstances under which a whole-building Option C M&V approach is recommended, as well as other available Options and how to select among them, taking into account factors such as the significance of interactive effects, the need to assess multiple measures individually, stability of conditions within the measurement boundary, required duration of the performance assessment, availability of baseline data, and other key considerations. Users of this guidance are encouraged to familiarize themselves with the various approaches included in the IPMVP to ensure that an Option C approach is suitable and viable.

There are various types of regression models that are used to generate baseline models for Option C M&V applications. The most common are linear and piecewise linear models[5, 6,7,8,9], many of which have been used by M&V practitioners for decades. Today, energy modeling techniques are being developed that incorporate complex statistical regression and machine learning methods applied to higher frequency meter data[10, 11].

---

[5] ASHRAE Guideline 14 (2014). ASHRAE Guideline 14-2014 for Measurement of Energy and Demand Savings, American Society of Heating, Refrigeration and Air Conditioning Engineers, Atlanta, GA.

[6] Research Into Action, et al. Regression for M&V: Reference Guide. Report prepared for Bonneville Power Administration. Bonneville Power Administration, May 2012.

[7] Fels, M.F., 1986. PRISM: an introduction. Energy and Buildings, 9(1), pp.5-18.

[8] Kissock, J.K., Haberl, J.S. and Claridge, D.E., 2002. Development of a Toolkit for Calculating Linear, Change-Point Linear and Multiple-Linear Inverse Building Energy Analysis Models, ASHRAE Research Project 1050-RP, Final Report. Energy Systems Laboratory, Texas A&M University.

[9] Mathieu, J. L., P. N. Price, S. Kiliccote, and M. A. Piette. 2011. "Quantifying changes in building electricity use, with application to demand response." IEEE Transactions on Smart Grid, 2(3), pp. 507–518.

[10] Heo, Y. and Zavala, V.M., 2012. Gaussian process modeling for measurement and verification of building energy savings. Energy and Buildings, 53, pp.7-18.

[11] Touzani, S., Granderson, J. and Fernandes, S., 2017. Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy and Buildings, *In Press*.

1. Baseline modeling narrative

Purpose: Normalized metered energy consumption-based savings estimation plans are expected to include a modeling narrative.

Modeling narrative guidance:

M&V Plans should include a description of:

- Why an Option C M&V approach is suitable given the expected program design and scope of associated energy efficiency measures.

- The mathematical form of the model(s), e.g. piece-wise linear regression, or artificial neural network.
  - If multiple models are being considered, briefly explain how they would be chosen for application to individual projects

- The dependent variables (e.g., therms, kWh, whole building combined Btu), and the independent variables used to predict consumption; the logic for including the specified independent variables, as well as logic for excluding others

- Additional building characteristics and information on monitoring infrastructure that may be collected to inform M&V activities.

- Why the model is expected to characterize energy well for the target building and or system types it will be applied to, given the program design
  - Include a description of how the independent variables relate to the measures and systems included in the program design
  - If available, note any similar programs in which the model was used successfully, or results of independent third-party model testing

- The time resolution (hourly, daily, etc.) of input data and output predictions
  - Note that buildings that are production or process driven, e.g. restaurants, may need additional variables to characterize the processes; the frequency of those data may be a limiting factor in model type and resolution.

- Planned typical and minimum duration and characteristics of the baseline period. Baseline period should follow Commission direction[12].
  - The typical duration should be reflected in the subsequent goodness of fit screening and uncertainty scenario analyses in Sections 2 and 3.

---

[12] Assigned Commissioner and Administrative Law Judge's Ruling Regarding High Opportunity Energy Efficiency Programs or Projects (12/30/2015) affirmed by CPUC Decision D.16-08-019. Documents available at:
https://apps.cpuc.ca.gov/apex/f?p=401:56:0::NO:RP,57,RIR:P5_PROCEEDING_SELECT:R1311005

- How site verification activities will be conducted, and documented including:
    - Measure installation and operation[13].
    - How measure implementation dates will be tracked and documented to establish the baseline and reporting periods for avoided energy use and normalized savings calculations, and documentation of savings.

- How missing, erroneous, or outlier data will be handled.

- How sites will be tracked to identify site/customer participation in multiple concurrent programs.

- How the model is implemented, e.g., in a packaged tool (provide the tool name and provider name, version number), coded in R or SAS, or other implementation
    - Note whether the tool or method has undergone any validation tests
    - Fixed versus user-defined model parameters.

- The anticipated sources and format for all meter data and independent variables, and the parties that will be responsible for providing the required data.
    - How the meters used in the Option C analysis will be mapped to accounts, premises, project measurement boundaries, and loads served in the building, as well as how any on-site generation will be treated
    - There are many possible configurations of buildings, customers, and meters, and this portion of the narrative should describe how implementers and utilities will collaborate to ensure that data is available for, and collected from all relevant and impacted meters that will be used to model the baseline and post-implementation performance period.

- Whether the meters used for the Option C analysis are expected to comprise utility account meters; and, for cases where there are no utility account meters, specify the calibration process that will be used to ensure data accuracy
    - In the case that sub-meters or data from control systems are used for independent variables, specify the process for ensuring data accuracy according to the specifications in the Rulebook.

- Narrative authors may find value in the modeling concepts and best practices that are presented in references such as *Applied Regression Analysis*, *Applied Statistics and Probability for Engineers,* and more domain-specific examples such as the *BPA Energy Smart Industrial Monitoring, Targeting and Reporting (MT&R) Reference Guide*.

[13] The objectives of measure installation verification are to confirm that: (1) the measures were actually installed, (2) the installation meets reasonable quality standards, and (3) the measures are operating correctly and have the potential to generate the predicted savings.

2. Baseline model goodness of fit screening for the target population

Purpose: M&V Plans should show that the intended baseline model forms are likely to characterize energy use well for the buildings in the program's target population. Industry standard goodness of fit metrics should be used, namely, coefficient of variation of the root mean squared error (CV(RMSE)), normalized mean bias error (NMBE), and coefficient of determination (R^2). The metrics are further defined and detailed in the examples shown in Appendix 1.

Goodness of fit screening guidance:

- Ability to sufficiently characterize energy use for the proposed population of target buildings should be demonstrated based on the fraction of analyzed buildings for which fitness metrics (see ASHRAE Guideline 14 and the IPMVP) are satisfied at the following thresholds:
    - CV(RMSE) < 25%[14]
    - NMBE[15] between -0.5% and +0.5%
    - $R^2 > 0.7$

- Baseline models planned for use in the program should be run against metered utility account consumption data from the program's target population, e.g. building types, sizes.
    - This analysis will be most meaningful if conducted using meters, or sums of meters that represent the totality of customer or buildings loads. For example, a building may have multiple meters, multiple accounts, and multiple potential program participants. This mapping is also referenced in the guidance for the Modeling Narrative.

- The baseline model goodness of fit screening analysis should be summarized to include:
    - The total number of buildings that were analyzed for model goodness of fit.

---

[14] For cases in which the CV(RMSE) does not meet the threshold, the reason may be due to missing independent variables, incorrect model form, the modeling time interval (hourly vs. daily), or other factors such as the presence of non-routine events. If desired, a data inspection may be conducted to identify large changes in consumption that may indicate a non-routine event. Suspected non-routine events may be removed from the data, if the event spans less than 25% of the data (by number of points), and CV(RMSE) may be re-computed; if the suspected event spans more than 25% of the data, the building may be removed from the analysis and the target population.

[15] NMBE refers to normalized mean bias error. Bias is the tendency of a statistical model to overestimate or underestimate the considered parameter. OLS regression models are the best linear unbiased estimators. When other methods are used there is a trade- off between decreasing the bias or the variance of the model; for this reason, cross validation is used to tune the model. For more information refer to: Friedman, J., Hastie, T. and Tibshirani, R., 2001. The elements of statistical learning. New York: Springer series in statistics.

- The number and fraction of buildings for which the CV(RMSE), R^2, and NMBE thresholds were all met – note that these do not comprise pass/fail criteria, but rather an analysis that will feed into an interpretation of model suitability (see end of Section 3).
- Optional; include only if buildings were removed from fitness screening: The number of buildings in which suspected non-routine events were identified, and whether the data or building was removed from the analysis.

- For example, results from the goodness of fit screening analysis would take the form of:
  - 250 buildings were selected and analyzed
  - For 200 buildings (80%), the CV(RMSE), R^2, and NMBE thresholds were met
  - For 50 buildings (20%), either the NMBE, R^2, or CV(RMSE) threshold was not met

- The results from baseline model goodness of fit screening analysis should also include an explanation and logic for the method used to select buildings, and the number of buildings selected.

- The results of the goodness of fit screening and if conducted, the uncertainty scenario analysis (see Section 3), should include an interpretation and discussion of the findings in terms of suitability of the planned model given the intended program design and implementation.
  - For example, if the models used in the analysis do not exhibit a good fit for a large fraction of the program's target population, will this be mitigated by individual site pre-screening? Will additional explanatory variables be acquired and documented in site-specific M&V plans, what are these variables, and how will they be obtained for program sites?

3. Scenario analysis of uncertainty due to baseline model error

Purpose: For cases in which baseline models are expected to use **monthly data**, proposed program M&V Plans should demonstrate that the proposed modeling approach is likely to produce results with acceptable levels of precision. Monthly baseline models' precision can be expressed in terms of uncertainty due to model error (non-routine adjustments are addressed in Section 5).

Scenario analysis:

- Uncertainty scenario analysis should be conducted as described in ASHRAE Guideline 14[16] and serves as an extension of the model goodness-of-fit analysis discussed in Section 2.

- Scenario analyses will define and/or set:
    - Ranges of CV(RMSE)[17] observed in model fitness screening
    - Range of expected fractional savings based on the program design, as referenced in program design documents
    - Planned number of data points in the baseline and performance period
    - A confidence level of 90%

- The uncertainty scenario analysis will determine the range of potential uncertainty[18] due to model error, at the 90% confidence level.

- Results from the uncertainty scenario analysis should be tabulated to show, for each building
    - CV(RMSE) from baseline fitness testing (Section 2)
    - Expected uncertainty due to model error
    - Expected fractional savings uncertainty due to model error

- Results of the scenario analysis should be summarized to indicate:
    - Fraction of buildings with uncertainty greater than expected savings, across the range of expected savings
    - Fraction of buildings with fractional uncertainty greater than 25%, and greater than 50%, across the range of expected savings

---

[16] See specifically Appendix B4 Uncertainty of Regression-Based Savings Models.

[17] In the 2014 version of Guideline 14, uncertainty equations are expressed in terms of MSE, i.e., the mean squared error; in the prior version they were expressed in terms of CV(RMSE). The CV(RMSE) can be converted to the MSE, or vice versa.

[18] Note that in the Guideline 14 formulation, uncertainty is directly proportional to CV(RMSE) and inversely proportional to the savings fraction. Therefore, high values of CV(RMSE) indicated through the analysis in Section 2, may be acceptable if savings are also expected to be high.

- An example of the form in which scenario analysis results could be presented is shown below in Table 1 and the associated results summary[19]. In this example, one year of monthly data is used in the baseline and performance periods, resulting in 12 post-installation performance data points. Based on the program design, the expected savings ranges from 5-7% at the whole building level. The CV(RMSE) values are those resulting from a goodness of fit analysis performed as described in Section 2.

Table 1: Uncertainty ranges due to model error: 12 baseline points, 12 expected post-installation data points; 5-7% expected savings; 90% confidence.

Building	CV(RMSE)	Expected Savings	Expected Uncertainty Due to Model Error	Expected Fractional Uncertainty due to Model Error
1	15	5%	1.3%	26%
		7%	.95%	14%
2	10	5%	1.1%	22%
		7%	.8%	11%
3	35	5%	4.5%	90%
		7%	2.5%	36%
…	…	…	…	…
n	20	7%	1.5%	21%

Summary of uncertainty scenario analysis results in Table 1:

Fraction of buildings with uncertainty greater than expected savings = 0%
Fraction of buildings with fractional uncertainty > 25% = 42%
Fraction of buildings with fractional uncertainty > 50% = 17%

- The results of the goodness of fit screening (Section 2) and uncertainty scenario analysis should include an interpretation and discussion of the findings in terms of suitability of the planned model given the intended program design and implementation.
  - For example, if the models used in the analysis do not exhibit a good fit for a large fraction of the program's target population, is the expected savings uncertainty reasonable? Will additional explanatory variables be acquired, what are these variables, how will they be obtained for program sites, and why are they expected to improve model fit?

[19] This example is purely illustrative, and included to provide a visual indication of how the results of this type of analysis might be presented. The numbers in the table are not based on analysis or modeling of actual building energy consumption data.

4. Coverage factor for independent variables in the baseline period

Purpose: Coverage factor refers to the range in observed values of independent variables during the baseline period. Baseline model projections for values of independent variables that are beyond those observed in the baseline period may under or over-estimate the counterfactual and associated savings estimates. For example, if a baseline model is constructed with baseline data that spans 50-75°F, it may not prove reliable in predicting consumption for 90°F conditions in the performance period. Analogous considerations apply to other potential independent variables such as those related to production.

Coverage factor guidance:

- M&V Plans may choose to adhere to ASHRAE Guideline 14, which advises: *"Apply the algorithm for savings determination for all periods where independent variables are no more than 110% of the maximum and no less than 90% of the minimum values of the independent variables used in deriving the baseline model."*
  - Alternative or enhanced assessments of coverage factor may be presented, but should include documentation sufficient to justify the approach

- M&V Plans should describe:
  - How a sufficient coverage factor will be verified
  - How the risk of insufficient coverage factor will be minimized, for example, through use of 12-month minimum baseline periods to capture annual variations in baseline model independent variables
  - How instances of insufficient coverage will be treated if they occur

5. Treatment of non-routine adjustments

Purpose: The importance of non-routine adjustments for events that occur during the performance period cannot be over emphasized; **they are critical to ensuring that a whole-building normalized metered energy consumption-based savings approach accurately reflects gross savings due to installation of program measures at an individual site.** M&V Plans should discuss how non-routine adjustments will be addressed.

Non-Routine adjustment guidance:

M&V Plans should include a description of:

- Common non-routine event types that are anticipated, given the program design, and associated building types, project types, and measures.

- The process that will be implemented to monitor projects to identify non-routine events including (measured or reported/surveyed) data that will be tracked, frequency of data collection, and methods to verify/validate data and/or reported/surveyed information.
    - This process may also include site-level data collection such as equipment operational or functional characteristics, operational parameters, and associated measurements.

- How non-routine events will be documented, and how their energy impacts will be quantified (e.g. simulation, engineering calculations) in accordance with the IPMVP.
    - Simple calculations may be sufficient for many adjustments, however in more complex cases simulation is preferred to address events with interactive effects, and prototype models may provide a useful beginning point.
    - Avoidance of bias in the directionality of adjustments (positive or negative adjustments of savings) should be discussed.

- The threshold for the magnitude of non-routine adjustments that will be quantified in determining of avoided energy use.

- Please see Appendix 2 for additional reference information on non-routine events and adjustments

6. Savings claims

Purpose: The previous sections in this document address model selection, suitability, and how non-routine events and adjustments will be addressed. Previous sections apply primarily to planning and development of the M&V Plan that is submitted with the Implementation Plan for a program, before specific buildings are recruited into the program, before projects are implemented, and before realized savings are quantified. Please see the Commission Rulebook for a full description of NMEC project documentation and filing requirements.

Appendix 1. Examples of model creation, variable selection, calculation of fitness metrics, and verification of coverage factor

Model creation and variable selection

Determination of appropriate baseline model resolution and independent variables is informed by assessment of model fitness metrics, including for example, the coefficient of variation of the root mean squared error, CV(RMSE), NMBE, and coefficient of determination, R^2.

In Example 1, whole building electric use is modeled using monthly electricity data and cooling degree days, and plotted in Figure A1. In the figure, the fit model is shown with a red line, and the metered consumption is plotted in yellow dots.

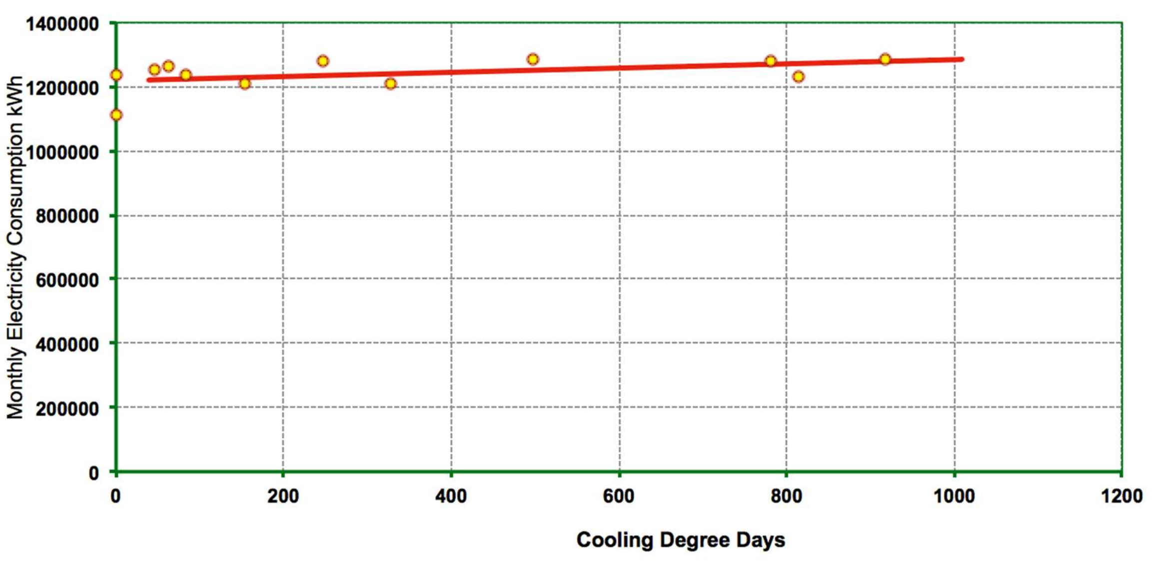

Figure A1. Monthly electricity consumption vs. cooling degree days

Although this model exhibits low CV(RMSE) and NMBE as desired, the R^2 metric fares much worse where higher values are desired. This is indicated in the summary of fitness metrics shown in Table A1 below:

Table A1. Fitness metrics for the model of monthly electricity consumption vs. cooling degree days

CV(RMSE)	3.7%
NMBE	0.25%
R^2	0.21

In Example 2, a second whole building electric model is tested, this time using hourly electricity consumption and outside air temperature, and plotted in Figure A2. In this model the independent variables used are time of week and outside air temperature. In the figure, the outside air temperature is plotted in red, the fit model is in blue, and the metered data is in pink.

Figure A2. Hourly electricity consumption based on time of week and outside air temperature

As indicated in the summary of fitness metrics in Table A2, the hourly model exhibits a better fit than the monthly model, with low NMBE and CV(RMSE) as well as high R^2.

Table A2. Fitness metrics for the model of hourly electricity consumption based on time of week and outside air temperature.

CV(RMSE)	11%
NMBE	-0.27%
R^2	0.95

Calculation of fitness metrics

Defined in Equation A1, the CV(RMSE) is the root mean square error normalized by the mean of the measured values. In the equation y_i is the actual metered value, $\hat{y}_i$ is the predicted value from the fit model, $\bar{y}$ is the average of the y_i, and n is the total number of data points. This metric provides a quantification of the typical size of the error relative to the mean of the observations. It indicates how much variation or there is between the data and the model, and reflects the model's ability to predict the overall energy use shape that is reflected in the data. Table A3 and Equation A2 provide an example calculation of the CV(RMSE), given twelve months of load data. In the case of interval data, the calculation remains the same, although the number of points, n, becomes much larger.

Equation A1. $$CV(RMSE) = \frac{\sqrt{\frac{1}{n}\sum_i^n (y_i - \hat{y}_i)^2}}{\bar{y}} \times 100$$

Table A3. Example calculation of parameters to calculate the CV(RMSE), R^2, and NMBE fitness metrics, given twelve months of load data.

Month	Metered load (y_i)	Predicted load ($\hat{y}_i$)	Metered-Predicted ($y_i - \hat{y}_i$)	(Metered-Predicted)^2 $(y_i - \hat{y}_i)^2$
1	394383	394320	63	3969
2	355120	377089	-21969	482636961
3	400758	390158	10600	112360000
4	423004	397406	25598	655257604
5	408421	406692	1729	2989441
6	421076	412458	8618	74269924
7	433731	432736	995	990025
8	452230	432995	19235	369985225
9	406071	417556	-11485	131905225
10	411741	424201	-12460	155251600
11	385556	380632	4924	24245776
12	385027	389090	-4063	16507969
Average, $\bar{y}$	406426			
Sum			21785	2026403719
Variance (y)				69580948

Equation A2. $$CV(RMSE) = \frac{\sqrt{\frac{1 \times 2026403719}{12}}}{406426.5} \times 100 = 3.19$$

Defined in Equation A3, R^2 is equal to one minus the mean square error divided by the variance of the actual energy use. In the equation y_i is the actual metered value, $\hat{y}_i$ is the predicted value from the fit model, var$(y)$is the variance of the y_i, and n is the total number of data points. It corresponds to the proportion of the energy use variance explained by the model. The R^2 value ranges between 0 and 1, with 0 indicating that the model explains none of the output variability, and 1 indicating that the model explains all the output variability. Using the values from Table A3, Equation A4 provides an example calculation of the R^2 given twelve months of load data.

Equation A3. $$R^2 = 1 - \frac{\frac{1}{n}\sum_i^n (y_i - \hat{y}_i)^2}{var(y)}$$

Equation A4. $$R^2 = 1 - \frac{\frac{1}{12}\text{x}2026403719}{642574282} = 0.73$$

Defined in Equation A5, NMBE represents the total difference between the actual metered energy use, and the energy use indicated with the fit model. In the equation y_i is the actual metered value, $\hat{y}_i$ is the predicted value from the fit model, $\bar{y}$ is the average of the y_i, and n is the total number of data points. Using the values from Table A3, Equation A6 provides an example calculation of the NMBE given 12 months of load data.

Equation A5. $NMBE = \frac{\frac{1}{n}\sum_{i}^{n}(y_i - \hat{y}_i)}{\bar{y}} \times 100$

Equation A6. $NMBE = \frac{\frac{1}{12}(21785)}{406426} \times 100 = 0.44$

Verification of coverage factor

ASHRAE Guideline 14 specifies: "apply the algorithm for savings determination for all periods where independent variables are no more than 110% of the maximum and no less than 90% of the minimum values of the independent variables used in deriving the baseline model." Table A4 provides an example of data used to verify sufficient coverage factor.

Table A4. Example of data used to verify sufficient coverage factor, given twelve months of load data, and a model that uses average outside air temperature (OAT) as the sole independent variable.

	Baseline		Performance Period	
Month	Consumption	Average OAT	Model-Predicted Consumption	Average OAT
1	394383	53.0	269831	54.1
2	355120	57.0	264236	57.4
3	400758	61.9	277054	58.1
4	423004	63.6	284204	61.2
5	408421	61.1	274539	59.9
6	421076	67.2	281134	67.1
7	433731	67.1	299625	69.5
8	452230	67.0	314535	70.2
9	406071	67.0	306156	69.1
10	411741	60.3	303321	66.3
11	385556	55.5	267428	53.0
12	385027	47.5	274512	50.6

In the data set shown in Table 4 above the range of observed values of average OAT during the baseline period range from 47.5 to 67.2 degrees. Applying the 90% of minimum and 110% of maximum criteria, the model could be confidently used to predict load during the performance period, for average outside air temperature conditions that range from 42.8 to 73.9 degrees. In the example data set, average outside air temperature ranges from 50.6 to 70.2 degrees, and therefore the coverage factor criterion is satisfied.

Appendix 2. Additional reference information on non-routine events and adjustments

Non-routine changes in building energy use are those that are not attributable to changes in the independent variables used in the baseline model, or to the efficiency measures that were installed. In the case of a non-routine event, the savings determined by subtracting the metered use in in the performance period from the baseline-predicted load may have to be adjusted to accurately determine the savings due to the installed measures. Figure A5 illustrates the presence of a potential non-routine event, as indicated by the building load profile.

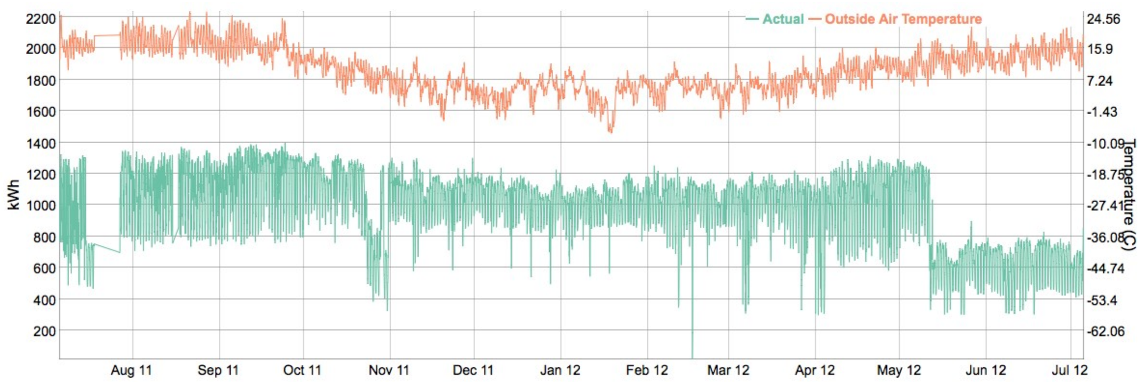

Figure A5. Approximately one year of metered electric load data (green), and outside air temperature (orange); the change in load in mid-May does not appear to be correlated with weather, and could indicate the presence of a non-routine reduction in consumption.

Some of the more frequently encountered types of non-routine events in commercial buildings include, but are not limited to those listed in Table A5.

Table A5. Frequently encountered non-routine event types in commercial buildings.

Services	# of rooms/beds
	food cooking/preparation
	# of registers
	#of workers
Equipment loads	# of computers
	# of walk-in or standard refrigeration units or open and closed cases
	# of MRIs
	# or capacity of HVAC units
Operations	hours of operation
	weekend operations
	heating and cooling setpoints
	system control strategies
Site characteristics	size
	% of building heated and cooled
	envelope changes

Non-routine events may be characterized as temporary or permanent, as load added or removed, and as constant or variable. A framework of assessing non-routine events may include

1. Determine whether an event is present
2. Determine whether the impact of the event is material, meriting quantification and adjustment (the threshold for what is considered 'material' should be specified in the M&V Program Plan)
3. Determine whether the event is temporary or permanent. Temporary events may be removed from the data set, however no more than 25% of the measured data should be removed, per ASHRAE Guideline 14, provided that a justifiable reason is provided.
4. Determine whether the event represents a constant or variable load
5. Determine whether the event represents added or removed load
6. Based on #3-5, the approach to measuring and quantifying the impact of the event may be determined.

Several methods may be used to determine whether an event is present. These include but are not limited to inspection of meter data, time series change detection or breakout analysis, periodic site visits and short-term measurements, and site surveys.

Determination of whether the impact of the event is material depends on engineering expertise, and the magnitude of the thresholds that are defined in the M&V Program plan.

Permanent events are those that are expected to last through the duration of the M&V analysis period.

Constant loads are understood to be those that do not fluctuate or change during a period of interest, such as when in the ‘on’ state.

Added loads are those that increase site energy consumption, while removed loads decrease site energy consumption.

Analogous to detecting the presence of an event, several methods may be used to quantify the impact or magnitude of the event. These include but are not limited to, engineering calculations, IPVMP Options A and B, simulation models, time series analysis of residuals, and the use of indicator variables in models fit to data before and after the event.

Appendix 3. Relevant Definitions

Topic	Definition
Bias	Bias is the tendency of a statistical model to overestimate or underestimate the considered parameter.
Calibrated simulation	Calibrated Simulation involves the use of computer simulation software to predict facility energy consumption during the baseline and/or reporting period. A simulation model must be "calibrated" so that it predicts an energy pattern that approximately matches actual metered data. (IPMVP)
Coefficient of determination (R^2)	The coefficient of determination (R^**2**) is the measure of how well future outcomes are likely to be predicted by the model. It illustrates how well the independent variables explain variation in the dependent variable. R^2 values range from 0 (indicating none of the variation in the dependent variable is associated with variation in any of the independent variables) to 1 (indicating all of the variation in the dependent variable is associated with variation in the independent variables, a "perfect fit" of the regression model to the data). (BPA).
Coefficient of Variation of the Root Mean Squared Error (CV(RMSE))	The coefficient of variation of the root mean squared error (CV(RMSE) is the RMSE expressed as a fraction or percentage of the mean of the actual data. (BPA).
Coverage factor	Coverage factor refers to the range in observed values of independent variables during the baseline period. Baseline model projections for values of independent variables that are beyond those observed in the baseline period may under or over estimate the normalized baseline consumption and associated savings estimates. (ASHRAE Guideline 14).
CZ2010 Weather Data	CZ2010 Weather Data is long term average weather data published by the California Energy Commission.
Early M&V	Early M&V refers to M&V conducted before required program reporting dates designed to identify and correct problems with program implementation. Same as "near term feedback."
Embedded M&V	The term "embedded M&V" refers to Commission direction that programs making claims based normalized metered energy consumption must collect sufficient data to validate the savings claims and document the financial incentives. Implementers must submit an Implementation plan consistent with D.15-10-025 Appendix 4 and include a program level measurement and verification (M&V) plan that defines the data collection activities. Financial data shall include the amount of financial incentives paid to customers or the amount of compensation offered to implementers or contractors .
Gross savings	Gross savings count the energy savings from installed energy efficiency measures irrespective of whether or not those savings are from free riders, i.e., those customers who would have installed the measure(s) even without the financial incentives offered under the program.
Interactive effects	Any energy effects occurring beyond the project measurement boundary are called 'interactive effects'. See Project Measurement Boundary. (IPMVP)

IPMVP	International Measurement and Verification Protocol (IPMVP) is a guidance document describing common practice in measuring, computing and reporting savings achieved by energy or water efficiency projects at end user facilities. The IPMVP presents a framework and four measurement and verification (M&V) Options (Options A, B, C and D) for transparently, reliably and consistently reporting a project's savings. www.evoworld.org
M&V	Measurement and Verification (M&V) is the process of using measurement to reliably determine actual savings created within an individual facility by an energy efficiency intervention. Savings cannot be directly measured, since they represent the absence of energy use. Instead, savings are determined by comparing measured use before and after implementation of a project, making appropriate adjustments for changes in conditions[20].
M&V Plan	The M&V Plan is a document describing the energy efficiency measures, data collection activities, data analysis methods and reporting activities. The preparation of an M&V Plan is a recommended part of savings determination. Advance planning ensures that all data needed for savings determination will be available after implementation of the energy efficiency measures. (IPMVP)
Material event	A non-routine event considered to have sufficient impact on the energy savings prediction that it must be included in the NMEC model.
Net savings	The savings realized when free ridership is accounted for. The savings is calculated by multiplying the gross savings by the net to gross ratio.
NMBE	NMBE refers to normalized mean bias error, which is the total error in the model expressed as a fraction of the total energy use, adjusted for the number of parameters in the model. (BPA).
Outlier	Data points that do not conform to the typical distribution. Graphically, an outlier appears to deviate markedly from other members of the same sample. (BPA).
Project Measurement Boundary	The Project Measurement Boundary refers to the portion of the building or facility included in the energy savings model. In the context of Option C (whole building) analysis, the measurement boundary encompasses the whole facility. For M&V Plans that utilize submetering, or selection of a subset of meters serving the building, the project measurement boundary is the portion of the building served by the selected meters or submeters. (IPMVP).
Residual	The residual is the difference between the predicted and actual value of the dependent variable in an energy consumption model. (BPA).
Root Mean Squared Error (RMSE)	The root mean squared error is typically referred to as a measure of variability, or how much spread exists in the predicted and the actual data. (BPA).