

Clean Energy Fund Commercial Chapter Impact Evaluation

Real Time Energy Management (RTEM) Program (2017 – Q1 2020 (Pre-COVID-19))

Final Report

Prepared for:

New York State Energy Research and Development Authority
Albany, NY

Dana Nilsson
Senior Project Manager, NYSERDA

Prepared by:

DNV

Chris Zimbelman, Associate Director, DNV
Alain Tayoun, Senior Engineer, DNV
Nathan Throop, Senior Engineer, DNV
Ben Jones, Vice President, DNV
Bradley Campbell, Director, DNV



Record of Revision

RTEM Impact Evaluation Final Report
RTEM Impact Evaluation Final Report October 2021

Revision Date	Description of Changes	Revision on Page(s)
10-8-2021	Original Issue	Original Issue



Notice

This report was prepared by the DNV Impact Evaluation Team in the course of performing work contracted for, and sponsored by, the New York State Energy Research and Development Authority (NYSERDA). The opinions expressed in this report do not necessarily reflect those of NYSERDA or the state of New York, and reference to any specific product, service, process, or method does not constitute an implied or expressed recommendation or endorsement. Further, NYSERDA, the state of New York, and the contractor make no warranties or representations, expressed or implied, as to the fitness for particular purpose or merchantability of any product, apparatus, or service, or the usefulness, completeness, or accuracy of any processes, methods, or other information contained, described, disclosed, or referred to in this report. NYSERDA, the state of New York, and the contractor make no representation that the use of any product, apparatus, process, method, or other information will not infringe on privately owned rights, and will assume no liability for any loss, injury, or damage resulting from, or occurring in connection with, the use of information contained, described, disclosed, or referred to in this report.

NYSERDA makes every effort to provide accurate information about copyright owners and related matters in the reports we publish. Contractors are responsible for determining and satisfying copyright or other use restrictions regarding the content of reports that they write, in compliance with NYSERDA policies and federal law. If you are the copyright owner and believe a NYSERDA report has not properly attributed your work or has used it without permission, please email print@nyserda.ny.gov.

Information contained in this document, such as web page addresses, are current at the time of publication.



Table of Contents

RECORD OF REVISION	I
NOTICE	II
TABLE OF CONTENTS	III
LIST OF FIGURES.....	III
LIST OF TABLES.....	III
1 EXECUTIVE SUMMARY.....	1
1.1 Approach.....	1
1.2 Results.....	3
1.3 Findings and Recommendations.....	4
2 BACKGROUND	7
2.1 Program Description.....	7
2.2 Evaluation Objectives.....	7
2.3 Previous Evaluations.....	8
3 METHODS	9
3.1 Pre-Post Consumption Analysis.....	9
3.2 Measure-Level Analysis.....	10
3.3 Aggregate Analysis.....	11
4 RESULTS, FINDINGS, AND RECOMMENDATIONS.....	15
4.1 Verified Gross Energy Savings (VGS) Results.....	15
4.2 Measure-Level Analysis Results.....	21
4.3 Findings and Recommendations.....	23
APPENDIX A: GLOSSARY OF TERMS	25

List of Figures

Figure 4-1. Savings Comparison for Nine-Year Weather Data vs. 30-Year Weather Data	18
Figure 4-2. Example Site Weather Correlation.....	18
Figure 4-3. Program and Evaluation Baseline Energy Consumption Comparison	19
Figure 4-4. Program and Evaluation Savings Comparison for the Sampled Sites	20

List of Tables

Table 1-1. Program Reported and Verified Gross Savings	3
Table 2-1. Facility Type Distribution	7
Table 3-1. Facility Type Distribution.....	12
Table 3-2. Weight Calculation	13
Table 4-1. Verified Gross Savings.....	15



Table 4-2. Percent Savings Comparison with Similar Program.....	15
Table 4-3. Analysis Differences	16
Table 4-4. Savings Comparison for Nine-Year Weather Data vs. 30-Year Weather Data	17
Table 4-5. Vendor Reported Energy Savings.....	21
Table 4-6. Capital vs Non-Capital Project Savings	22
Table 4-7. Measure EUL Based on Different State TRMs.....	22

1 EXECUTIVE SUMMARY

The CEF was designed to support New York State's clean energy agenda by working with market participants to develop clean energy market opportunities at scale and advance progress toward the State's nation-leading clean energy goals. NYSERDA has the following programs that are currently funded through the CEF Commercial Chapter, with impact evaluations planned under a common contract to help synchronize insights for this sector:

- Commercial Real Estate (CRE) Tenant
- REV Campus Challenge
- P-12 Schools
- Remote Energy Management (REM)/Real-Time Energy Management (RTEM)

This report focuses on the REM/RTEM program and describes an impact evaluation that assessed the energy savings for projects installed between October 31, 2016, and February 1, 2019. This period immediately precedes the widespread of COVID-19 and was chosen to help avoid uncertainty arising from the effects the pandemic had on typical operations.

Energy Management (EM) is the common name for the management of building energy consumption from a combination of building data collection systems (e.g., meters, sensors, equipment feeds), analytics, and building data information services. There is a full spectrum of EM sophistication, ranging from the basic, REM, to the more advanced RTEM.

The Program provides cost-sharing incentives for both the information-gathering systems and the vendors reviewing the captured data. Vendors provide recommendations for changes based on the information and customers decide what to implement. The systems do not include automated controls (though these could be recommended by the vendors). Incentives are based on the project cost (30% for system installation and 30% for service years 1 – 3; reduces to 20% service cost share for years 4 - 5) and are not directly proportional to energy savings achieved by the project.

The objectives of this impact evaluation are as follows:

- Estimate the evaluated and verified gross first-year energy impacts for RTEM projects, which includes electric energy (kWh) and fossil fuel energy (MMBtu) savings and provide recommendations to improve Program effectiveness.
- Develop the verified gross savings realization rate (SRR) for the Program period.
- Investigate savings persistence, which includes quantifying the annualized evaluated gross energy savings in the second and subsequent years for electric (kWh) and fossil fuel energy (MMBtu) at customer sites.

1.1 Approach

The RTEM Program funded 293 projects between January 1, 2017, and February 28, 2020. These projects received incentives to help offset installation costs of either the information-gathering portion of the RTEM system, the engineering review of the data, or both. DNV was tasked to conduct an evaluation of the RTEM program to develop the SRR, verify the gross electric and gas savings, and investigate the expected life of the

savings. To date, DNV has reviewed program data and performed savings analysis using two separate approaches, first using a billing analysis and second with a more detailed ground-up approach compiling measure details from program documents. This section outlines the two approaches examined in evaluating the RTEM program.

1.1.1 Billing Analysis

RTEM¹ works in real time to monitor building systems' current and historical performance data. Data points such as set points, power loads, flow rates, temperature, and humidity, are collected and processed on-site, or on a cloud-based server, or a combination of the two to fine-tune the building energy system operations and identify capital projects. Depending on the technology, RTEM systems can also provide predictive analytics, fault detection and diagnostics and performance optimization. Due to this setup, and with the possibility of interaction occurring throughout the building, for this study, evaluators chose to use International Performance Measurement and Verification Protocol (IPMVP) Option C, or whole building pre-post analysis, to estimate the verified gross savings in the first round of review.

The Program received utility data from 42 of 293 participating sites to perform measurement and verification (M&V) and to estimate the Program's first year savings. Due to the timing for this initial study, evaluators used the same data to perform the evaluation pre-post consumption analysis. The collected data was regressed against historical weather data to generate a linear regression model that could be applied to "typical" weather conditions to calculate savings. The savings calculated from the sampled sites were then extrapolated to the population using a sample weights and the statistical method ratio estimation. The weights take into account the segment type – commercial office, commercial retail, commercial other, and multifamily – and facility size – small, medium, and large. The weight calculations and extrapolation are further discussed in Section 3.3.2.

Evaluators also performed an in-depth review of Vendor reports and Program documentation for the period immediately preceding COVID; however, no direct facility survey was completed during this work. Any other possible non-routine events were not described in Vendor reports or Program documentation and therefore were not addressed.

1.1.2 Measure-Level Analysis

The evaluation also explored an alternative approach to assess and compare savings results to the billing analysis. This alternative method looks at the same sites studied in the billing analysis but from the bottom up: the measure-level analysis generates site-level savings through measure-level energy savings calculations. This analysis was intended to provide additional context to the billing analysis results and provide additional information on the persistence and long-term savings expectations for the Program.

The vendors participating in the Program collect, store, and analyze their installed RTEM system data to provide their customers with insightful findings that improve the participating facility's operation, performance, and occupancy comfort. As part of the program requirements, vendors submit service reports to both NYSERDA and customers at regular intervals. The evaluators requested those reports and their supporting trend data for further review. DNV examined the service reports associated with each of the sampled sites and extracted the list of measures recommended, their installation status, the energy savings associated with those measures for each reporting period, any non-energy savings impacts and the calculation methodology when available. The evaluators compared the vendor claimed savings to the billing analysis savings to see how the measure-level analysis SRR would fair against the former approach. In addition, the evaluators attempted to calculate measure-level savings from the information provided in the reports and the trend data to compare

¹ <https://www.nyscrda.ny.gov/ny/PutEnergyToWork/Energy-Technology-and-Solutions/Building-Operations-and-Performance/Real-Time-Energy-Management>

with the vendor claimed savings. As will be discussed in depth in Section 3, the evaluators found that there was not enough data to validate the claimed savings suggested in the service reports.

Finally, the evaluators also leveraged the collected measure-level information to examine the distribution of measures recommended and installed. This analysis separated measures into capital (new equipment) and non-capital (operations & maintenance) measures and looked at TRM values for the estimated useful life of the measures. The goal of this exercise is two-pronged:

1. Identify potential savings the Program can expect beyond year one as savings from capital projects, which are more likely to be achieved in later years given a longer timeline for installation and commissioning relative to non-capital measures.
2. Identify the expected life of savings achieved by the Program.

1.2 Results

The primary focus of this study was to develop the verified gross energy savings estimates and the SRR for the RTEM Program. After reviewing multiple approaches to estimate the savings for the Program, the evaluators recommend the use of the weighted billing analysis data as the evaluated savings of the study. These values were calculated by fuel type. Table 1-1 outlines the gross savings, the SRR by fuel type, and the relative precisions derived from the billing analysis.

Table 1-1. Program Reported and Verified Gross Savings

Parameter	Sample Size	Gross Savings Realization Rate (SRR)	RP at 90% Confidence	Absolute Precision	Program Reported Savings	Evaluation Verified Gross Savings	Evaluated %Savings Relative to Baseline
Electric energy (MWh/yr.)	39	20%	76%	15%	159,309	32,043	3.2%
Fossil fuel energy (MMBtu/yr.)*	39	42%	105%	44%	131,003	55,029	3.0%

*Includes natural gas, fuel oil, and district steam

The SRRs for electric and fossil fuels were 20% and 42%, respectively. The total combined fuel MMBtu realization rate is 24%. The evaluated savings were 3% of baseline consumption, which is within the range of savings typically found for energy management programs in other jurisdictions.

While both the Program and evaluation used IPMVP Option C for their analyses, there were several differences in their specific methods. There were also differences with respect to how the savings were extrapolated to the population. Details surrounding the differences between the evaluation and program analyses can be found in Section 4.1.2.

1.2.1 Overall Precision

The calculated relative precisions for the saving estimates are 76% and 105% for electric and fossil fuels respectively at the 90% confidence interval. The high relative precision is due to the large variation in evaluated results at each site relative to Program claimed savings and the small proportion of the population included in the sample. The variation is further outlined by the absolute precision which was 15% and 44% for electric and fossil fuel respectively.

1.3 Findings and Recommendations

The evaluators offer the following two findings and eight recommendations to help improve program performance moving forward.

1. **Finding:** The Program conducts measurement and verification (M&V) and works to accurately capture Program savings. The Program has been collecting M&V data for a sample of sites that have been installed for at least 12 months for each program year, with the number of sampled sites growing as program participation grows. Moving forward, they are collecting baseline data for every site enrolled in the Program, which will provide more granularity on the baseline consumption and breakdown of fuels.

Recommendation: Evaluators recommend stratifying by two dimensions to weight the sampled projects. The first dimension is facility type, and the second is facility size. This approach will allow for more accurate representation of the population along these dimensions. Below are the recommended stratification segments, based on the population of 293 sites that were evaluated to date.

Facility Type:

- a. **Commercial Office:** These account for 38 out of 293 projects, and 50% of the total population energy use.
- b. **Commercial Other:** These account for 141 out of 293 projects, and 30% of the total population energy use. The facility types under this segment will depend on the population.
- c. **Commercial Retail:** These account for 51 out of 293 projects, and 4% of the total population energy use.
- d. **Multifamily:** These account for 63 out of 293 projects, and 16% of the total population energy use.

Facility Size:

- a. **Projects greater than 1,000,000 sq ft.**
- b. **Projects between 100,000 sq ft and 1,000,000 sq ft.**
- c. **Projects less than 100,000 sq ft.**

Future program stratification should be a function of the population and may be broken down by a different variety of representative facility types and sizes. Additionally, once the program increases its available data, we recommend sampling and extrapolating savings within the expected fuel use type.

Recommendation: We also strongly recommend that within the utility data collection, the Program should retain the actual read dates from the utility bills in order to use those dates to best align with historical weather data. In addition to those dates, we recommend retaining information on estimated vs. read values, if available. Consider revising the Program to require opting into utility data sharing with NYSERDA (not just vendors).

2. **Finding:** The Program implementers are taking many steps to calculate energy savings with reasonable and appropriate methods. They collect utility info on a sample of sites each year and are working to continuously improve those savings methodologies. The evaluators have the following recommendations with respect to that analysis.

Recommendation: When collecting utility data and conducting the billing analysis, evaluators recommend accounting for an installation period. Currently, 24 months of consecutive utility data are

used in most cases for the savings analysis (targeting 12 months of pre-installation data and 12 months of post-installation data). It is likely that there was a period within that date range where the installation of the RTEM system was in progress, and it was not immediately fully operational. There is also the potential for a time lag between the installation of RTEM and operational changes that are the result of the newly installed monitoring system. Accounting for these periods and starting the post-installation billing period after installation is complete and potentially has had some time to function will likely increase savings for the program.

Recommendation: Distinguish between forecasts of savings and acquired savings. While forecasts may include predicted, but not yet realized savings in the future, acquired savings should be based on observed savings in a sample that has been extrapolated to the population (where full population observation is not reasonable). Do not apply the capital projects adjustment to acquired savings and switch to looking at a savings persistence approach, or the percent savings achieved each year. Capital projects that are recommended and installed as a result of the Program should not be claimed until they are installed. If capital projects are installed at a facility, they will be captured within the billing analysis that is conducted and will contribute to the savings once they are captured. These savings can then be mathematically extrapolated to the population using the appropriate sample weights. We also caution that many of these projects are likely to receive incentives from other programs, such as utility implementation programs. Overlap is not within the scope of this study; however, it is a factor that should be considered.

Recommendation: Evaluators recommend collecting detailed information from each site receiving M&V moving forward prior to using 2020 data in billing analyses. This information should be targeted at understanding how building operations in 2020 compared to a normal year. If facilities were operating as they typically do for part of the year, then that portion may be used. However, it is likely that 2020 data may not be representative of typical facility operation.

3. **Finding:** The evaluators extracted measure-level information from the sites' service reports and classified them as capital and non-capital to analyze the impact of both in the future. The evaluators then assigned EUL values based on the measures' closest description in the referenced TRMs. The majority of the installed savings stem from non-capital measures that typically have an EUL of 5-7 years.

Recommendation: The evaluators recommend investigating the program persistence further. Measure-level review findings show that the majority of the installed measure types include: control settings, schedules, retrocommissioning or repairs. Based on the referenced TRMs, those type of measures have a useful life of 5-7 years. As a result, the preliminary findings indicate that the program persistence is between 5-7 years.

4. **Finding:** In reviewing the service reports provided by the vendors, the evaluators identified inconsistent and missing information. The evaluators recognize that the reports are tailored to their end user, but the current structure does not allow the program to get a full picture of the activities happening at sites due to the RTEM system. This finding is consistent with Program findings as well, and the Program is working to address this.

Recommendation: The evaluators recommend requiring site and measure-level savings information from the vendors, outlining what measures were recommended, their installation status, the energy savings by fuel associated with them, and a brief description or narrative of how the measure contributes to energy savings. This will allow the Program to both understand participant actions better

and can provide supporting evidence for M&V activities. This recommendation is consistent with the Program direction moving forward.

Recommendation: The evaluators recommend including information on the mechanical equipment affected by the measures as part of the reporting requirement. The equipment size, efficiency, etc. was often missing from the reports, and having that information would allow an assessment of the vendor estimates of measure savings.

2 BACKGROUND

This section presents a Program description, the evaluation goals, and a summary of the previous evaluations.

2.1 Program Description

RTEM works in real time to monitor building systems' current and historical performance data. Data points such as set points, power loads, flow rates, temperature, and humidity, are collected and processed on-site, or on a cloud-based server, or a combination of the two to fine-tune the building energy system operations and identify capital projects. Depending on the technology, RTEM systems can also provide predictive analytics, fault detection and diagnostics and performance optimization. RTEM systems are observation only, control of the systems for which RTEM systems monitor is through separate control systems.

EM techniques are applicable to all building types and organizational structures. Existing and new construction, including commercial, industrial, and multifamily buildings, can benefit. The initial targeted sector for the program is existing commercial buildings, with uptake likely higher in subsectors with significant existing penetration of Building Management Systems—commercial office, retail, university/college, non-profit, and healthcare. These sectors also have large, centrally managed buildings or portfolios, and therefore are more likely to have the human resources necessary to capitalize on the potential of EM.

The Program provides cost-sharing incentives for both the information-gathering systems and the vendors reviewing the captured data. Incentives are based on the project cost (30% for up to a five-year vendor contract) and are not directly proportional to energy savings achieved by the project.

2.1.1 Evaluated Population

This evaluation included 293 RTEM projects installed between October 31, 2016, and February 1, 2019, and includes data for the period immediately preceding COVID (February 28, 2020) A breakdown of the facility types for the population of projects can be found in Table 2-1.

Table 2-1. Facility Type Distribution

Customer Facility Type	Number of Facilities in Population	% of Population
College/University	6	2%
Commercial Office	38	13%
Commercial Retail	51	17%
Food/Beverage	75	26%
Government	17	6%
Healthcare	14	5%
Hospitality	7	2%
K-12 School	14	5%
Manufacturing	1	0%
Multifamily	63	22%
Not-for-Profit	7	2%
Total	293	100%

2.2 Evaluation Objectives

This is an ongoing evaluation that has and will continue to occur in multiple phases. The objectives of this impact evaluation are as follows:

1. Estimate the evaluated and verified gross first-year energy impacts for RTEM projects, which includes electric energy (kWh) and fossil fuel energy (MMBtu) savings and provide recommendations that seek to improve program effectiveness.
2. Develop the verified gross SRR for the program period.
3. Investigate savings persistence, which includes quantifying the annualized, evaluated gross energy savings in the second and subsequent years for electric (kWh) and fossil fuel energy (MMBtu) at customer sites.

2.3 Previous Evaluations

This is the first evaluation of the RTEM program; no previous evaluations have been conducted.

3 METHODS

Section 3 describes the methods used to develop impact estimates for the RTEM population from project years 2016 through early 2019.

3.1 Pre-Post Consumption Analysis

Evaluators used data collected as part of the program M&V process to complete a utility pre-post billing analysis. Program staff provided utility information for 42 sites. Evaluators reviewed the data and performed a quality check for completeness and usability. Upon review, three of the sites were dropped from the evaluation sample due to a lack of the minimum required amount of data to complete the analysis, which was set at nine months of pre-install and nine months of post-install data. Evaluators completed a savings analysis on 39 of the sites.

For each site within the sample population, evaluators plotted the utility data against historical weather data to establish the relationship between utility consumption and both heating and cooling degree days. Evaluators used a dual-variable regression, which allowed them to check for both heating and cooling dependence for each of the fuels consumed at the site.

Once the relationship between the site's energy use and outside temperature was established using actual weather data, these regressions were then used to calculate the normal long-term annual use by using recent nine-year average weather data.

No R² filter was applied to the pre-electric regressions for sites with existing fossil fuel equipment for inclusion, as the pre-electric is not expected to have the same weather dependence as the post. The R² is a measure of how well a regression model represents the data. In evaluations with a big enough sample size, the evaluators typically set an R² threshold below which a site is removed to uphold the quality of the sample.

The following equations calculate the normal modeled energy consumption for a given fuel and case.

$$E_{\text{case}} = \sum_{i=1}^{12} C_{\text{case}} \times CDD_i + H_{\text{case}} \times HDD_i + B_{\text{case}}$$
$$EH_{\text{case}} = \sum_{i=1}^{365} H_{\text{case}} \times HDD_i$$
$$EC_{\text{case}} = \sum_{i=1}^{365} C_{\text{case}} \times CDD_i$$

where:

E_{case} =	Annual energy of given case, pre or post, and fuel type (kWh or MMBtu)
EH_{case} =	Annual heating energy of given case, pre or post, and fuel type (kWh or MMBtu)
EC_{case} =	Annual cooling of given case, pre or post, and fuel type (kWh or MMBtu)
HDD_i =	Monthly heating degree days of typical year weather for month i
CDD_i =	Monthly cooling degree days of typical year weather for month i
C_{case} =	Cooling degree day constant for given case, pre or post, and fuel type

H_{case} = Heating degree day constant for given case, pre or post, and fuel type

B_{case} = Non-weather-dependent constant for given case, pre or post, and fuel type

Savings were then calculated as the difference in the weather-normalized pre-installation (base case) consumption and the weather-normalized post-installation (post-case) consumption.

3.2 Measure-Level Analysis

The evaluators received project files supporting the 39 sampled sites. Project files included the vendor's service reports, RTEM meter trend data, project documentation, and, in some instances, calculation spreadsheets. Following the low SRR obtained from the billing analysis, the evaluators leveraged the project information provided to assess how reasonable the billing analysis results were. The following section describes the steps taken by the evaluators in conducting the measure-level analysis.

3.2.1 Service Report Review

The evaluators conducted a detailed review of the service reports of all 39 projects. The goal of this review was to understand the scope of the projects and to extract information pertaining to the recommended measures resulting from the RTEM system observations. The extracted information included:

- Recommended measure description and type (capital vs. non-capital)
- Measure installation status
- Non-energy impacts
- Installation dates
- Service reports start and end dates
- Energy and cost savings resulting from the measure (electric, gas, or steam)

The information above was collected for all service reports within a project. Service reports predominantly cover a period of six months, with a few exceptions covering quarters. The number of service reports provided depends on how long the RTEM system has been installed. The evaluators encountered one to four service reports per site.

Through the review, the evaluators identified a single vendor with 11 projects in the sample that provides service reports that are less quantitative and more qualitative than the other vendors. The evaluator excluded these 11 projects from the measure level analysis, since no data conducive to an analysis could be extracted from the reports. NYSERDA staff are working to obtain data sufficient for savings analysis from this vendor for future evaluations.

3.2.2 Trend Data Analysis

Following the measure-level data collection, the evaluators proceeded to verify the energy savings reported by the vendors by conducting an analysis on the trend data provided. The following steps were taken in attempting to conduct the analysis:

- Examine the trend data to identify points that are relevant to the installed measures.
- Identify key information related to the affected mechanical equipment (size, efficiency, flow rate, etc.)
- Obtain weather data from the closest weather station for the same time as the data on hand.

- Check the measure implementation date against the time series start and end date to determine whether enough pre and post installation data is provided.
- Conduct analysis based on the measure description to identify whether the recommended changes are reflected in the data.

The approach described was attempted on all 28 sampled sites. The analysis proved to be inconclusive for various site-dependent factors. The leading factors were sites having either pre or post installation data, no relevant data-points to the recommended measure, and missing equipment specs.

The evaluators elected to adhere to the billing analysis savings for three main reasons. First, the vendor claimed savings were on par with the billing analysis savings, as a result, the vendor claimed savings did not show a potential for an increase in the evaluation's SRR. Second, the vendor reported savings cover 28 out of the 39 sampled sites. The 11 sites that were omitted were all from one vendor, so while the billing analysis and vendor report savings generally align in aggregate, the evaluators could not verify if that's the case for a major participating vendor. Third, the trend data analysis proved to be inconclusive and not viable due to lack of information and data, therefore, the evaluators did not have the means to validate the reported savings.

3.3 Aggregate Analysis

The following describes the key steps and factors for the aggregate analysis.

3.3.1 Site Baseline Energy Consumption

For this analysis, evaluators used the utility bills to calculate the baseline load for each of the 39 sites. Both the electric and fossil fuel (gas and steam) data were weather normalized using the most recent nine years of weather data.

3.3.2 Savings Extrapolation

The baseline and post-installation energy usage as well as the energy savings were established for both electric and gas usage through the billing analysis as described in Section 3.1. As a result, savings from all 39 sites in our sample are calculated. To extrapolate to the entire population, the evaluators assigned each of the 39 sites a sample weight based on their customer segment and size. The tracking data set is considered as the reference in the size weighting. Since the program calculated the savings by applying a factor to the baseline energy usage, which in turn is based on square footage and the corresponding CBECS energy intensity value, the square footage was used as the proxy for a site's size as follows:

- Small: sites with square footage less than or equal to 100,000 sq. ft.
- Medium: sites with square footage between 100,000 sq. ft. and 1,000,000 sq. ft.
- Large: sites with square footage greater than or equal to 1,000,000 sq. ft.

The evaluators found that the three sector types most represented in the sample and population are multifamily, commercial retail, and commercial office. To properly represent the remaining sector types and avoid having the main three segments dominate the sample in weight, the remainder of the sites were grouped under "commercial other." Table 3-1 shows the population spread across the four sector types and facility size.

Table 3-1. Facility Type Distribution

Evaluation Assigned Sector	Customer Facility Type	Facility Size	Number of Facilities in Sector and Size	Total number facilities in sector
Commercial Other	Food/Beverage	Medium	1	75
		Small	74	
	Government	Medium	4	17
		Small	13	
	Healthcare	Medium	12	14
		Small	2	
	K-12 School	Medium	8	14
		Small	6	
	Hospitality	Medium	4	7
		Small	3	
	Not for Profit	Small	7	7
College/University	Medium	6	6	
Manufacturing	Large	1	1	
Commercial Other Total				141
Multifamily	Multifamily	Medium	36	63
		Small	27	
Commercial Retail	Commercial Retail	Medium	10	51
		Small	41	
Commercial Office	Commercial Office	Large	14	38
		Medium	20	
		Small	4	
Total				293

Following the segmentation, the sampled sites were assigned a weight based on how many sites of the same size and sector they represent in the population. The weight is the ratio of the number of sites within a particular category in the sample to the number of sites in the same category in the population. In the instance where there are no large sites of a particular sector in our sample, the medium and large sites are grouped together. Table 3-2 shows the weight calculations.

Table 3-2. Weight Calculation

Sector	Size	Number of sites in sample	Number of sites in population	Weight
Commercial Other	Large/Medium*	6	36	6.0
	Small	6	105	17.5
Multifamily	Medium	6	36	6.0
	Small	5	27	5.4
Commercial Retail	Medium	3	10	3.3
	Small	1	41	41.0
Commercial Office	Large	3	14	4.7
	Medium	7	20	2.9
	Small	2	4	2.0

*Large and medium “commercial other” sites were grouped together since there were no large sites in the sample.

Subsequently, the weight of each site was applied to both the evaluated savings and the program level savings within the sample. The realization rate was then calculated as the ratio of the weighted evaluated savings to the weighted program savings. The evaluated program level savings were obtained by applying the realization rates for each fuel type to the program reported savings.

3.3.3 Savings Realization Rate

Once the evaluators calculated savings for each of the sites within the sample, they then quantified the overall SSRs by fuel type and facility type. The facility size was embedded in the weights assigned to sites of different sizes within a particular facility type. As noted above, this breakout was chosen in order to properly stratify the population and be consistent with future evaluations of the program. The realization rate calculation is as follows.

Notation: The following terms are used in calculating the realization rate for each fuel type:

- T_j = Tracking estimate of gross savings for site j
- V_j = Verified estimate of gross savings for site j
- W_j = Sample weight for site j
- S = Number of sites in the sample

The realization rates are calculated directly:

$$RR = \frac{\sum_{j=1}^S W_j * V_j}{\sum_{j=1}^S W_j * T_j}$$

3.3.4 Precision and Limitations

Relative precision was calculated using the procedures described in Chapter 13 of the California Evaluation Framework².

The evaluation was limited by the following factors:

² http://www.calmac.org/publications/California_Evaluation_Framework_June_2004.pdf

- Sample of convenience: Both the program and the evaluators selected the sites in the sample based on availability of data. The sites were not selected at random within each stratum, which could result in potential bias.
- Initial program year: Since the program period evaluated included only the first year of implementation, the vendors had to undergo a learning curve in understanding the process from the installation of meters to implementation of measures.
- First year savings: In the first year of an Energy Management program, vendors had the opportunity to identify numerous operations and maintenance measures that are easy to identify and implement. Once those are identified, implemented, and exhausted, there is a possibility the program will face difficulties in maintaining the level of first year savings as new incremental savings in future years.
- Capital measures: First year measures are less likely to include savings from capital projects, particularly large and expensive measures, since these require more time to identify, budget for, and implement than operations and maintenance type measures.

4 RESULTS, FINDINGS, AND RECOMMENDATIONS

This section presents the results and findings from the verified gross savings evaluation. The section concludes with recommendations.

4.1 Verified Gross Energy Savings (VGS) Results

The first objective of this evaluation was to determine the verified gross savings for the program during the evaluated time period.

4.1.1 Overall VGS Results

The VGS for projects completed between October 31, 2016, and February 1, 2019 is calculated by conducting a pre-post billing analysis on a representative sample of sites from the population and extrapolating the results by applying the site-specific weight to its energy savings. The baseline energy use, VGS, VGS RR, and percent saved are presented in Table 4-1.

Table 4-1. Verified Gross Savings

Parameter	Sample Size	Gross Savings Realization Rate (SRR)	RP at 90% Confidence	Absolute Precision	Program Reported Savings	Evaluation Verified Gross Savings	Evaluated %Savings Relative to Baseline
Electric energy (MWh/yr.)	39	20%	76%	15%	159,309	32,043	3.2%
Fossil fuel energy (MMBtu/yr.) *	39	42%	105%	44%	131,003	55,029	3.0%

*Includes natural gas, oil, and district steam

The program savings identified through the billing analysis was 3.2% and 3% for electric and fossil fuels, respectively. The total electric energy savings was 32,043 MWh, and the total fossil fuel savings was 55,029 MMBtu. The calculated relative precision for this result is 76% and 105% for electric and fossil fuels respectively at 90% confidence. The latter indicates that the realization rate is in the 5%-35% and -2%-86% range for electric energy and fossil fuels respectively. The wide ranges on the precision are driven by several factors: the variability in site results is the most important, but sample size, billing data that includes usage by non-RTEM monitored equipment, data reliability, and the use of a savings factor that attempt to account for non-implemented capital projects by the program team are also contributing factors.

The evaluators researched evaluations of analogous programs to serve as a comparison to the RTEM evaluation findings. Even though the programs are not identical, they are all energy management programs that work with commercial and multifamily facilities. The basis of all these programs is analyzing collected data from the facility and producing actionable recommendations that predominantly save energy. This exercise serves to estimate the range of savings only and is not intended to inform the reported RTEM program savings. Table 4-2 summarizes the savings percent from these studies. As can be seen in the table, the savings from this evaluation are in line with those found for the programs below.

Table 4-2. Percent Savings Comparison with Similar Program

Study	% Savings
-------	-----------

NYSERDA RTEM	3%
CT EEB Strategic Energy Management (SEM) Best Practices and Evaluation ³	3%-5%
Impact Evaluation of Commercial Strategic Energy Management - Energy Trust of Oregon ⁴	1-7%
ComEd and Nicor Gas Strategic Energy Management (SEM) ⁵	1%-1.5%
Enbridge Gas RunItRight (2017-2019) ^{6 7 8}	2%-4%
Union Gas RunSmart Program (2018) ^{6 7 8}	4%
Proving the Business Case for Building Analytics – LBNL. Median Energy Information System (EIS) Energy Savings ⁹	3%
Proving the Business Case for Building Analytics – LBNL. Median Fault Detection and Diagnostics (FDD) Energy Savings ⁹	9%

The LBNL referenced in the table above studies both energy information systems (EIS) and fault detection and diagnostics (FDD) systems. EIS track energy use and identify opportunities whereas FDD reduce maintenance cost, improve comfort, and find hidden energy waste. Both share attributes with RTEM which made the study a suitable candidate for comparison.

4.1.2 Savings Comparison for Sampled Sites

The evaluators used the billing data collected by the program to perform a billing analysis for the sampled sites, though evaluators took a somewhat different approach than the program with respect to the billing analysis.

The main differences between the evaluator analysis and the program analysis are summarized in Table 4-3, with additional description of each.

Table 4-3. Analysis Differences

Analysis Step/Component	Program Method/Value	Evaluation Method/Value
Utility Data	42 sites	39 sites
Weather Data	Historical and NYSERDA normal 30-year values	Historical and NYSERDA nine-year average
Regression Method	Fuels combined: single-variable regression	Fuels separated: dual-variable regression
Savings % Calculation	Sum of savings	Sum of savings
Population Extrapolation	Multifamily, Comm <25,000 ft ² , >25,000 ft ²	Ratio estimator: Calculated the SRR by calculating the ratio of weighted evaluated savings to the weighted program savings of sites

³ https://energizect.com/sites/default/files/2021-06/C1906%20SEM%20Evaluation%20Best%20Practices%20Report_FINAL.pdf

⁴ https://www.energytrust.org/wp-content/uploads/2017/03/FinalReport_EnergyTrust_CommSEM_ImpactEvaluation_wStaffResponse.pdf

⁵ https://library.cee1.org/system/files/library/13219/ComEd_Nicor_SEM_EPY8_GPY5_Evaluation_Report_2016_12_16_Final.pdf

⁶ <https://www.oeb.ca/sites/default/files/2017-2018-DSM-Custom-Evaluation-Executive-Summary.pdf>

⁷ <https://www.oeb.ca/sites/default/files/2018-DSM-Free-Ridership-Evaluation.pdf>

⁸ <https://www.oeb.ca/sites/default/files/2019-Natural-Gas-Demand-Side-Management-Annual-Verification-Report.pdf>

⁹ https://eta-publications.lbl.gov/sites/default/files/kramer_provingbuildinganalytics_october2020.pdf

		within the sample. The SRR was then applied to the program savings to get the evaluated savings.
--	--	--

The first three differences represent changes to the calculations with respect to the sampled sites, and the remaining are with respect to how those savings were extrapolated to the population. Below, we discuss the first three differences, and then show a comparison of the sampled site results before getting into the population extrapolation.

4.1.2.1 Utility Data

The program selected a random sample of projects across the population, ultimately collecting data for 42 sites. Evaluators also used the data collected by program staff; however, evaluators eliminated three of the sites that didn't have the minimum amount of data the analysts deemed was necessary to perform a billing analysis, which was set as a minimum of nine months of pre-installation and nine months of post-installation billing data. If either or both of those periods had less than nine months, the site was not used in the evaluator analysis. The utility data collected by the program was also limited by not having billing period start/end dates, or information on actual vs. estimated reads. Not knowing this information adds error to the weather normalization.

4.1.2.2 Weather Data

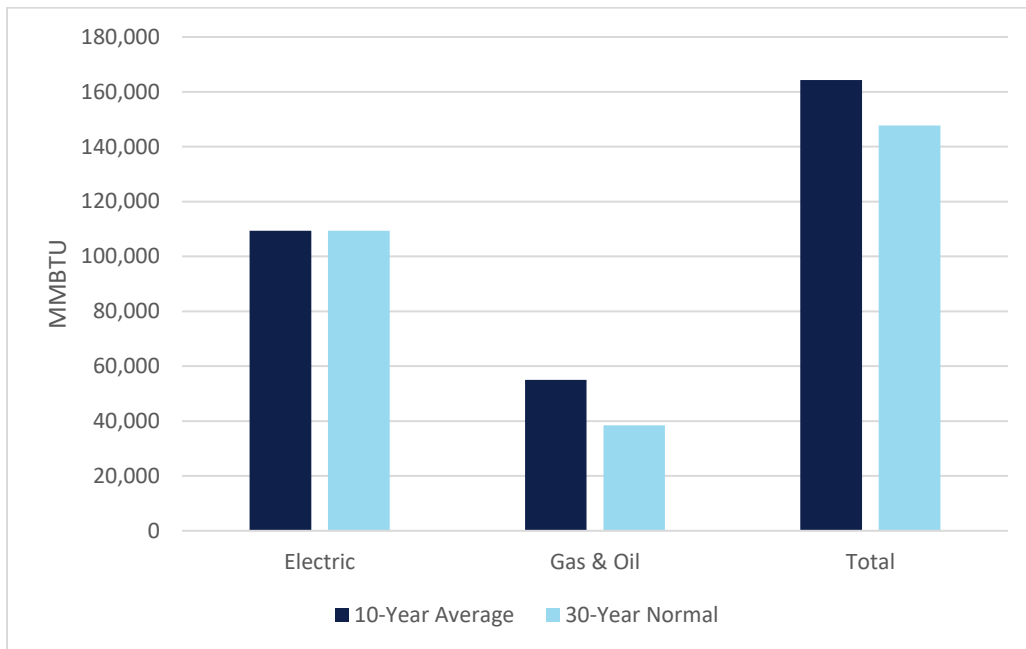
In both cases, the analysis used historical data to get a weather correlation, then applied it over conditions expected in a typical weather year. For this step, both the evaluators and program staff used weather data from the NYSERDA website; however, evaluators used the most recent nine years of weather data, rather than the "normal" year that is provided, which is representative of a 30-year timeframe. Evaluators made this selection in order to represent more recent weather patterns. Evaluators chose to use nine years instead of 10 because of a lack of complete data for the 10th year.

Evaluators also compiled savings estimates using the normal year weather data provided on NYSERDA's website in order to compare the results and estimate the difference between the two. Overall, the different weather data did not have a significant impact on the savings, and the nine-year data generally provided higher savings. Table 4-4 compares the evaluated savings values using each of the weather data sets, and Figure 4-1 presents these values graphically.

Table 4-4. Savings Comparison for Nine-Year Weather Data vs. 30-Year Weather Data

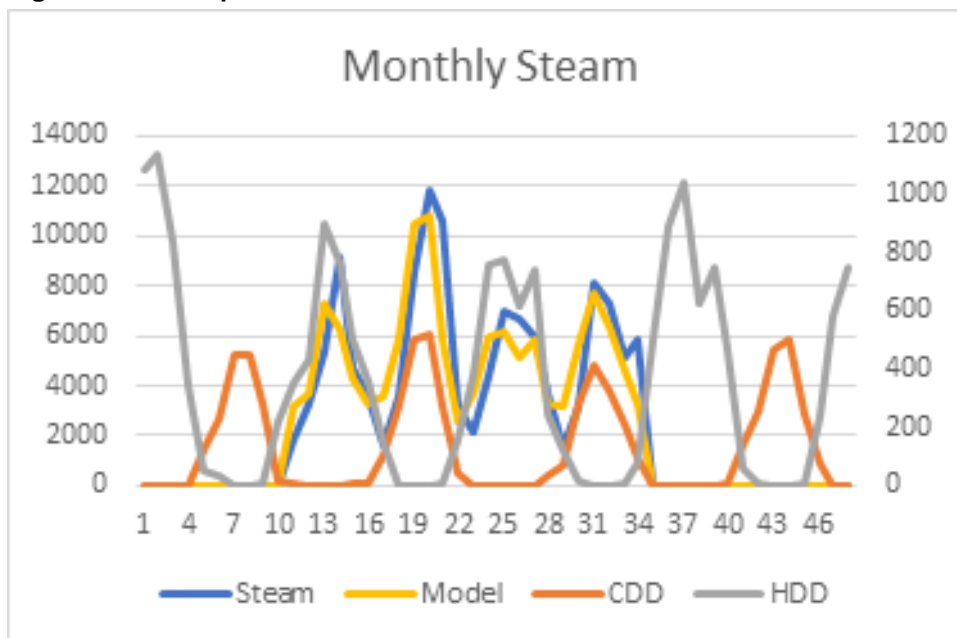
Fuel	Evaluated Savings (Nine-Year Data)	Evaluated Savings (30-Year Data)
Electric (MWh)	32,043	32,043
Gas & Oil (MMBtu)	55,029	38,401

Figure 4-1. Savings Comparison for Nine-Year Weather Data vs. 30-Year Weather Data



It is important to note that generally, we would expect that more HDD would equate to more energy savings on the “heating fuels.” However, several sites in the sample use those fuels for cooling as well. In this case, the change in savings between the 30-year data and the 10-year data is driven primarily by one site that uses steam chillers for cooling. Figure 4-2 is a plot of the HDD, CDD, and steam use for that site.

Figure 4-2. Example Site Weather Correlation



You can clearly see that an increase in CDD will also equate to an increase in summer steam use. The overall savings in this sample are being driven by a few larger sites generally, and in this case the overall steam

savings are being significantly impacted by the change in CDD, which has a fairly large impact on the overall sample savings.

4.1.2.3 Regression Method

The regression methods between the two analyses differed as well. The program approach was to convert all fuels to MMBtus and regress the combined fuels against total degree days, summing the Heating Degree Days (HDD) and Cooling Degree Days (CDD) into a single variable regression. Evaluators regressed the fuels separately, and against the degree days separately, and performed a dual-variable regression to develop separate coefficients for HDD and CDD. This approach can account for electric and fuel dependency on both HDD and CDD, where it exists.

4.1.2.4 Sampled Site Savings Comparison

With these differences described, the savings calculated for each of the sampled projects varied between the evaluation analysis and the program analysis. The figures below show a comparison of the evaluated site baseline energy consumption (Figure 4-3) and a comparison of individual site savings for the sampled sites (Figure 4-4). In each of these plots, each point represents a site, with the x coordinate representing the program value for the site and the y coordinate representing the evaluated value. The line represents a slope of 1, which would be where the site would fall if the evaluated value equaled the program-reported value. Sites above the line are sites where the evaluated value is higher than the program value, and sites below the line are sites where the evaluated value is lower.

Figure 4-3. Program and Evaluation Baseline Energy Consumption Comparison

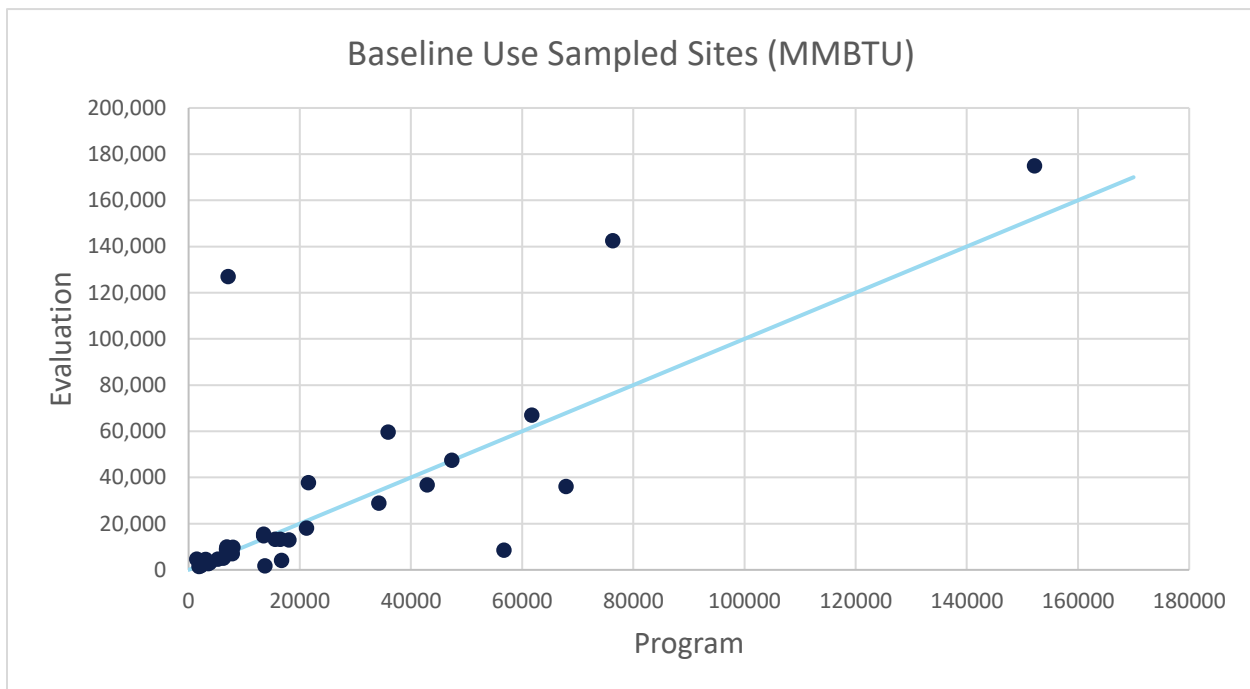


Figure 4-4. Program and Evaluation Savings Comparison for the Sampled Sites

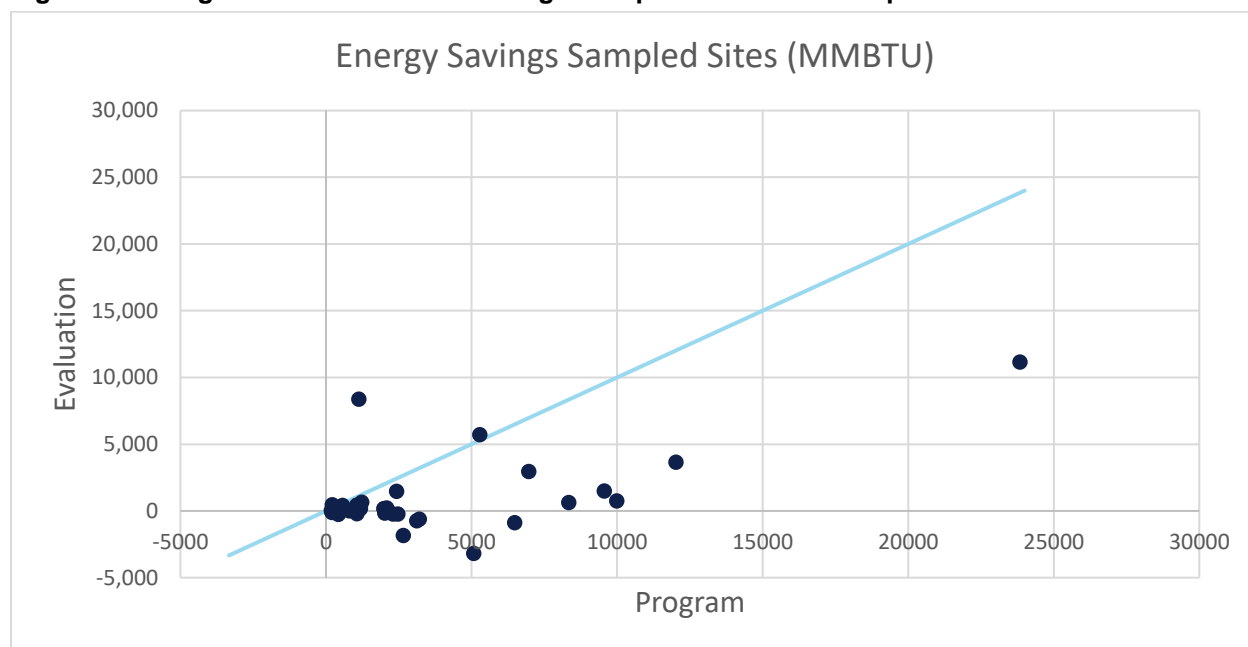


Figure 4-3 shows that, generally, the evaluator and program calculated baseline energy use for each of the 39 sampled sites are comparable. The evaluators calculated the baseline energy use based on provided utility data, whereas the program utilized the facility square footage and the appropriate CBECS EUI value to estimate baseline usage. As a result, the outlier sites (points that are not on or close to the $x=y$ line) on the figure can be explained by over or under-estimated energy usage based on the size of the facility.

The energy savings varied more significantly as demonstrated in Figure 4-4, with several of the larger sites having low realization rates that drove the overall Program SRR values. This difference is largely due to the difference in methodologies, and a spreadsheet calculation error made by the program. These discrepancies are further discussed in the section that follows.

Finally, the program assumed that all participating projects implemented measures and realized savings. However, the evaluator's billing analysis showed otherwise – to the extent that some sites showed negative savings as shown in figure 4-4 above.

4.1.2.5 Savings Calculation and Population Extrapolation

The program analysis used a percent savings calculated from the sample and extrapolated it over the population and stratified for commercial and multifamily separately. However, the evaluation calculated the savings for a sample of projects using the billing analysis approach and extrapolated by using the ratio estimator method. Ratio estimation applies the sites' weight to both the program and evaluated savings for the sampled sites and calculates the realization rate as the ratio of the weighted evaluated savings to the weighted program savings. The realization rate is then multiplied times the population claimed program savings to calculate the evaluated savings. To develop weights, the evaluation stratified over four sector types: commercial office, commercial retail, commercial other, and multifamily; and three different facility sizes: large, medium, and small.

The program included an adjustment for capital projects where they posited that sites larger than 25,000 square feet were more likely to install capital projects, such as variable frequency drives (VFDs) or other new equipment, that could have been identified by the RTEM system. Part of the RTEM Program process is for vendors to make recommendations with respect to capital projects as well. The evaluation did not include this adjustment, as any capital projects that are installed will be captured through the billing analysis of future years. Capital measures and potential savings are further discussed in Section 4.2.1 below.

When reviewing Program calculations, evaluators also discovered one additional deviation that was not due to a difference in analysis approach, but rather a calculation error. This was that the savings percentages in the Program analysis were based on an Excel pivot table that contained older, or otherwise incorrect, data, and therefore the applied values were not reflective of the most up-to-date Program M&V results. This error alone accounted for about 25% of the difference between the evaluated and the Program-reported values.

4.2 Measure-Level Analysis Results

Evaluators extracted individual measure information from the service reports provided by the RTEM vendors for 28 sites from the sample. Table 4-5 compares the energy savings of the sampled sites by fuel type for year one of the program with the billing analysis results of those same sites.

Table 4-5. Vendor Reported Energy Savings

Status	Vendor Reports					Unweighted Billing Analysis Savings for Sites with Quantified Savings in Vendor Reports		
	Number of sites	Measure count	Electric (MWh)	Fossil Fuel (MMBTU)	Combined (MMBTU)	Electric (MWh)	Fossil Fuel (MMBTU)	Combined (MMBTU)
All Recommended	28	119	8,030	35,269	62,669	3,746	9,239	22,022
Implemented	28	63	4,549	6,623	22,145			
Not Implemented	28	56	3,481	28,646	40,524			

As can be seen in Table 4-5, the savings that stem from implemented measures generally support the results from the billing analysis in aggregate for these sites. The billing analysis resulted in slightly lower electric savings and somewhat higher fuel savings. The combined electric and fuel savings are marginally different to the vendor estimated savings from the service reports. Evaluators were unable to verify the vendor provided measure level savings using system level trend data, however this comparison of the billing analysis savings with the vendor estimates for measure savings suggests that the vendor estimates are fairly reasonable in aggregate. It is important to mention that the vendor reported savings cover 28 out of the 39 sampled sites. The 11 sites that were omitted were all from one vendor, while the billing analysis and vendor report savings generally align in aggregate, this was not true for individual sites and the evaluators could not verify if that was the case for a major participating vendor.

4.2.1 Program Savings Potential

The program is interested in understanding potential savings offered beyond year one, especially as it pertains to projects that may take time to develop savings (i.e., capital projects). Upon extracting the information from each of the reports, the evaluators classified the measures as capital and non-capital to analyze the impact of both in the future.

Capital measures are projects with a capital cost associated with them. Those are projects where equipment replacement or upgrades are required to acquire energy savings. These types of measures have multiple year payback, take longer to implement, and, as a result, achieve savings after the installation is complete.

Non-capital measures are projects with little to no capital cost. Those are predominantly operational, behavioral, or control modifications that lead to energy savings. These types of measures typically present immediate energy savings and are less of a financial burden on the customer. Table 4-6 outlines the recommended, installed, and non-installed measure savings by fuel type for capital and non-capital measures.

Table 4-6. Capital vs Non-Capital Project Savings

Status	Capital*			Non-Capital		
	Measure count	Electric (MWH)	Fossil Fuel (MMBTU)	Measure count	Electric (MWH)	Fossil Fuel (MMBTU)
All Recommended	5	409.5	18,177.5	115	7,621	17,092
Installed	1	15.9	-	63	4,533	6,623
Not Installed	4	393.6	18,177.5	52	3,088	10,468

*The recommended, not installed fossil fuel savings stem from a single fuel switching measure that also contributes to electric savings penalties. Excluding this measure, the recommended electric and fossil fuel savings would be 891 MWH and 0 MMBTU respectively.

As shown in Table 4-6, 5% and 52% of electric and fossil fuel energy savings are attributed to capital measures respectively. However, only 2% of the electric and none of the fossil fuel capital energy savings measures have been installed. In contrast, 60% of the electric and 40% of the fossil fuel non-capital energy savings measures have been installed. It is important to note that the savings reported in table 4-6 are based on the vendors' service reports. The non-installed recommended savings serve as an estimate of potential future savings rather than an actual forecast.

The recommended capital measures were comprised of a single chiller replacement and four VFD measures, the average useful lives of those measures, according to the TRM values shown in Table 4-7, are 20 and 15 years respectively. On the other hand, the majority of the non-capital projects are control, tune-up, and repair measures; all of which have an average useful life of five to seven years. As a result, based on the data currently at hand, the savings persistence of the first-year savings for this program would fall in the five- to seven-year range, since most of the installed measures are non-capital projects.

Table 4-7. Measure EUL Based on Different State TRMs

Category	NY TRM	IL TRM	MD TRM	MA TRM	CA TRM	WI TRM
Capital Measures						
Chiller	20	23	23	23	20	20
VFD	15	15	15	15	15	5
Non-Capital Measures						
Boiler Controls				15 ^a	7	5
Boiler Oxygen Trim						5

Boiler Tune-Up	5	3				1
Chiller Plant Set-Point adjustment						5
Cooling System Tune-Up	5	3	5			5
Demand Control Ventilation (DCV)	15	10		10	15	10
Economizer	10 ^a	5	10 ^a	10 ^a		5
HW Supply Reset						5
Outside Air Intake Control Optimization						5
Schedule Optimization						5
Steam Trap	6		6	6		6
Supply Air Temp Reset						5
Temp Sensor Calibration						5
Valve Repair						5

^aThe TRMs for these measures accounted for the life of the installed equipment in the EUL and is therefore not a proper representation of the EUL of the control measure.

4.3 Findings and Recommendations

The evaluators offer the following three findings and ten recommendations to help improve program performance moving forward.

1. **Finding:** The Program conducts measurement and verification (M&V) and works to accurately capture Program savings. The Program has been collecting M&V data for a sample of sites that have been installed for at least 12 months for each program year, with the number of sampled sites growing as program participation grows. Moving forward, they are collecting baseline data for every site enrolled in the Program, which will provide more granularity on the baseline consumption and breakdown of fuels.

Recommendation: Evaluators recommend stratifying by two dimensions to weight the sampled projects. The first dimension is facility type, and the second is facility size. This approach will allow for more accurate representation of the population along these dimensions. Below are the recommended stratification segments, based on the population of 293 sites that were evaluated to date.

Facility Type:

- a. **Commercial Office:** These account for 38 out of 293 projects, and 50% of the total population energy use.
- b. **Commercial Other:** These account for 141 out of 293 projects, and 30% of the total population energy use. The facility types under this segment will depend on the population.
- c. **Commercial Retail:** These account for 51 out of 293 projects, and 4% of the total population energy use.
- d. **Multifamily:** These account for 63 out of 293 projects, and 16% of the total population energy use.

Facility Size:

- a. **Projects greater than 1,000,000 sq ft.**
- b. **Projects between 100,000 sq ft and 1,000,000 sq ft.**

c. **Projects less than 100,000 sq ft.**

Additionally, once the program increases its available data, we recommend sampling and extrapolating savings within the expected fuel use type.

Recommendation: We also strongly recommend that within the utility data collection, the Program should retain the actual read dates from the utility bills in order to use those dates to best align with historical weather data. In addition to those dates, we recommend retaining information on estimated vs. read values, if available. Consider revising the Program to require opting into utility data sharing with NYSERDA (not just vendors).

2. **Finding:** The Program implementers are taking many steps to calculate energy savings with reasonable and appropriate methods. They collect utility info on a sample of sites each year and are working to continuously improve those savings methodologies. The evaluators have the following recommendations with respect to that analysis.

Recommendation: When collecting utility data and conducting the billing analysis, evaluators recommend accounting for an installation period. Currently, 24 months of consecutive utility data are used in most cases for the savings analysis (targeting 12 months of pre-installation data and 12 months of post-installation data). It is likely that there was a period within that date range where the installation of the RTEM system was in progress, and it was not immediately fully operational. There is also the potential for a time lag between the installation of RTEM and operational changes that are the result of the newly installed monitoring system. Accounting for these periods and starting the post-installation billing period after installation is complete and potentially has had some time to function will likely increase savings for the program.

Recommendation: Do not apply the capital projects adjustment to acquired savings and switch to looking at a savings persistence approach, or the percent savings achieved each year. Capital projects that are recommended and installed as a result of the Program should not be claimed until they are installed. If capital projects are installed at a facility, they will be captured within the billing analysis that is conducted and will contribute to the savings once they are captured. These savings can then be mathematically extrapolated to the population using the appropriate sample weights. We also caution that many of these projects are likely to receive incentives from other programs, such as utility implementation programs. Overlap is not within the scope of this study; however, it is a factor that should be considered.

Recommendation: Evaluators recommend collecting detailed information from each site receiving M&V moving forward prior to using 2020 data in billing analyses. This information should be targeted at understanding how building operations in 2020 compared to a normal year. If facilities were operating as they typically do for part of the year, then that portion may be used. However, it is likely that 2020 data may not be representative of typical facility operation.

3. **Finding:** The evaluators extracted measure-level information from the sites' service reports and classified them as capital and non-capital to analyze the impact of both in the future. The evaluators then assigned EUL values based on the measures' closest description in the referenced TRMs. The majority of the installed savings stem from non-capital measures that typically have an EUL of 5-7 years.

Recommendation: The evaluators recommend investigating the program persistence further. Measure-level review findings show that the majority of the installed measure types include: control settings, schedules, retrocommissioning or repairs. Based on the referenced TRMs, those type of

measures have a useful life of 5-7 years. As a result, the preliminary findings indicate that the program persistence is between 5-7 years.

4. **Finding:** In reviewing the service reports provided by the vendors, the evaluators identified inconsistent and missing information. The evaluators recognize that the reports are tailored to their end user, but the current structure does not allow the program to get a full picture of the activities happening at sites due to the RTEM system. This finding is consistent with Program findings as well, and the Program is working to address this.

Recommendation: The evaluators recommend requiring site and measure-level savings information from the vendors, outlining what measures were recommended, their installation status, the energy savings by fuel associated with them, and a brief description or narrative of how the measure contributes to energy savings. This will allow the Program to both understand participant actions better and can provide supporting evidence for M&V activities. This recommendation is consistent with the Program direction moving forward.

Recommendation: The evaluators recommend including information on the mechanical equipment affected by the measures as part of the reporting requirement. The equipment size, efficiency, etc. was often missing from the reports, and having that information would allow an assessment of the vendor estimates of measure savings



Appendix A: Glossary of Terms

census – All individuals in a group. In evaluations of energy-efficiency programs, census typically refers to all the projects in a stratum of program projects.

evaluated gross savings – The change in energy consumption and/or demand that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated, as calculated by the program evaluators.

evaluated net savings – The total change in load that is attributable to an energy efficiency program, as calculated by the program evaluators. This change in load may include, implicitly or explicitly, the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, and other causes of changes in energy consumption or demand.

net savings – The total change in load that is attributable to an energy efficiency program. This change in load may include, implicitly or explicitly, the effects of spillover (SO), free riders, energy efficiency standards, changes in the level of energy service, and other causes of changes in energy consumption or demand.

nonparticipants/nonparticipating – Any customer or contractor who is eligible but did not participate in the program under consideration. Nonparticipating contractors can include contractors who have never participated in the program and contractors who formerly participated prior to the year(s) being evaluated but have not participated since.

normal replacement – The replacement of equipment that has reached or passed the end of its measure-prescribed expected useful life (EUL).

overlap (OL) – The proportion of installed measures for which customers received funding from other NYSERDA programs or other sources.

participant – An end user who receives an assessment or a service provider—assessment provider, expeditor, or finance partner—associated with the program.

relative precision – Reflects the variation due to sampling as compared to the magnitude of the mean of the variable being estimated. It is a normalized expression of a sample's standard error from its mean. It represents only sampling precision, which is one of the contributors to reliability and rigor and should be used solely in the context of sampling precision when discussing evaluation results.

About DNV

DNV is a global quality assurance and risk management company. Driven by our purpose of safeguarding life, property, and the environment, we enable our customers to advance the safety and sustainability of their business. We provide classification, technical assurance, software, and independent expert advisory services to the maritime, oil & gas, power, and renewables industries. We also provide certification, supply chain and data management services to customers across a wide range of industries. Operating in more than 100 countries, our experts are dedicated to helping customers make the world safer, smarter, and greener.