Industrial and Process Efficiency Program Impact Evaluation (2010-2012)

Final Report

April 20, 2015

Prepared for:

The New York State
Energy Research and Development Authority

Carley Murray
Project Manager

Prepared by:

ERS
120 Water Street, Suite 350
North Andover, MA 01845
Phone: 978-521-2550

Principal Investigators:
ERS
West Hill Energy and Computing, Inc.

NOTICE

ABSTRACT

This report describes the impact evaluation of the Industrial and Process Efficiency Program (IPE or, the Program). This study represents the second impact evaluation of the IPE program.

The Impact Evaluation Team assessed the measurement and verification (M&V) of energy savings of the Program for projects with measures completed between July 1, 2010, and June 30, 2012. The Impact Evaluation Team performed measurement-based engineering analysis on a sample of completed projects to quantify the evaluated energy savings by project. The evaluation did not include research into net-to-gross factors associated with free ridership and spillover. The Program electric realization rate is 0.91, demand reduction realization rate is 1.10, and the natural gas realization rate is 0.96.

Through project-level research, the Impact Evaluation Team concludes that the IPE Program has effectively implemented a large range of project types and sizes with varying technologies and stake holders. The Program produces savings estimates that are accurate and defensible.

This report presents the measured and verified energy savings of the Program at participating projects. As noted, the net impacts are not addressed in this report. This report also provides recommendations that seek to improve Program effectiveness, with consideration to potential forthcoming Program revisions associated with the Clean Energy Fund Proposal[1] and the Reforming the Energy Vision (REV) regulatory proceeding[2].

[1] *Clean Energy Fund Proposal*, New York State Energy Research and Development Authority, Proceeding on Motion of the Commission to Consider a Clean Energy Fund – Case 14-M-0094, 9/23/14

[2] *Reforming the Energy Vision*, NYS Department of Public Service Staff Report and Proposal – Case 14-M-0101, 4/24/14

ACKNOWLEDGMENTS

This final report was prepared by the Impact Evaluation Team led by ERS with Nick Collins as the project manager. Betsy Ricker of ERS was the technical lead for the project. ERS conducted the on-site verification. The Impact Evaluation Team wishes to acknowledge the significant contribution of NYSERDA Associate Project Manager Carley Murray and Senior Project Manager Cheryl Glanton, the input of the NYSERDA Impact staff, and also the valuable suggestions provided by Bill Saxonis of the DPS and its consultants, Ralph Prahl and Rick Ridge of TecMarket Works.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

SECTION 1: EXECUTIVE SUMMARY

This report describes the impact evaluation of the Industrial and Process Efficiency Program (IPE or Program). Between July 1, 2010, and June 30, 2012, 305 projects had at least one measure installed.

The objective of this impact evaluation was to estimate the evaluated savings for the Program, which includes energy and demand savings for electricity, and energy savings for natural gas. The evaluated savings are based on project-specific measurement and verification (M&V) performed on a statistically valid sample of 49 electric energy savings and 13 natural gas savings projects from the population. Many projects contained multiple measures. Table 1-1 summarizes the results of the evaluation. With realization rates (RRs) within 10% of 1.0 for all three savings parameters, the Impact Evaluation Team concludes that the program accurately estimated electric energy, electric demand, and natural gas savings.

Table 1-1. IPE Program-Reported Savings Evaluation Results (July 1, 2010 through June 30, 2012)

Parameter	Program-Reported Savings	Realization Rate	Evaluated Savings	Relative Precision
Electric energy (MWh/yr)	207,577	0.91	188,020	6.5%
Peak demand reduction (MW)	19.5	1.10	21.5	8.5%
Natural gas (MMBtu/yr)	338,385	0.96	324,071	3.8%

1.1 APPROACH

Retrospective Evaluation

The Impact Evaluation Team developed estimates of the RRs for program-reported electricity, natural gas, and demand savings by verifying for a representative sample of projects the installation of efficiency measures and through the estimation of savings. The evaluation did not include research into net-to-gross ratios (NTGR) and spillover and its hypothesized causal mechanisms. The RR was applied to the program-reported savings, resulting in the evaluated savings estimates. RRs represent and adjust the program-reported savings, upward or downward, to account for differences between the evaluated savings and program-reported savings.

The impact evaluation sample was developed with the goal of performing a rigorous and statistically significant evaluation of data center process efficiency, industrial process efficiency, and non-process efficiency projects. To do this, the evaluation electric energy sample was grouped into three categories: data center process efficiency, industrial process efficiency, and non-process efficiency projects. Projects were categorized based on the project's characterization in NYSERDA's tracking database.

The Impact Evaluation Team was able to make statistically significant statements at the Program level and for the data center efficiency projects, but not for the industrial process efficiency or non-process efficiency projects. The Impact Evaluation Team's review of each sampled project found that many projects that were characterized as process efficiency actually included significant savings from non-process measures such as lighting upgrades. Similarly, many of the non-process projects included large process efficiency measures such as industrial equipment upgrades. In the interest of providing feedback to regulators and Program staff regarding the different project types, the Impact Evaluation Team reclassified project categories and calculated RRs for the industrial process efficiency and non-process efficiency projects. This reclassification was not extrapolated to the Program level; the Program-level savings and RRs are based on the original impact evaluation sample to ensure their statistical validity. This topic is discussed further in Section 4.1.1.

Concurrent Evaluation

In parallel with the retrospective evaluation, the Program and evaluation staff are engaged in a concurrent review process whereby the Impact Evaluation Team reviews projects early in the application process and provides feedback to Program staff on baseline characterization, metering strategies, and analysis methods. This typically occurs before the Program sets incentives or program-reported savings are submitted to the New York Department of Public Service (DPS).

1.2 FINDINGS AND RECOMMENDATIONS

The Clean Energy Fund proposal and the Reforming the Energy Vision (REV) proposal suggest significant changes to NYSERDA's programs, but details of future activities and programs are limited at this point. What is clear is that NYSERA's future efforts will place emphasis on market transformation and animation. Recommendations are offered based on the findings of the evaluation with consideration to the future of NYSERDA based on the REV and Clean Energy Fund proposal[3]. The Impact Evaluation Team offers five findings and four recommendations based on the impact evaluation research.

[3] *Clean Energy Fund Proposal*, New York State Energy Research and Development Authority, Proceeding on Motion of the Commission to Consider a Clean Energy Fund – Case 14-M-0094, 9/23/14

1. The Program estimates savings well.

The most significant and overarching finding is that the Program's savings projection and review process results in accurate estimates. Not only did the RRs average near 1.0[4], the variability between the Program and Impact Evaluation Team estimates was exceptionally low. The RR error ratios ranged from 0.10 to 0.32[5], depending on the parameter. Error ratios of this magnitude are not typical of programs that fund predominantly unique projects that are subject to independent savings estimation with relatively high rigor. Error ratios this low are more typically associated with programs that fund replicable projects with consistent behavior, or with evaluations that apply low rigor to simply verify installation of the measures. This finding demonstrates the continued success of the Program in estimating project savings. The prior evaluation also found RRs near 1.0 and low error ratios, although not as low as this current evaluation study.

2. Data center cooling systems have particularly challenging baselines

While the Program and data center projects overall averaged electric energy RRs near 1.0, the unweighted RR for data center cooling measures was 0.43. The primary reason for the low RR was that the Impact Evaluation Team frequently concluded that a baseline reflecting a more efficient data center cooling operation was more appropriate than the baseline chosen by the applicant. The New York Energy Conservation Construction Code (ECCC) generally was not relevant, requiring custom baseline development for each project.

Recommendation: Expand the existing concurrent evaluation review process to include at least a sample of data center cooling and IT projects, even if they do not meet the existing size threshold for concurrent evaluation eligibility. With this mechanism, the Impact Evaluation Team and Program staff can work together to align or at least understand baseline characterizations early in the project development process.

3. Data center equipment has a short effective useful life (EUL).

In one-third of the evaluated data center information technology (IT) projects, the as-built equipment had already been replaced with newer equipment at the time of the evaluation site

---

[4] A realization rate (RR) of 1 means that the Program and the evaluators estimated savings for a project are equal. An RR of less than 1 indicates that the evaluators found less savings then the Program; an RR greater than 1 indicates the evaluators found savings greater than the Program.

[5] In energy efficiency evaluation, the error ratio is a measure of the degree of variance between the program-reported savings estimates and the evaluated estimates.

visit. Nearly half of contacts at the data center sites state that they replace their IT equipment every three years on a continuous basis. Certain facilities and customers refresh equipment as frequently as every 18 months. This rapid market evolution presents a particular challenge in identifying both baseline and high efficiency equipment or performance levels, and in keeping ahead of the pace of natural change.

Recommendations: Conduct research into data center operations, equipment obsolescence, IT equipment efficiency improvements, baseline identification and measure life to inform program support. Use the concurrent review mechanism or a rolling sample approach to identify IT projects before the traditional retrospective evaluation cycle enabling evaluation staff to visit sites while it is still possible to evaluate the pre-existing, baseline equipment performance.

4. NYSERDA is immersed in the data center equipment market.

In the course of the Impact Evaluation Team's research on data center IT equipment baselines and related topics, it was noted that NYSERDA has implemented nearly 40 data center projects throughout New York and has established relationships with the qualified engineering firms and vendors that support these data centers.

Recommendation: NYSERDA should consider leveraging existing relationships and knowledge of the data center market to identify and promote efficiency best practices in a more market-animating or transformational basis that convinces actors in this fast-paced market to choose higher efficiency options as a matter of course, perhaps reducing the need for direct project-by-project intervention by the Program.

5. The concurrent review process is effective.

The three projects that underwent concurrent review all had RRs near 1.0. This is particularly notable given the exceptional size and/or complexity of these projects. In addition, the process is allowing the Impact Evaluation Team to provide input on data collection strategies early thus increasing the level of M&V rigor for a low incremental cost. Last, the Impact Evaluation Team believes that the many discussions held between the Impact Evaluation Team, Program, site staff, and Program technical assistance staff have educated all parties regarding project costs, savings, and ratepayer funded incentives.

Recommendation: Continue to support the concurrent review process through 2015 or until the existing funds are exhausted.

SECTION 2: INTRODUCTION

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

2.1 PROGRAM DESCRIPTION

The IPE Program was developed in 2009 as an additional component to NYSERDA's Existing Facilities (EFP), New Construction (NCP), FlexTech, and Research & Development (R&D) incentive programs in order to provide incentives for cost-effective energy efficiency improvements in data centers and key manufacturing sectors in New York. This included industries such as chemicals (including pharmaceuticals), printing and publishing, automotive, food processing, forest products, agriculture, mining, extraction, and water/wastewater. Data centers are included in the Program because their process energy consumption is similar to manufacturing consumption in its load shape, process-oriented characteristics, economic development impact, power quality requirements, mission-critical nature, and load growth potential. The Program focuses on projects that reduce electric energy use per unit of production. Non-process measures, such as lighting or space conditioning of industrial facilities, are eligible as well. In 2010, incentives became available for projects that save natural gas.

Performance-based incentives are offered for projects that reduce energy in all of the above-mentioned sectors. Custom projects that save more than 500,000 kWh (1,000,000 kWh for standard facility improvements and LED lighting projects) or 10,000 MMBtu in natural gas are subject to Program measurement and verification (M&V) requirements, including pre-retrofit and/or up to 12 months of post-installation metering.

Industrial processes require customized approaches to identifying, implementing, and quantifying the savings associated with energy efficiency projects. Production lines and processes have unique characteristics and functions. IPE's project- and sector-specific approaches ensure that the best energy efficiency opportunities are identified and addressed at participating facilities. This approach maximizes process and energy reliability, productivity, and energy savings. NYSERDA also works with Outreach Contractors and Technical Reviewers who are experts in particular industrial processes and data centers. The credibility and quality of the technical assistance are essential to Program success, as are customer and stakeholder engagement.

2.1.1 Summary of Program-Reported Savings

This evaluation included projects with at least one measure installed between July 1, 2010, and June 30, 2012, The Program incentivized 305 projects during this time. Table 2-1 presents the program-reported savings for these projects.

Table 2-1. IPE Program-Reported Savings (July 1, 2010 through June 30, 2012)

Measure Type/Facility Type	# Projects with Completed Measures	Annual Savings	Percentage of Installed Savings
Electric			
Non-process/ all facility types	209	100,487,621	49%
Process/industrial	58	37,748,127	18%
Process/data center	28	29,610,967	14%
Concurrent evaluation projects	3	39,730,342	19%
Electric savings total (kWh/yr)	**298**	**207,577,057**	**100%**
Natural Gas			
All natural gas projects (MMBtu/yr)	31	338,385	100%
Natural gas savings total (MMBtu/yr)	**31**	**338,385**	**100%**
Total	**305**[1]		

[1] Both electricity and natural gas savings were claimed for 24 of the projects in the sample frame, bringing the total number of projects to 305 rather than 329 (the sum of the total electric and natural gas projects in Table 2-1).

2.2 EVALUATION OBJECTIVES

The objective of this impact evaluation was to estimate the gross savings for the Program, which includes energy and demand savings for electricity, and energy savings for natural gas. The evaluated savings are based on rigorous project-specific M&V and calculation of representative realization rates (RRs) from a statistically valid sample of projects from the population. This evaluation places particular emphasis on process-oriented projects, as this is the focus of the Program.

NYSERDA has provided the DPS with key documents for review and comment throughout the evaluation. This report complies with the M&V savings-related requirements listed in *New York Evaluation Plan Guidance for EEPS Program Administrators*[6], which was issued by the DPS and is intended to provide robust, timely, and transparent results. The methods also comply with the

[6] http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

guidelines of the *National Action Plan for Energy Efficiency (NAPEE) Model Energy Efficiency Program Impact Evaluation Guide*[7].

The overall evaluation scope and objectives are identified in Table 2-2.

Table 2-2. IPE Evaluation Scope and Objectives

Objectives	Outputs	Method Used
Evaluated energy savings	Annualized first-year evaluated energy savings based on electric (kWh) and natural gas savings (MMBtu) at the customer meter	On-site measurement and verification (M&V) for projects using on-site logging, custom engineering assessments, billing analysis, and building simulation of a representative sample of Program participants
Evaluated demand savings	Peak electrical demand savings at the customer meter calculated in a manner consistent with the New York Technical Manual (NYTM[1]) definitions	
Realization rate (RR)	Ratio of the sum of the evaluated savings divided by the sum of the program-reported savings	
Statistical validity	The sample design will target a 10% relative precision or better for Program evaluated energy savings variables at 90% confidence.	Stratified ratio estimation sample design

2.3 PREVIOUS EVALUATIONS

The IPE Program has undergone one previous impact evaluation covering measures installed during the period from Program inception in 2009 through June 30, 2010, a market characterization and assessment evaluation in 2012, and a process evaluation in 2011.[8] The prior impact evaluation estimated a 0.89 electric RR with approximately 35,000 MWh/yr of annual savings, a 1.14 RR on the two natural gas projects, and a 1.04 net-to-gross ratio.

[7] http://www.epa.gov/cleanenergy/documents/suca/evaluation_guide.pdf

[8] *NYSERDA 2009 – 2010 Industrial and Process Efficiency Program Impact Evaluation Report*, Megdal & Associates ERS with Megdal & Associates and West Hill Energy & Computing, September 24, 2012, https://www.nyserda.ny.gov/-/media/Files/Publications/PPSER/Program-Evaluation/2012ContractorReports/2012-IPE-Impact-Report.pdf

Industrial And Process Efficiency Program, Market Characterization And Market Assessment Evaluation, GDS Associates, Inc. and Navigant Consulting, Inc., May, 2012, http://www.nyserda.ny.gov/-/media/Files/Publications/PPSER/Program-Evaluation/2012ContractorReports/2012-IPE-MCA-Report-with-Appendices.pdf

and *Process Evaluation: Industrial and Process Efficiency Program*, Research Into Action, Inc., NYSERDA November 2011. , http://www.nyserda.ny.gov/-/media/Files/Publications/PPSER/Program-Evaluation/2011ContractorReports/2011-IPE-Process-Report.pdf.

SECTION 3: METHODS

The scope of work was to determine the savings of the Program. This impact evaluation consisted of two components: a retrospective M&V impact evaluation of completed IPE measures, and a concurrent evaluation of ongoing IPE projects. The methods used to estimate the savings are discussed in the following sections. The majority of this discussion pertains to the retrospective impact evaluation, which constituted the bulk of the evaluation activities and resulted in an estimate of the RR. The concurrent review was a smaller component of the evaluation and had only a small bearing on the results.

3.1 EVALUATION METHODS

A critical component of the evaluation was the development of rigorous estimates of the RRs for program-reported electricity, natural gas, and demand savings, which includes verifying the installation of efficiency measures and the generation of an independent savings analysis. The RR is applied to the program-reported gross savings, resulting in the evaluated gross savings estimates. RRs represent an adjustment to the program-reported savings, upward or downward, to account for differences between the evaluated savings and program-reported savings.

$$RR = \frac{Savings_{Evaluated}}{Savings_{Reported}}$$

where,

RR = Realization rate

$Savings_{Evaluated}$ = Savings as per measurement and verification (M&V) evaluation

$Savings_{Reported}$ = Savings as reported by the Program

Figure 3-1 provides an overview of how the evaluated projects were assigned their level of engineering analysis rigor. The level of rigor assigned to each project was based on the evaluation manager's review and consideration of the complexity of the analysis, the availability of or ability to obtain data, the magnitude of savings, customer sensitivities, and the overall budget available. Baseline characterization is a primary evaluation M&V task. The general process by which the Impact Evaluation Team characterized baseline is described in Appendix C.

Figure 3-1. Method for Assignment of Rigor

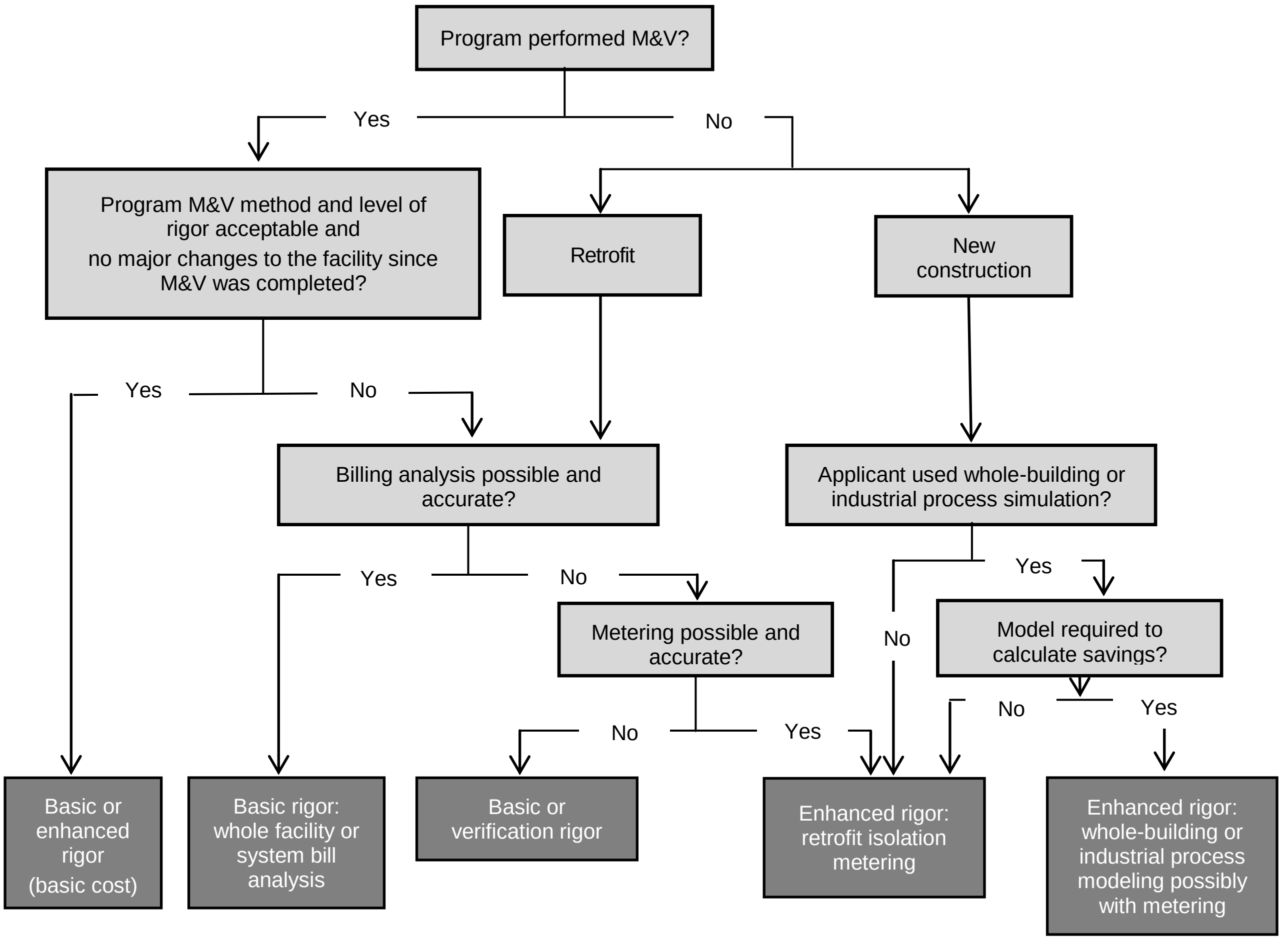


Table 3-1 summarizes the levels of rigor employed in the evaluation.

Table 3-1. Projected Level of Rigor amongst Sampled Projects

Level of Rigor	Description	Percent of Projects Receiving Level of Rigor
Enhanced IPMVP Option B or D	Typically consistent with IPMVP Option B (retrofit isolation) or Option D (calibrated simulation level analysis); this level of rigor usually includes metering of end-use energy for between 4 and 8 weeks with advanced modeling of the building or process, calibrated against field measurement of specific equipment. Enhanced rigor projects may also include whole-building simulations for projects where such methods are appropriate, such as comprehensive new construction efforts.	73%
Basic IPMVP Option A or C[1]	Typically consistent with IPMVP Option A (partially measured retrofit isolation) or Option C (whole-facility level analysis); an analysis based on production-normalized billing analysis often is this level of rigor. If projected savings exceed about 10% of the pre-installation utility bills for a specific project, bill data is available, and savings can be normalized for production, then billing analysis is used to assess the change in use for that project. Utility meters or customer sub-meters that are dedicated to the specific processes being evaluated are used if they can better isolate the energy use of changed process and increase the savings relative to metered use. Lacking billing analysis, analysts design M&V plans around spot measurements and short-term metering to supplement previously performed Program M&V with additional data collection as needed to enhance the previously reported results to the level of accuracy required for evaluation M&V.	19%
Verification	Inspection or review-only verification. If pertinent equipment is not accessible or cannot be metered over the long term, the Impact Evaluation Team uses a combination of research-based methods to determine evaluated savings. The Impact Evaluation Team verifies the installation of all or a sample of project measures through on-site inspection. To calculate savings, the Impact Evaluation Team relies on deemed savings references, measure evaluation results from other similar projects in the sample, and/or past equipment-specific performance studies either conducted or supported by NYSERDA. Verification is for those projects where a basic or enhanced level of rigor is not possible.	8%

[1] Refer to Table 2: Overview of M&V Options in *International Performance Measurement & Verification Protocol, Concepts and Options for Determining Energy and Water Savings, Volume I*, January 2012, http://www.nrel.gov/docs/fy02osti/31505.pdf.

As a practical matter, the only IPE projects evaluated at a verification level of rigor were those for which higher rigor was not possible.

3.2 SAMPLE DESIGN

Stratified ratio estimation (SRE) was used for the sample design because it allows for efficient sampling design and generally requires a lower sample size for a targeted level of precision if there is a strong correlation between the program-reported savings and the evaluated savings. As noted, the sample frame constructed includes all projects with at least one measure completed between July 1, 2010, and June 30, 2012. A summary of the sampling plan is represented in Table 3-2.

Table 3-2. Summary of the Sampling Plan

	Sample Approach	Comments
Sample frame	Program-reported data; all projects with at least one measure completed between July 1, 2010, and June 30, 2012	Program-reported data was provided by NYSERDA.
Method	Stratified ratio estimation	Correlation between program-reported and evaluation savings was expected to be strong; the kWh error ratio from the previous (2009–2010) evaluation was 0.33.
Variable to estimate	Realization rate (RR) for annual electric (kWh, kW) or natural gas (MMBtu) savings	M&V to establish evaluated savings and RR is calculated as the ratio of the evaluated savings to the program-reported savings.
Primary sampling unit	Project	A “project” refers to any project with at least one measure completed during the 7/1/10 through 6/30/12 time period. Many projects have multiple measures.
Upper-level stratification variables	Measure type (non-process, process, data center, completed projects that earlier had received concurrent reviews by evaluation team) and fuel type	Separate sampling for each fuel type and facility/measure type; fuel types are separated due to few projects with natural gas savings.
Lower-level stratification variables	Size	Size was determined by the annual kWh savings (for projects with electric savings) and MMBtu savings (for projects with natural gas savings).

3.1.1 Upper-Level Stratification

The sample design stratified projects by four project types:

1. Process efficiency projects in data centers
2. Process efficiency projects in industrial and manufacturing facilities
3. Non-process projects in all facility types
4. Concurrent review projects

Each project was assigned to a single category, based on the project or measure type shown in NYSERDA’s tracking database. Projects with both process and non-process measures or indeterminate measure types were assigned as process efficiency projects. The sample sizes were designed to increase the relative number of process efficiency and data center projects and decrease the relative number of non-process projects included in the evaluation, as these project types are the focus of the IPE Program. The sample design proposed increased the focus on M&V of data center process efficiency and industrial process efficiency projects to make the results of this evaluation more relevant to future Program activities.

Sample sizes within each upper-level stratification category were calculated by establishing a confidence/precision target for each one. Because the focus of this evaluation is on process efficiency projects, more stringent precision/confidence targets were applied to the process/data center and process/industrial upper-level stratification categories than the non-process category.

The error ratios for the non-process measures (both electric and natural gas) were estimated based on the results of the 2009/2010 impact evaluation. The target precision of the non-process measures was relaxed since these measures were extensively evaluated in the 2009/2010 evaluation and were found to have RRs that both were relatively low in variability and near 1.0, meaning that the evaluation savings and Program savings estimates were largely the same. The confidence/precision targets and assumed error ratios by stratification category are shown in Table 3-3.

Table 3-3. Populations and Target Samples by Upper-Level Strata

Measure Type/Facility Type	# Projects with Completed Measures	% Energy Savings	Target Precision	Initial Sample	Error Ratio
Electric					
Non-process/ all facility types	209	48%	20%	16	0.4
Process/industrial	58	18%	11%	24	0.6
Process/data center	28	14%	10%	12	0.7
Concurrent evaluation projects	3	19%	0%	3	N/A
Total Electric	**298**	**100%**	**10%**	**55**	**N/A**
Natural Gas					
All natural gas projects	31	100%	9%	15	0.6
Total Natural Gas	**31**	**100%**	**9%**	**15**	**N/A**
Total Combined	**305**[1]	**100%**		**70**	**N/A**

[1] Both electricity and natural gas savings were claimed for 24 of the projects in the sample frame, bringing the total number of projects to 305 rather than 329 (the sum of the total electric and natural gas projects in Table 2-1).

Non-process projects were assigned an error ratio of 0.4, which reflects the error ratio of the previous evaluation cycle (ER=0.33) that consisted largely of non-process related measures. The 0.6 and 0.7 error ratios for industrial and data center process efficiency projects, respectively, reflected the expectation for higher variation in the evaluated results of these project types and increased the number of process efficiency projects included in the evaluation sample.

3.1.2 Lower-Level Stratification

The lower-level stratification variable is project size. Size categories were based on the magnitude of project savings. Four size categories were defined per upper-level stratification category. Cutoffs were established using the method described in the *2004 California Evaluation Framework*.[9]

For each upper-level stratification category, the project size was defined based on the program-reported project electric or natural gas savings. The largest size stratum in each segment is a census stratum (all projects will be evaluated). Two additional strata were defined to allow for random sampling of the medium-sized projects in each upper-level stratification category. Table 3-4 and Table 3-5 show the evaluation electric and natural gas participant samples, broken out by upper- and lower-level stratification variables.

Projects in the lowest size stratum that accounted for less than 2% of the total energy savings for the upper-level stratification categories were not evaluated. While there are many of these small projects, they account for a small part of the overall program-reported savings and have little effect on the RR. The RR developed for the sample frame was applied to these smaller projects.

[9] TecMarket Works, et al. *The California Evaluation Framework*. Project Number: K2033910. Prepared for the California Public Utilities Commission and the Project Advisory Group. June, 2004. Pages 327 to 339 and 361 to 384. http://www.cpuc.ca.gov/NR/rdonlyres/F14E59AF-25B9-45CE-8B3C-D010C761BE8D/0/CAEvaluationFramework.pdf

Table 3-4. Electric Projects – Upper- and Lower-Level Stratification Results[1]

Upper-Level Stratum	Sampling Method	# of Projects	Maximum Savings	% of Total Electric Savings in the Stratum	Sample Projects with Electric Savings	Sample Projects –% Total Electric Savings
Non-process	Census	8	4,420,279	10%	8	10%
	Random	165	3,531,188	37%	8	3%
	None	36	125,472	2%	0	0%
	Subtotal	**209**	**8,076,939**	**48%**	**16**	**13%**
Industrial process	Census	5	5,361,274	8%	5	8%
	Random	39	2,403,894	10%	19	7%
	None	14	121,246	0%	0	0%
	Subtotal	**58**	**7,886,414**	**18%**	**24**	**14%**
Data center process	Census	7	9,053,565	10%	7	10%
	Random	15	887,588	4%	5	1%
	None	6	238,717	0%	0	0%
	Subtotal	**28**	**10,179,870**	**14%**	**12**	**11%**
Concurrent review	Census	3	16,466,200	19%	3	19%
	Subtotal	**3**	**16,466,200**	**19%**	**3**	**19%**
Electric totals		**298**	**N/A**	**100%**	**55**	**58%**

[1] Of the electric saving sampled projects, one includes program-reported fuel savings.

Table 3-5. Natural Gas Projects – Upper- and Lower-Level Stratification Results[1, 2]

Upper-Level Stratum	Sampling Method	# of Projects	Maximum Source Savings	% of Total Natural Gas Savings in the Stratum	Sample Projects with Natural Gas Savings	Sample Projects – % Total Natural Gas Savings
Natural gas	Census	8	93,299	69%	8	69%
	Random	17	13,140	29%	7	15%
	None	6	2,159	2%	0	0%
Natural gas totals		**31**	**N/A**	**100%**	**15**	**84%**

[1] Of the natural gas savings sampled projects, nine include program-reported electric savings.

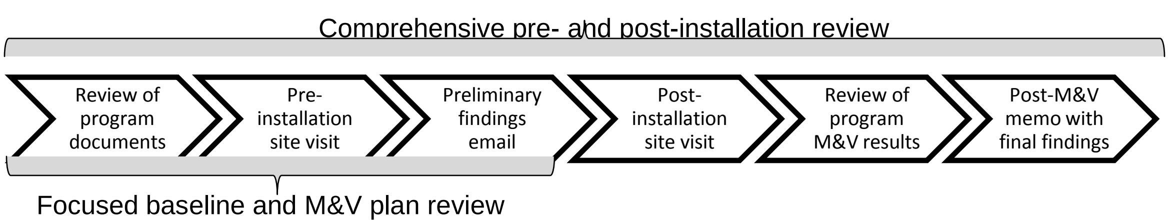
[2] There are four projects that are in both gas and electric samples.

3.2 CONCURRENT REVIEW

The Impact Evaluation Team implemented a concurrent review process on select projects in parallel with the retrospective impact evaluation discussed in this report. Concurrent, or pre-installation, review occurs in parallel with the Program's application and review process, including the development and execution of Program M&V. This work is designed to include in-depth review and feedback on large, complex projects selected by the IPE Program staff and the Impact Evaluation Team project manager. It provides the Program staff with feedback on the baseline conditions, M&V methods, measure characterization, building operating assumptions, and possible lost opportunities based on information provided by the implementer or site engineer.

The concurrent review process includes two levels of rigor: comprehensive pre- and post-installation review and focused baseline and measurement and verification plan review. A level of rigor is assigned based on the consensus of Program staff and evaluators after considering various project attributes including measure type, seasonality, interactivity, magnitude of savings, and the availability and quality of data. The steps in the typical concurrent review process are shown in Figure 3-2.

Figure 3-2. Typical steps in the concurrent review process



For all concurrent review projects, the Impact Evaluation Team works with the implementation staff and their technical review contractors to incorporate evaluation-grade techniques into their scope of work.

Focused Baseline. Baseline conditions are assessed using data collected during pre-installation site inspections in combination with Program data, which include design reports, pre-installation metering and program forms required by NYSERDA, and (if needed) data obtained from the participant, including existing equipment, production volumes, and operating schedules. For projects undergoing a focused baseline and M&V plan review, the concurrent review process concludes with a memo that is filed with NYSERDA evaluation and implementation staff following the evaluator's pre-installation site visit.

Comprehensive. For projects undergoing comprehensive pre- and post-installation review, the concurrent review process continues through project completion. The Impact Evaluation Team reviews the Program's post-installation M&V reports and performs post-installation site inspections. Upon completion of each project review, the Impact Evaluation Team provides a report summarizing the actions taken and the recommendations made to NYSERDA evaluation and implementation staff.

The concurrent review process was initiated in May 2011 and has a planned completion in December 2015[10]. As noted in Table 3-5, three of the projects that underwent the concurrent review process were installed between July 1, 2010 and June 30, 2012[11] are included in the retrospective impact evaluation. There were no other concurrent review projects installed between July 1, 2010 and June 30, 2012. The results of those three projects are discussed in Sections 4.1.1 and 4.3. Concurrent projects included in this retrospective evaluation were assigned to an independent third party firm for review to prevent any conflict of interest.

3.3 Revisions to Program-Reported Savings During the Course of Evaluation

The NYSERDA IPE program initially reports completed project savings to the DPS based on the estimate available at the time of project installation; once savings are reported to DPS, the project is subject to evaluation. However, the Program requires that large projects receive up to one year of post-installation M&V. This means that the savings initially reported to the DPS may be updated a year or more after project installation based on the results of post-installation Program M&V. This reporting approach causes challenges for retrospective evaluation sampling, as the savings for some subset of projects in the evaluation sample frame is not final.

The Impact Evaluation Team used the following approach to calculate an adjusted RR for this subset of projects. This approach is detailed in Appendix B, but fundamentally, the approach is as follows:

1. The Impact Evaluation Team performed all sampling, evaluation M&V, and related work in the same fashion as if there was no ongoing Program M&V that could modify the originally reported savings. The result is the original program RR.

[10] It should also be noted that the concurrent review work is expected to extend through December 2015. This report provides a summary of the concurrent review process and projects through December 2014.

[11] The sample frame for the retrospective impact evaluation included projects installed between July 1, 2010 and June 30, 2012.

$$RR_{original} = \frac{Evaluated}{Reported_{original}}$$

2. At the conclusion of Program M&V, the Program provided a new data set of record for all projects and measures in the original frame subjected to post-installation Program M&V. The new data set included projects where program-reported savings changed compared to the original data set as a result of Program M&V.

3. The Impact Evaluation Team and Program staff reviewed the projects with changes to ensure that the new data set did not include savings for new measures added to old projects or reasons other than post-installation M&V adjustment.

4. The adjusted RR was calculated by dividing the evaluated savings by the revised and adjusted program-reported savings.

$$RR_{adjusted} = \frac{Evaluated}{Reported_{adjusted}}$$

The adjusted RRs, those reflecting the most current program-reported savings data, are used in the final savings analysis.

SECTION 4: RESULTS, FINDINGS, AND RECOMMENDATIONS

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

4.1 ELECTRIC ENERGY SAVINGS RESULTS

This section summarizes the results of the measurement and verification (M&V) activities for electric energy projects.

4.1.1 Program Electrical Energy Savings and Realization Rates

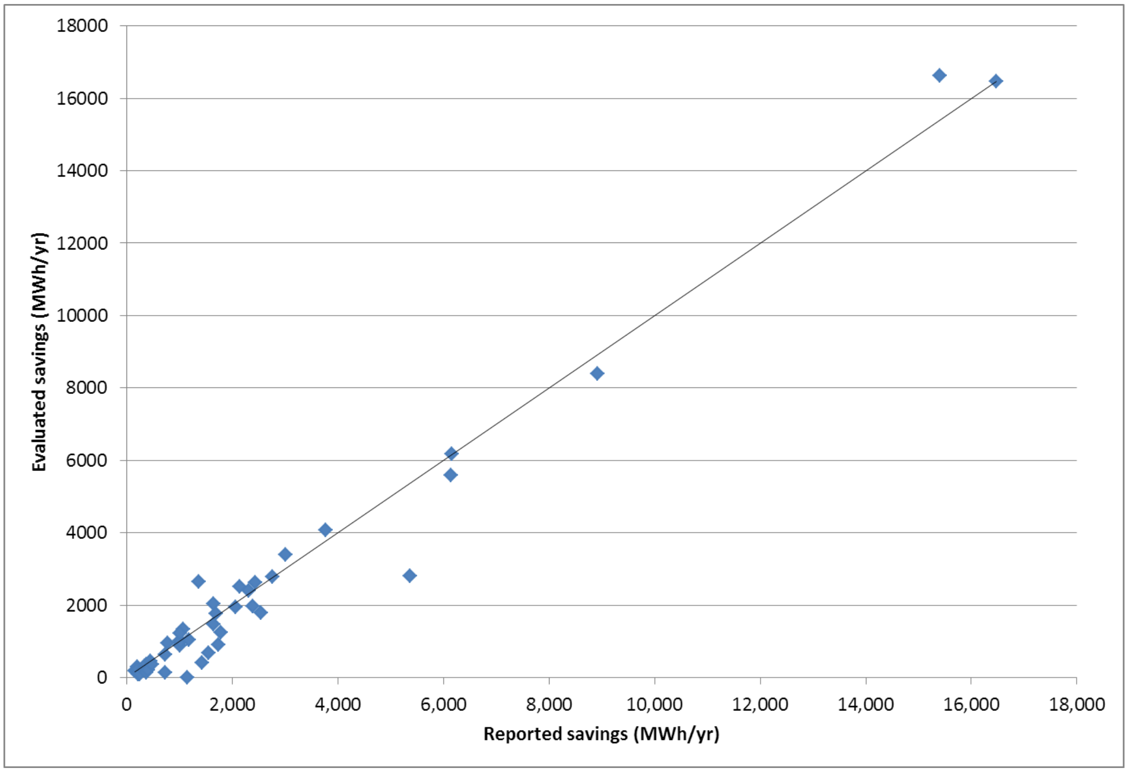
The realization rate (RR) for the electric energy savings of the program, calculated as the evaluated savings divided by the program-reported savings, is 0.91. Table 4-1 provides the key program results including the reported savings, realization rate, evaluated electric energy savings, relative precision and error ratio.

Table 4-1. Reported and Evaluated Electric Energy Savings (July 1, 2010 through June 30, 2012)

Parameter	Program-Reported Savings (MWh/yr)	Realization Rate	Evaluated Savings (MWh/yr)	Relative Precision	Error Ratio
Electric energy (MWh/yr)	207,577	0.91	188,020	6.5%	0.29

Figures 4-1 illustrates the evaluated annual electric energy savings compared with that reported by the Program. For a RR of 1, the evaluated savings would match the program-reported savings. This ideal is shown as a solid black line on the charts. Actual findings are plotted as points on the graph. A pattern of points below the ideal line illustrates an RR of less than 1; points above the line illustrate an RR greater than 1. The error ratio is a measure of the degree of variance between the program-reported savings estimates and the evaluated estimates. The higher the error ratio, the greater the amount of scatter between points.

Figure 4-1. Reported and Evaluated Electric Energy Savings (July 1, 2010 through June 30, 2012)



The program-level electricity savings error ratio was calculated to be 0.29. This is lower than the 0.4, 0.6, and 0.7 error ratios shown in Table 3-3, that were assumed in estimating the non-process, industrial process, and data center process sample sizes, indicating less scatter in the point distribution than was assumed when estimating sample sizes, and meaning that the evaluation was able to attain better sampling precision than originally anticipated.

Upper-Level Stratification Category Results

The evaluation sample included four upper-level stratification categories that were generated based on the project categories included in NYSERDA's tracking database: 1) data center process efficiency ("data centers"), 2) industrial process efficiency ("process"), 3) non-process efficiency, and 4) concurrent projects. These categories were chosen with the goal of being able to make statistically valid statements about each type of project. Data center and industrial process efficiency projects were of particular interest because they were under-represented in the previous impact evaluation of the Program. As shown previously in Table 3-3, the data center, process, and non-process categories were sampled with target precision values of 0.10, 0.11, and 0.20, respectively. The fourth upper-level stratification category, concurrent, was a census stratum that included three projects that were subject to review by the Impact Evaluation Team through the concurrent evaluation. Table 4-2 provides the evaluation results by upper-level stratification category.

Table 4-2. Electrical Energy Results by Upper-Level Stratification (July 1, 2010 through June 30, 2012)

Stratum	N (pop)	n (samp)	Realization Rate	Percent of Evaluated Savings	Relative Precision	Error Ratio
Non-process/all facility types	209	16	0.90	49%	13%	0.29
Process/industrial	58	21	0.77	15%	6%	0.40
Process/data centers	28	9	0.95	13%	13%	0.35
Concurrent evaluation projects	3	3	1.02	23%	Census	0.05

Data Center, Process, and Non-Process Project Results

The results in Table 4-2 reflect the project category assigned by NYSERDA and entered into their project tracking tool. During the sample design, the Impact Evaluation Team did not have any project-specific information with which to assess the assigned project category. Project review conducted after the sample design was complete and project-specific data was received found that the data center efficiency projects were correctly categorized; however, some of the projects classified as non-process based on Program tracking were actually process projects. Similarly, some of the process efficiency projects were actually non-process projects. This difference in project categorization does not impact the statistical validity of the Program-level RRs or the data center RRs. However, it does prevent the Impact Evaluation Team from making statistically significant statements about the performance of process efficiency projects vs. non-process efficiency projects.

Given these misclassification issues, the Impact Evaluation Team re-categorized the process and non-process projects after the impact evaluation was complete to determine if feedback could be provided about the performance of process and non-process efficiency projects. The Impact Evaluation Team did not include the concurrent review projects in the process/non-process categories because it was more informative to present the process/non-process results independent of those concurrent projects that included early evaluator involvement. The reclassification of process/non-process projects after sample design did not include the original sample weights as these were no longer valid after the projects were re-categorized. Instead, the Impact Evaluation Team calculated the average RR for the projects. In this average, each RR was given the same weight to ensure that bias was not unintentionally introduced based on project size. This reclassification resulted in three projects being moved from the non-process to the process efficiency category and changed the average RR for each category by less than 2%[12]. Despite this small change, there is still too much uncertainty as to the proper weight of each project within the population to conclude conclusively that there is a material difference between the performance of process and non-process projects.

Concurrent Project Results

The impact evaluation included three projects that had previously undergone concurrent review. These three projects were found to have RRs of 1.08, 1.00, and 0.94. Together, these three projects achieved an RR of 1.04. The proximity of these RRs to 1.00 suggests that the concurrent review process is effective. However, these projects also revealed ways in which the process could be improved, including increasing the rigor of the Impact Evaluation Team's baseline research and final project savings review.

4.1.2 Program Electrical Peak Demand Savings and Realization Rates

The RR for the peak demand reduction, calculated as the evaluated reductions divided by reported reductions, is 110%.

The Impact Evaluation Team found that peak demand savings were not consistently accounted for in the applicants' analyses and often were not reported by the Program. Demand savings as calculated by the applicants did not always make use of the peak demand period definition prescribed in the New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs (also referred to as the New York Technical Manual or NYTM)[13]. The Impact Evaluation Team calculated the peak demand reductions of the sampled projects in accordance with the peak periods as defined in the NYTM.

Table 4-3 provides key program results including the reported peak reductions, the realization rate, the evaluated peak reductions, the relative precision and error ratio.

[12] This is the average of the realization rates for the process and non-process projects. This average is exclusive of the sample weights derived in the evaluation sample that was described in Section 3.2.

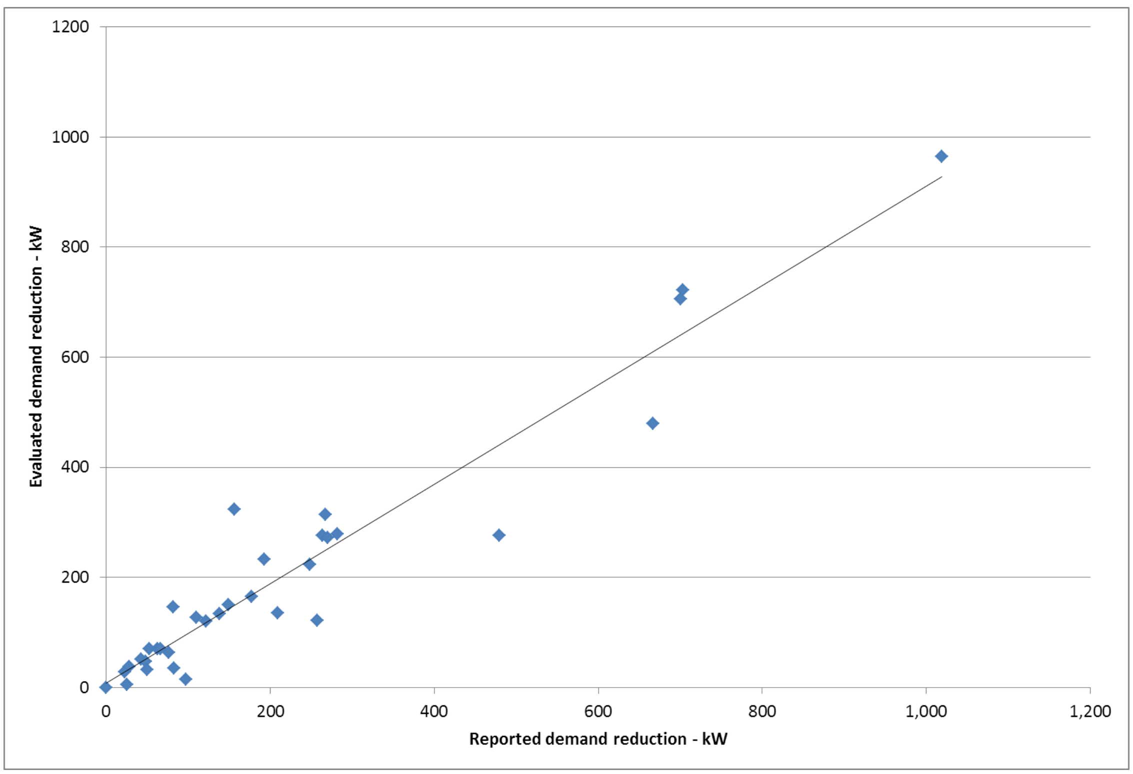
[13] Peak demand calculated using the approach recommended on page 8 of the New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs (also referred to as the New York Technical Manual)
http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/06f2fee55575bd8a852576e4006f9af7/$FILE/TechManualNYRevised10-15-10.pdf

Table 4-3. Program Reported and Evaluated Electric Demand Savings (July 1, 2010 through June 30, 2012)

Parameter	Program-Reported Peak Reductions (MW)	Realization Rate	Evaluated Peak Reductions (MW)	Relative Precision	Error Ratio
Peak demand reduction (MW)	19.5	1.10	21.5	8.5%	0.32

Figures 4-2 illustrates the evaluated peak demand reduction compared with that reported by the Program. As with Figure 4-1, the ideal result, an RR of 1 is shown as a solid black line on the charts. Actual findings are plotted as points on the graphs. The peak demand error ratio is 0.32.

Figure 4-2. Program Reported and Evaluated Peak Demand Reductions July 1, 2010 through June 30, 2012



4.1.3 Difference between Program and Evaluated Electrical Energy Savings

For each project with a RR other than 1, a difference analysis was performed to identify the major driver or drivers of the RR. The difference analysis results are aggregated in an attempt to identify systematic differences in methods between the Program and the Impact Evaluation Team. The results of the difference analysis are presented in Figure 4-3. This figure shows the impact of the difference in terms of increased and decreased savings. Differences that increased RR are shown to the right of the 0.0% line, and differences that reduced RR are to the left.

Figure 4-3. Differences Analysis Results – Electric Energy

Consolidated Categories	# of occurances	Impact on kWh
Operating Characteristics	11	-9.59%
Pre-post inspection	3	-3.39%
Analysis Assumptions	3	-1.52%
Administrative	2	-0.01%
Baseline	1	4.77%
Total	**20**	**-9.74%**

Figure 4-3 above demonstrates key contributors to the 0.91RR for electric energy savings. Some of the major drivers to RR are discussed below:

- Different operating characteristics were the largest difference category that reduced the RR. This category includes differences in equipment efficiency, load profile, inoperable installed equipment, and run hours.
  - A single project accounts for nearly 60% of the differences associated with the operating characteristics.
  - The remaining differences associated with this category are associated with a project where the energy consumption was closely tied to ambient weather conditions. Program M&V occurred for three months, and the typical energy use was extrapolated out over a year. The Impact Evaluation Team reviewed 12 months of data and found the system to be more dynamic in response to ambient conditions then was extrapolated by the applicant.
- Pre- and post-inspection is the second-largest category driving the RR down. This category includes differences in installed technologies, quantities, or projects where the pre-project characterization was found to be different.
  - The primary driver in this category was a single project where nameplate data was used to predict the proposed energy consumption. The trended data analyzed by the Impact Evaluation Team found significantly less energy use than indicated by nameplate data.
- These differences generally represent stand-alone events. The analysis found no systematic differences when compared with the evaluation.

4.1.4 Additional Electrical Energy Savings Results

The Impact Evaluation Team also reviewed the evaluated projects to determine if patterns could be identified or feedback provided about project performance based on measure type, project location, and other key project features. The resulting observations are presented below. Although these results do not adhere to the same 90/10 confidence precision targets as the evaluation sample, the Impact Evaluation Team is presenting the results as they are useful for providing feedback on project and measure performance.

Electric Energy Savings by Measure Type

Miscellaneous data center process efficiency and lighting were the most common measure types evaluated. The miscellaneous data center process efficiency category includes several technologies such as data center cooling, server virtualization, power distribution upgrades, and physical server upgrades.

While the program and data center projects overall averaged electric energy RRs near 1.0, the unweighted RR for data center cooling measures was 0.43. The primary reason for the low RR was that the Impact Evaluation Team concluded on some projects that a different baseline reflecting more efficient data center cooling operation was more appropriate than that chosen by the applicant. The New York Energy Conservation Construction Code (ECCC) generally was not relevant, requiring custom baseline development for each project. In two other examples, the energy demand associated with the computer room air conditioning (CRAC) units was assumed to drop to 0 kW when in economizing mode; however, CRAC unit fans continue to operate while in economizing mode and this fan energy use was accounted for by the Impact Evaluation Team.

The third most prevalent was the miscellaneous industrial process efficiency group. This group contains numerous and varied specific measures including process consolidation (lean manufacturing), injection molding machines, industrial dryers, and industrial gas generation. Figure 4-4 provides a summary of measures within the sampled projects.

Figure 4-4. Number of Evaluated Measures by Measure Group (July 1, 2010 through June 30, 2012)

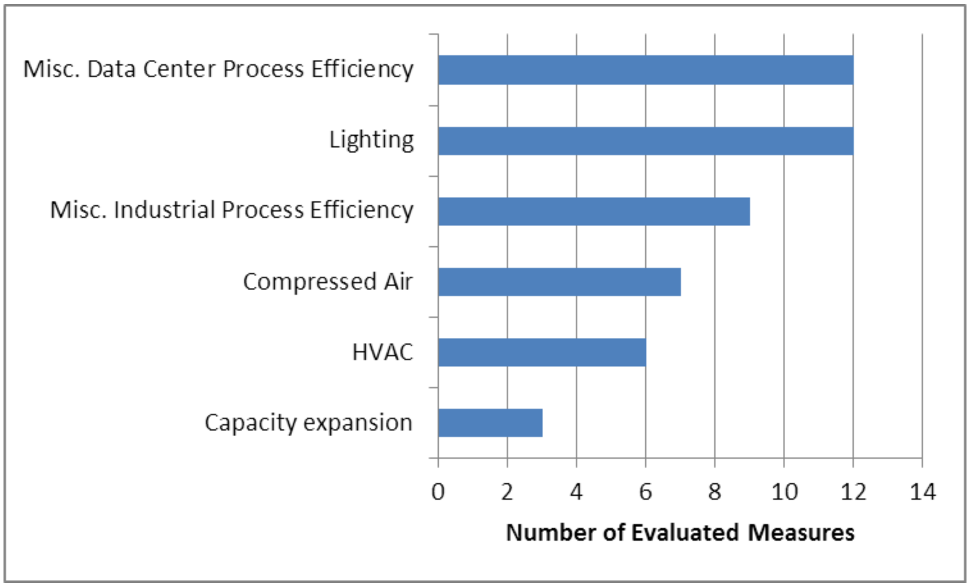


Figure 4-5 provides the savings by measure type for the sampled projects.

Figure 4-5. Electric Energy Savings by Measure Group (July 1, 2010 through June 30, 2012)

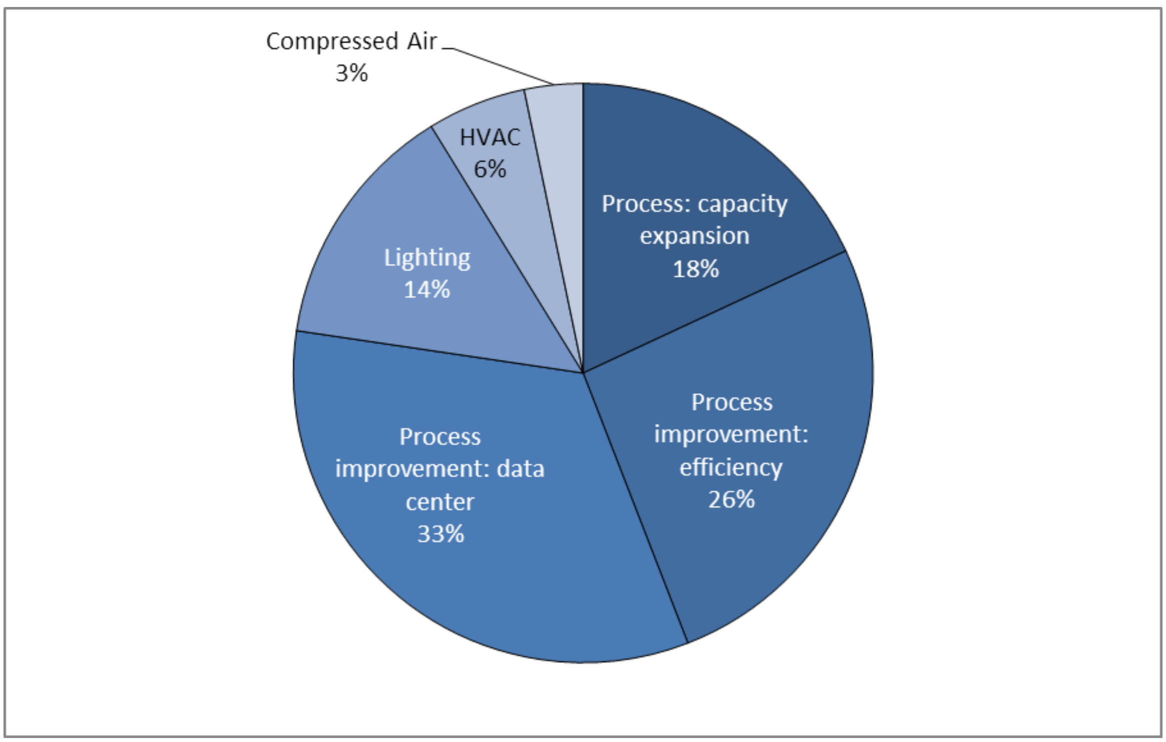


Table 4-4 presents the RRs for each measure group. Since measure type was not a unique sampling stratum, the measure level results are only representative of those projects evaluated. These RRs should not be extrapolated out to the Program level. Only RRs based on sampled stratum can be used to make statistically valid statements about the Program (see Table 4-2 above).

Table 4-4. Unweighted Electric Energy Realization Rates by Measure Group (July 1, 2010 through June 30, 2012)

Measure Group	Realization Rate	Savings as percent of total
Process: capacity expansion	0.99	18%
Compressed Air	0.55	3%
HVAC	0.84	6%
Lighting	1.01	14%
Process improvement: data center	0.89	33%
Process improvement: efficiency	1.07	26%

It was noted that as a group, compressed air measures experienced the lowest RR. All seven compressed air measure projects received RRs of less than 1. In almost all cases this is due to the evaluation M&V demonstrating a different load profile than that included in the program-reported savings. Baseline characterization differences also drove the RR down; one capacity expansion project did not account for load growth in the baseline characterization, which reduced the savings. Four of these seven projects received Program M&V, but the RR for these projects differs little from the RR for projects that did not receive Program M&V. File data suggests that the Program M&V periods were insufficient in length or, more importantly, did not record a period of time which was representative of typical operation.

Although the compressor projects had a minor influence on the overall Program results, the opportunity exists to improve Program savings estimates by deferring Program M&V until new or retrofitted compressed air systems are operating in their typical state.

Electric Energy Savings for Update and Downstate Projects

NYSERDA has historically been interested in whether there are material differences in the performance of projects located in upstate New York and downstate New York. The Impact Evaluation Team classified each of the projects in the sample frame into one of these two categories based on their electrical utility[14] and determined that 83% of the program-reported savings was attributable to upstate projects while 17% of the program-reported savings was attributable to

[14] The Impact Evaluation Team classified all projects served by Con Edison as downstate projects. All other projects were classified as upstate projects.

downstate projects.[15] The Impact Evaluation Team then reviewed the electrical energy savings of the sampled sites and calculated the upstate project RR to be 92% and the downstate project RR to be 87%. Based on this, and previous review of the impact evaluation results by project type, the Impact Evaluation Team does not believe there is a correlation between project location and RR for the IPE program.

4.1.5 Influence of Program Measurement and Verification Activities on Electric Energy Savings Estimates

For the evaluation period, the Program requires that any project with greater than 500,000 kWh in electrical savings[16] receive M&V. M&V is performed by the applicant and a third-party Technical Reviewer assigned to the project.

The RRs for projects with and without completed Program M&V are presented in Table 4-5.

[15] As presented and discussed in the final *NYSERDA Industrial and Process Efficiency Program FINAL Detailed Evaluation Plan* that was submitted to the DPS on June 21, 2013

[16] Lighting projects require measurement and verification if savings are greater than 1,000,000 kWh.

Table 4-5. Program M&V Verification Results for Electric Savings (July 1, 2010 through June 30, 2012)

Program M&V	n (sample)	Electric Energy Realization Rate
Conducted	34	0.92
Not-conducted	15	0.85

For electrical energy projects, the projects that completed Program M&V showed higher RRs than those projects that did not receive Program M&V. This finding suggests that the Program M&V efforts have been successful in improving the electrical energy savings estimates associated with the projects, but given the relatively close RRs for the two types of projects, the Impact Evaluation Team does not find this result to be conclusive. Additionally, these values are subject to change as projects proceed from the PIR stage to the M&V stage of their Program review.

Of the 15 projects where Program M&V was not conducted, seven are lighting projects and account for 59% of the weighted savings of projects that did not receive Program M&V.

4.2 NATURAL GAS SAVINGS RESULTS

This section summarizes the results of the M&V activities for natural gas projects. In 2010 funding for natural gas energy savings projects became available; as such, natural gas project participation has increased since the last evaluation.

4.2.1 Program Natural Gas Savings and Realization Rates

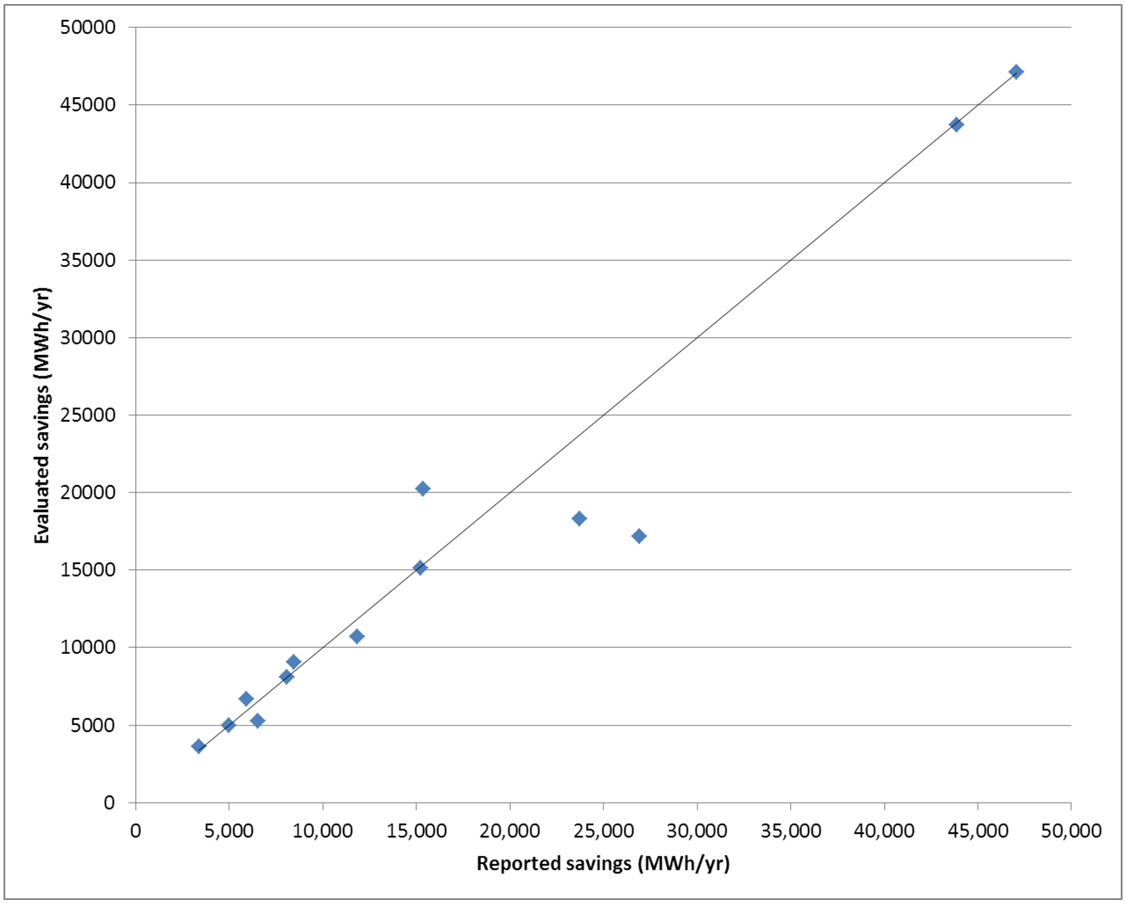
The RR for the natural gas energy savings of the program, calculated as the evaluated savings divided by the program-reported savings, is 0.96. Table 4-6 provides the key results including reported savings, the realization rate, evaluated natural gas energy savings and error ratio.

Table 4-6. Program Reported and Evaluated Natural Gas Energy Savings (July 1, 2010 through June 30, 2012)

Parameter	Program-Reported Savings (MMBtu/yr)	Realization Rate	Evaluated Savings (MMBtu/yr	Relative Precision	Error Ratio
Natural gas (MMBtu/yr)	338,385	0.96	324,071	3.8%	0.10

Figures 4-6 illustrates the evaluated annual natural gas energy savings compared with that reported by the Program. Ideally, the evaluated savings would always match the program-reported savings. This ideal is shown as a solid black line on the charts. The actual findings are plotted as points on the graphs. The error ratio assumed for estimating sample size is shown in Table 3-3 above.

Figure 4-6. Program Reported and Evaluated Natural Gas Energy Savings (July 1, 2010 through June 30, 2012)



The natural gas energy savings error ratio was calculated to be 0.1. This is lower than the 0.6 error ratio that was assumed in estimating the sample size, indicating less scatter in the point distribution than was assumed when estimating sample sizes, and meaning that the evaluation was able to attain better sampling precision than originally anticipated.

4.2.2 Differences between Program Reported and Evaluated Natural Gas Savings

For each project with an evaluated RR other than 1, a difference analysis was performed to identify the major driver or drivers to the evaluated RR. The difference analysis results are aggregated in an attempt to identify systematic differences in approach between the Program and the Impact Evaluation Team. The results of the difference analysis are presented in Figure 4-7.

Figure 4-7. Differences Analysis Results – Natural Gas (July 1, 2010 through June 30, 2012)

Consolidated Categories	# of occurances	Impact on MMBtu
Operating Characteristics	5	-7.9%
Pre-post inspection	1	-0.9%
Analysis Assumptions	1	3.1%
Administrative	1	0.0%
Total	**8**	**-5.8%**

Figure 4-7 above illustrates key contributors to the 0.96 RR for natural gas energy savings. Some of the major drivers to RR are discussed below:

- Different operating characteristics category has the largest impact on RR. This category includes differences in equipment efficiency, differences in load profile, inoperable installed equipment, and differences in run hours.
    - Two projects constitute the bulk of the differences associated with this category. In one project, the as-built run times were found to be almost half of what was predicted. For the second project, the as-built energy intensity found through the evaluation varied from the energy intensity estimated by the applicant.
- The difference analysis did not find any systematic differences; the Program is accurately estimating savings with methodologies and data sources similar to those used in the rigorous post-installation evaluation.

4.2.3 Additional Natural Gas Savings Results

The Impact Evaluation Team also reviewed the evaluated projects to determine if patterns could be identified or feedback provided about project performance based on measure type, project location, and other key project features. The resulting observations are presented below. Although these results do not adhere to the same 90/10 confidence precision targets as the evaluation sample, the Impact Evaluation Team still feels that the results presented are useful for providing feedback on project and measure performance.

Natural Gas Savings by Measure Type

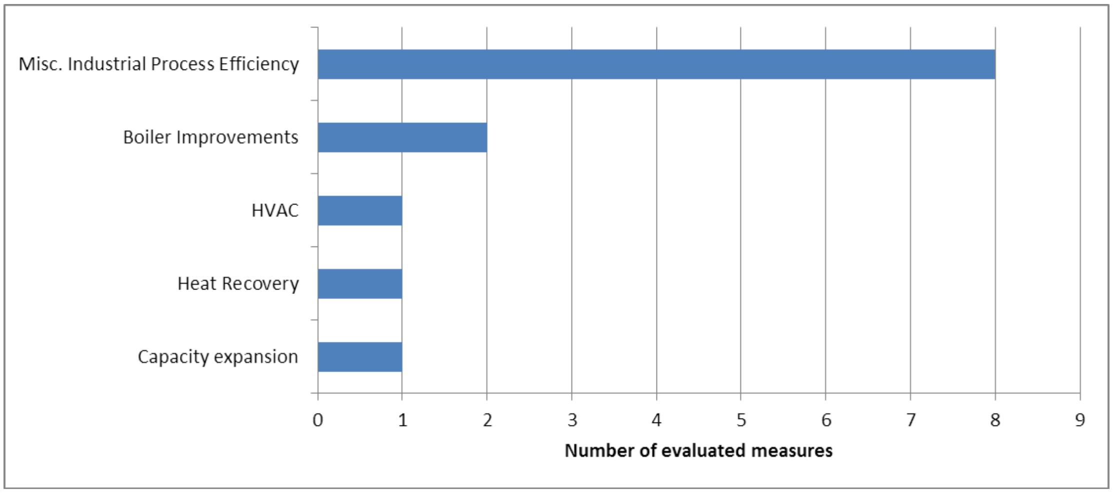
Figure 4-8 reflects the measure types evaluated on-site. There is little comparison to be made to the previous evaluation cycle, as it only included two natural gas projects. This evaluation by comparison drew a sample of 15 projects from a population of 31.

Process efficiency measures dominated the sample with more such measures (eight) than all other types combined (five). Examples of evaluated process efficiency measures include:

industrial ironers, process consolidation (lean manufacturing), regenerative thermal oxidizers, and industrial dryers.

Boiler improvements, the second most common measure type includes boiler control improvements and boiler economizers.

Figure 4-8. Number of Evaluated Measures by Measure Group – Natural Gas (July 1, 2010 through June 30, 2012)



Industrial process efficiency improvements made up 59% of the weighted evaluated savings of the Program. Boiler improvements accounted for 22% the total as shown in Figure 4-9.

Figure 4-9. Natural Gas Savings by Measure Group (July 1, 2010 through June 30, 2012)

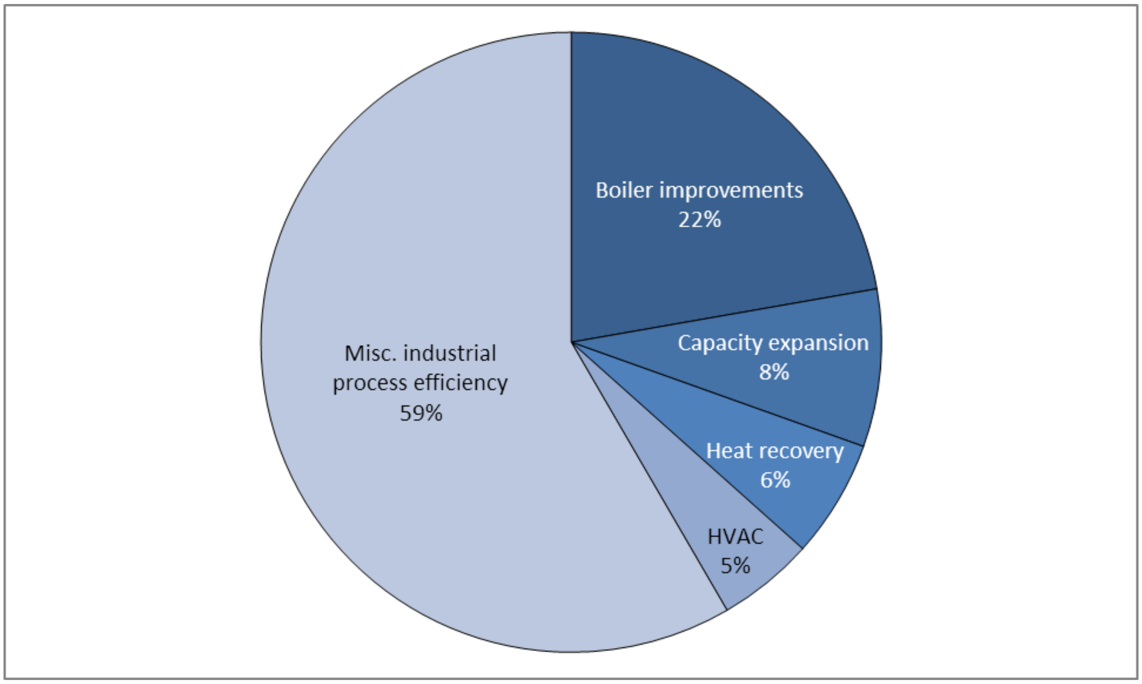


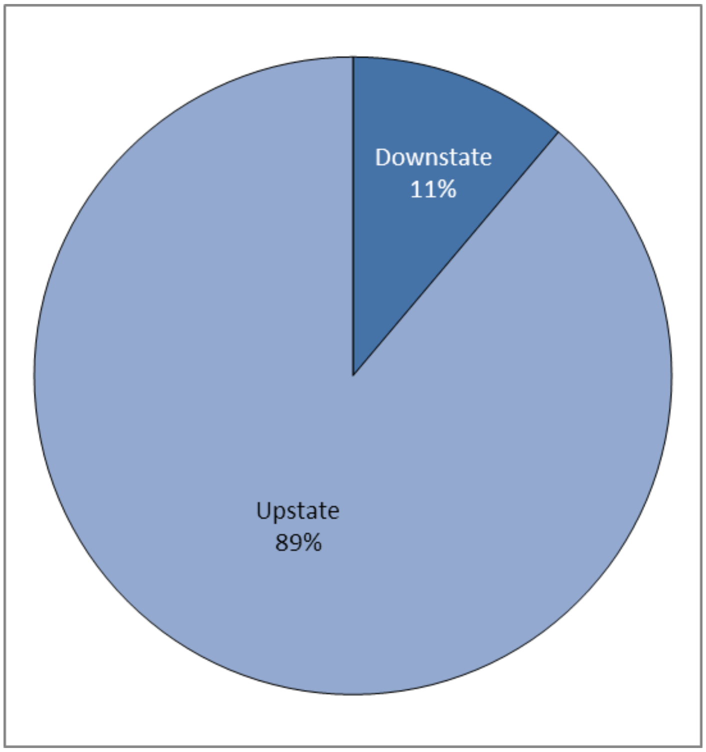
Table 4-7 presents the RRs for each measure group. Since measure type was not a unique sampling stratum, the measure level results are only representative of those projects evaluated. These RRs should not be extrapolated out to the Program level. Only RRs based on sampled stratum can be used to make statistically valid statements about the Program (see Table 4-6 above).

Table 4-7. Realization Rates by Measure Group – Natural Gas (July 1, 2010 through June 30, 2012)

Measure Group	Realization Rate
Boiler improvements	0.92
Capacity expansion	1.32
Heat recovery	1.00
HVAC	0.80
Misc. Industrial Process Efficiency	0.92

Sampling did not consider whether the project was located upstate or downstate; the weighted savings of the sampled projects by their geography[17] is presented in Figure 4-10.

Figure 4-10. Natural Gas Savings Upstate and Downstate (July 1, 2010 through June 30, 2012)



[17] The Impact Evaluation Team classified all projects served by Con Edison as downstate projects. All other projects were classified as upstate projects.

4.2.4 Influence of Program Measurement and Verification Activities on Natural Gas Savings Estimates

The Program requires that any projects with savings greater than 10,000 MMBtu in natural gas savings receive Program M&V. M&V is performed by the applicant and a third-party Technical Reviewer assigned to the project.

The RRs for projects with and without completed Program M&V are presented in Table 4-8.

Table 4-8. Program M&V Results for Natural Gas

Program M&V	n	Realization Rate
Conducted	10	0.97
Not-conducted	3	0.93

The RRs for projects with completed Program M&V are similar to the RRs for those projects that did not receive formal Program M&V. However, the savings for three of the four projects that did not receive formal Program M&V are based on measured and trended data that had been collected as part of previous project reviews or by the applicant. While these projects did not receive formal Program M&V, the key variables used in estimating savings were in fact measured.

4.3 CONCURRENT EVALUATION REVIEW RESULTS

The Impact Evaluation Team implemented a concurrent review process in parallel with the retrospective impact evaluation discussed in this report. Although the retrospective evaluation has concluded, the concurrent evaluation review process is ongoing with a proposed close-out of December 2015. Through the concurrent review process, evaluators provide feedback to NYSERDA Program Staff and their contractors regarding project baselines, measurement and verification strategies, and analysis methodologies and results. The IPE Program Staff are not mandated to comply with the evaluator recommendations, although thoughtful discussion and resolution of differences is a core component of this review process and often results in the Program and evaluation perspectives converging on a mutually agreed upon solution. The evaluation feedback provided through this process is documented with NYSERDA. However, should a project that undergoes concurrent review be pulled in a future retrospective evaluation, there is no guarantee that the concurrent evaluation results will be adopted; changes in site operations, differences in evaluation contractor perspectives, or other factors may result in adjustments or exclusion of the concurrent review results.

This concurrent evaluation review process was performed on 36 projects from September 2013 through December 2014. Table 4-9 breaks down the 36 concurrent review projects by their facility type.

Table 4-9. Summary of Concurrent Review Project Count

Facility Type	Number of Projects
Data center	8
Industrial	28
Total	**36**

In aggregate, the thirty-six projects included in the concurrent review process account for more than 127,000 MWh, 6.1 MW, and 319,000 MMBtu/yr in projected savings. These projects represents a large amount of savings; by comparison this equates to more than 60% and 90% of the Program-reported kWh and MMBtu savings evaluated in the retrospective evaluation of projects completed between July 1, 2010, and June 30, 2012 (see Table 2-1, above, for the retrospective evaluation sample frame).

As of December 2014, preliminary reviews were complete on 33 of 36 projects. Table 4-10 summarizes the status of the 36 projects in the concurrent review process.

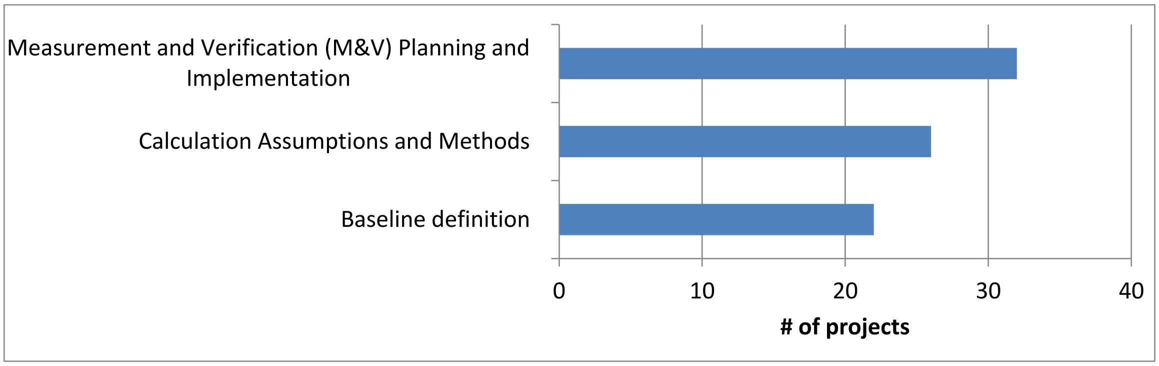
Table 4-10. Summary of Concurrent Review Project Status

Status of Concurrent Review	Number of Projects
Final post-installation feedback provided to Program	10
Preliminary feedback provided to Program	23
No feedback provided to Program	3
Total	**36**

Of the 10 projects where the post-installation concurrent review feedback has been finalized, the Impact Evaluation Team calculated the evaluated energy savings to be within 10% of the program-reported savings for eight of them. These savings were arrived at after extensive review and vetting of questions and concerns between the Impact Evaluation Team and Program staff on each individual project. For the two projects where the final evaluated energy savings were more than 10% different than the program-reported savings, the reasons for the differences were either errors noted in the final M&V calculations submitted by the Program or fundamental differences in the calculation methods applied by the Impact Evaluation Team and the Program.

For the 23 projects with completed preliminary reviews, as well as the 10 finalized post-installation reviews, the Impact Evaluation Team noted several recurring issues. Figure 4-11 summarizes these issues and indicates the number of projects in which each issue was noted; some projects had multiple issues raised during the concurrent review process. The significance of these issues and the steps taken to address them are discussed in the paragraphs that follow.

Figure 4-11. Most Frequently Occurring Concurrent Review Issues



M&V planning and implementation was the issue most often occurring during the concurrent review process, followed by calculation methods and assumptions and baseline definition. These three issues include a variety of subtopics and are discussed in greater detail below.

- **M&V Planning and Implementation** – The Impact Evaluation Team made recommendations to revise M&V plans on just under half of the projects included in the concurrent review process. This included approximately 75% of the data center projects and 80% of the industrial projects. These recommendations included collecting additional data to triangulate whole-facility analysis results and providing better resolution on the M&V sampling, metering duration and data collection strategies. On most projects, the Impact Evaluation Team's recommendations were implemented by the Program; however, there were several instances where the Program staff either did not collect all of the recommended data or where it was impractical for the measurements to be taken given on-site limitations. For such projects, the Impact Evaluation Team either solicited additional data from the site, or, where this was not feasible, included recommendations for such data collection in future evaluations. In general, the Impact Evaluation Team found that coordinating M&V activities with the Program staff increased the likelihood that the requested data would be collected from the site, ensured the proposed M&V was of a high level of rigor, and provided an opportunity for sharing M&V strategies with the Program's contractors to improve future project data collection and results.
- **Calculation Assumptions and Methods** – Most of the projects where the Impact Evaluation Team found issues with calculation methods or assumptions were capacity expansion projects that required regression analysis or natural gas projects where whole-facility energy usage was used to calculate annual energy savings. This issue arose in approximately 65% of both the data center and industrial projects

In recognition of this issue, and after discussion with both NYSERDA evaluation and program implementation staff, a rigorous review of the Program's M&V results was incorporated into the Impact Evaluation Team's review scope. The Impact Evaluation Team found that incorporating a detailed evaluation review of the Program's analysis and results provided the opportunity to discuss any questions or concerns regarding analysis methods and increased the likelihood that the Program and Impact Evaluation Team's results converged on similar solutions.

- **Baseline definition –** The most often discussed baseline issues surrounded data center process efficiency projects and industrial new construction and capacity expansion projects. Baseline definition concerns were raised by the Impact Evaluation Team on more than 85% of the data center projects and 50% of the industrial projects. Early in the concurrent review process, the Impact Evaluation Team provided recommendations for baseline documentation and research but did not perform research to validate the project baseline. In reviewing early concurrent review projects, it was found that additional research was warranted to both inform project-specific baselines and to better inform baseline definition for projects moving forward. On more recent concurrent review projects, the Impact Evaluation Team worked closely with the Program staff to gather site-specific and third-party data to document project baselines. This coordinated effort produced more defensible and better documented project-specific baselines and eliminated the room for interpretation that existed in early concurrent project reviews. Appendix C provides an overview of the process used to characterize project baselines.

Challenges to the concurrent review process include long project lead times, the need for timely and effective communication among all involved parties, and the demand for open information exchange at key project milestones. The Impact Evaluation Team observed that the concurrent review process was most successful on projects with technical reviewers who understood the process and its intent and evaluation review engineers who were able to provide timely and useful feedback. The concurrent review process enabled this communication and information exchange by implementing email correspondence and feedback tracking during the early stages of a project's concurrent review, with formal report deliverables after the final pre- and post-project installation feedback was provided by the Impact Evaluation Team. The projects where the process was most effectively implemented also had the same technical reviewer or evaluation engineer from review inception to completion. This continuity enabled smoother communication and more effective reviews because topics were not being revisited as staff turned over on the project. The need to continue to involve Impact Evaluation Team early in project development is paramount as it helps to mitigate lost opportunities. On at least one large project the opportunity to perform pre-installation metering was missed because the Impact Evaluation Team was not involved early enough in the process.

4.4 RECOMENDATIONS

The Impact Evaluation Team offers four recommendations based on the impact evaluation research.

1) Expand the concurrent review process to include at least a sample of data center cooling and IT projects even if they do not meet the standard size threshold for concurrent evaluation eligibility.

   As discussed in Section 4.1.4, data centers present challenges to retrospective M&V evaluation. Evaluation fatigue and concerns with the interruption of critical systems at data centers means that the Program generally has only one chance to meter data center projects post-installation. Most sites in the sample did not want the Impact Evaluation Team to install equipment for a second round of metering. Additionally, the evaluators rely on project narratives and site-supplied data to verify baseline conditions. However, in many cases, long before the post installation site visit, baseline equipment has already been removed, thus limiting the Impact Evaluation Team's ability to modify the applicant baseline.

   The Program staff should expand the concurrent review process through 2015 (or until existing funds are exhausted) to include a selection of data center projects. This will allow the Program and the Impact Evaluation Team early access to data center projects, providing the opportunity to characterize the baseline, and to ensure that Program M&V plans will capture the key variables for a representative time period.

2) As noted, the RR for data center cooling projects was 0.43. The Impact Evaluation Team generally found that a lower level of rigor was applied to the data center cooling projects, and baseline characterization often differed significantly from the Impact Evaluation Team's baseline characterization.

   A baseline determination flow chart, similar to the baseline characterization flowchart contained in Appendix C, should be developed for data center cooling project baseline determination. Research conducted as part of this evaluation has generated a body of knowledge on which to develop this flowchart, which could be added to the existing baseline characterization flowchart.

3) Conduct data center research. Conduct research into data center operations, equipment obsolescence, IT equipment efficiency improvements, and baseline identification.

   In addition or in lieu of a specific research effort, use the concurrent review mechanism to identify IT projects outside of and before the traditional retrospective evaluation cycle and enable evaluation staff members to visit sites while it is still possible to evaluate the equipment performance, or introduce a rolling sample approach to the next impact evaluation design to enable early access.

   This recommendation is based on the finding that in one-third of the evaluated data center IT projects, the as-built equipment had already been replaced with newer equipment at the time of the evaluation site visit. This is a reflection of the speed of obsolescence in the data center industry. Nearly half of the data center sites state that they replace their IT equipment every three years due to market forces. Certain facilities and customers replace equipment as frequently as every 18 months. This rapid market evolution presents a particular challenge in identifying both baseline and high efficiency equipment or performance levels.

4) Consider leveraging existing relationships and knowledge of the data center market.

   NYSERDA should consider leveraging existing relationships and knowledge of the data center market to identify and promote efficiency best practices in a more market-animating or transformational basis that convinces actors in this fast-paced market to choose higher efficiency options as a matter of course even though the required payback for investing in efficient measures must be very short given that the EULs are in the range of 1.5 to 3 years. The need for direct project-by-project intervention by the Program could be reduced. This approach is in line with proposed CEF strategies. As an example, the Impact Evaluation Team was able to demonstrate to a data center how they could use their existing systems and software to immediately begin to display and trend energy use of their systems. They were already in possession of the software, and were unaware of its capabilities and functionality. NYSERDA can play a role in helping data center facilities leverage existing resources while providing education on best practices and energy savings opportunities.

APPENDIX A: GLOSSARY OF TERMS[18]

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

error ratio – In energy efficiency evaluation, the error ratio is a measure of the degree of variance between the program-reported savings estimates and the evaluated estimates. For a sample, the error ratio is:

$$er = \frac{\sqrt{\sum_{i=1}^{n} w_i \frac{e_i^2}{x_i^{\gamma}} \sum_{i=1}^{n} w_i x_i^{\gamma}}}{\sum_{i=1}^{n} w_i y_i}$$

where,

n is the sample size

w_i is the population expansion weight associated with each sample point i

x_i is the program-program-reported savings for each sample point i

y_i is the evaluated savings for each sample point i, the constant gamma, $\gamma = 0.8$ (typically), and the error for each sample point $e_i = y_i - bx_i$, where b is the program realization rate

evaluated 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 Impact Evaluation Team.

IPMVP Option A – This M&V option involves the partial measurement of isolated equipment affected by the evaluated measure. Relevant equipment variables are spot-measured when possible or stipulated when necessary.

IPMVP Option B – This M&V option involves full measurement of the isolated equipment affected by the evaluated measure. No stipulations are allowed. Both short-term and continuous data monitoring are included under Option B.

IPMVP Option C – This M&V option involves the use of utility meters to assess the performance of a total building. Option C addresses measure impacts in aggregate, not individually, if the affected equipment is connected to the same meter.

IPMVP Option D – This M&V option involves the use of computer modeling to determine facility or equipment energy use. Option D requires calibration with actual utility consumption data for either the pre-project or post-project period.

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, free riders, energy efficiency standards, changes in the level of energy service, and other causes of changes in energy consumption or demand.

net to gross, net-to-gross ratio (NTG, NTGR) – The relationship between net energy and/or demand savings – where net is measured as what would have occurred without the program, what would have occurred naturally – and gross savings (often evaluated gross savings). The

[18] NYSERDA generally follows and uses the terms as defined in the "Northeast Energy Efficiency Partnerships Glossary of Terms," found at http://neep.org/uploads/EMV%20Forum/EMV%20Products/EMV_Glossary_Terms_Acronyms.pdf. This glossary defines those terms absent from the NEEP report or provides more-specific definitions to generalized NEEP terms.

NTGR is a factor represented as the ratio of net savings actually attributable to the program divided by program gross savings. For NYSERDA programs the NTGR is defined as 1 minus free ridership plus spillover.

.

population expansion weight – The total number of units in a population divided by the number of units in the sample.

realization rate (RR) – The ratio of the evaluated gross savings to the Program's program-reported savings. The RR represents the percentage of program-estimated savings that the Impact Evaluation Team estimates as being actually achieved based on the results of the evaluation M&V analysis. The RR calculation for electric energy for a sampled project is shown below:

$$RR = \frac{kWh_{evaluation}}{kWh_{program}}$$

where,

RR is the realization rate

$kWh_{evaluation}$ is the evaluation M&V kWh savings (by evaluation M&V contractor)

$kWh_{program}$ is the kWh savings claimed by program

relative precision – 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 deviation 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.

Relative precision is calculated as shown below. It must be expressed for a specified confidence level. The relative precision (*rp*) of an estimate at 90% confidence is given below:

$$rp = 1.645\ \frac{sd(\mu)}{\mu}$$

where,

μ is the mean of the variable of interest

sd(μ) is the standard deviation of μ

1.645 is the *z* critical value for the 90% confidence interval

For the 90% confidence interval, the error bound is set at 1.645 standard deviations from the mean. The magnitude of the *z* critical value varies depending on the level of confidence required.

stratified ratio estimator (SRE) – An efficient sampling design combining stratified sample design with a ratio estimator. It is most advantageous when the population has a large coefficient of variation, which occurs, for example, when a substantial portion of the projects have small savings, and a small number of projects have very large savings. The ratio estimator uses supporting information for each unit of the population when this information is highly correlated with the desired estimate to be derived from the evaluation, such as the program-reported savings and the evaluated gross savings.

summer coincident peak demand period – For this evaluation NYSERDA defined the summer coincident peak demand period as the energy reduction during the hottest non-holiday summer (June through August) weekday during the hour ending at 5 p.m.

APPENDIX B: MEMORANDUM – ACCOUNTING FOR PROGRAM ADJUSTMENTS TO REPORTED PROJECT SAVINGS AFTER DELIVERY OF DATA SET OF RECORD DELIVERY TO EVALUATION TEAM

710 Park Place
College Station, TX 77840
Phone: (979) 696-5465
Web: www.ers-inc.com

MEMORANDUM

DATE: June 18, 2014

TO: Judeen Byrne, Carley Murray, Jennifer Meissner, Dakers Gowans

FROM: Jon Maxwell

CC: Sue Haselhorst, Kathryn Parlin, Nick Collins, Betsy Ricker, Yogesh Patil

RE: ACCOUNTING FOR PROGRAM ADJUSTMENTS TO REPORTED PROJECT SAVINGS AFTER DELIVERY OF DATA SET OF RECORD DELIVERY TO EVALUATORS

This memorandum proposes a method to account for changes that NYSERDA's program administrators make to reported savings estimates on projects after they have provided a data set of record to evaluators for sampling. It applies most immediately to IPE but the approach is general and will work for other programs.

If NYSERDA evaluation staff finds it reasonable we can circulate it to Program staff and the DPS consultants.

Accounting For Program Adjustments to Reported Project Savings After Delivery of the Data Set of Record to Evaluators

Background

NYSERDA uses a cumulative approach to reporting program energy savings. For example, the savings reported for 2013 in the spring of 2014 is the sum of all savings for all programs reported through 2013 less the amount reported a year earlier. The cumulative approach enables program administrators to adjust reported values from earlier years such as 2012. The intent is that the reported savings represent NYSERDA's most recent and best estimate of all savings to date.

This approach to tracking and reporting savings is particularly relevant for programs such as Industrial Process Efficiency and Existing Facilities. NYSERDA program administrators require that large projects in these programs receive up to a year of post-installation measurement & verification (M&V). The programs initially report completed project savings based on the estimate available at the time of installation. After M&V, the programs re-state project savings based on the M&V results if necessary.

The Problem

The cumulative approach causes challenges for ex post evaluation sampling. Evaluators request a data set of record from which they will build a frame and draw a sample. For the evaluation to be timely and relevant, this request is usually for projects completed over a relatively recent period, often for a time range that ends less than a year prior to the time of request. For example, an evaluation data request issued in April 2014 might request tracking data on all projects completed between January 1, 2012 and December 31, 2013.

The problem is that the savings for some subset of projects in the evaluator's population frame is not final. The Program will eventually change savings estimates due to ongoing M&V. The evaluator cannot wait for the updated estimate to execute the sample design, and cannot change the estimate in the middle of the analysis period because the sample design, stratification, weighting, etc. all is based on the savings in the data set of record. Excluding projects pending M&V revision would bias the results. Creating a separate census stratum for them is not practical. But if evaluators ignore the Program's subsequent program-required M&V changes, the Program will be subject to double correction of savings. Applying the evaluation's realization rate to the continuously improved program savings estimate will give the wrong result.

The Proposed Solution

This proposed approach recalculates the realization rate and program reported savings to account for the presence of sites where savings adjustments will occur.

1. The impact evaluation team will perform all sampling, evaluation M&V, and related work in the same fashion as if there was no ongoing program-driven M&V. The result will be an original program realization rate.

$$RR_{original} = \frac{Evaluated}{Reported_{original}}$$

2. After all site-specific evaluation M&V is complete, the impact evaluation team will ask the Program to provide a new data set of record for all projects and measures in the original frame subjected to post-installation M&V. The new data set will have program-reported savings that for some projects will have changed compared to the original data set as a result of program M&V.
   Program staff must ensure that the new data set does not include savings for new measures added to old projects or other reasons than post-installation M&V adjustment.
3. The adjusted realization rate is calculated by dividing the evaluated savings by the revised and adjusted reported savings.

$$RR_{adjusted} = \frac{Evaluated}{Reported_{adjusted}}$$

4. The site-specific M&V reports will present both the original and adjusted program reported savings and corresponding realization rates.

Attachment 1 provides an example calculation. With this technique, there are several advantages:

- The evaluated net energy savings remains the same; only the realization rate and program reported savings are adjusted.
- It will accurately present the program performance using the most contemporary data available
- It will not contaminate the original sample design.

Attachment 1: Example

Column No.:	A	B	C	D	E	F	G	H	I
							= (F / B)		= (F / H)
Row No.	Population ID	Dec 31, 2013 Program Reported Savings	Stratum	In Sample?	Expansion Weight	May - Jul 2014 Evaluated Savings	Original Evaluation Realization Rate	Aug 31, 2014 Program Reported Savings (KWh)	Adjusted Realization Rate Also Presented in Sep 2014 Site Report
1	1	100,000	1	n				100,000	
2	2	100,000	1	y	2	100,000	1.000	100,000	
3	3	200,000	2	n				200,000	
4	4	200,000	2	y	2	200,000	1.000	200,000	
5	5	300,000	3	n				300,000	
6	6	300,000	3	y	2	300,000	1.000	300,000	
7	7	400,000	4	n				400,000	
8	8	400,000	4	y	2	400,000	1.000	400,000	
9	9	500,000	5	n				500,000	
10	10	500,000	5	y	2	475,000	0.950	*475,000*	1.000
11	11	600,000	6	n				*570,000*	
12	12	600,000	6	y	2	500,000	0.833	600,000	
13	**Total**	**4,200,000**					**0.940**	**4,145,000**	

14	Original program reported gross savings	= B13	4,200,000	
15	Original program realization rate	= G13	0.940	
16	Original program evaluated gross savings	= B13 * G13	3,950,000	
17	Adjusted program evaluated gross savings	= B13 * G13	3,950,000	*same as G16*
18	Adjusted program reported gross savings	= H13	4,145,000	
19	Adjusted program realization rate	= G17 / G18	0.953	

APPENDIX C: NYSERDA BASELINE DETERMINATION METHODOLOGY FOR PROGRAM EVALUATION

NYSERDA Industrial Process Efficiency Baseline Determination Methodology for Program Evaluation

--- Final 2/12/2010 ---

---

by

Jon Maxwell, ERS

and Megdal & Associates Impact Evaluation Team

prepared for

Cherie Gregoire, NYSERDA Energy Analysis

Section 1:

OBJECTIVE

NYSERDA's Industrial Process Efficiency (IPE) program helps New York manufacturers and data centers improve their operations to reduce energy use per unit production. Most studies and projects funded through this program will face the issue of defining baseline energy use and efficiency for their project. Often the project will be unique, there will be no legislated code or standard that specifies minimum efficiency, and there will no known prior research on common practice.

The objective of this document is to establish a standard procedure by which evaluators will define baseline energy use for manufacturing and data center projects, and to do so early in the program's implementation so that Program staff can use the same procedure or at least be aware of the standard against which their projects will be evaluated when calculating energy savings for incentive calculation and reporting purposes.

Baseline is a complex idea that can be hard to separate from free ridership, measure life, and other factors. This report describes the scope and limitations of the NYSERDA baseline determination process in Section 2. In Section 3 we define terms later used in the core Section 4, the baseline determination process itself. Section 4 is largely a logic flowchart with accompanying explanatory text.

Section 2:

APPROACH AND SCOPE

The intent of this document is to offer a procedure for defining baseline. It does not attempt to specify the actual baseline values, applicable codes or standards, or appropriate technology baselines for any particular type of project.

The underlying principle for defining baseline is straightforward: *The baseline is the least efficient option specific to a particular facility and application that the customer technically and economically would have reasonably considered to deliver the post-retrofit level of production.* Application of this concept can be complicated. This memorandum provides guidance in interpreting this statement for a variety of circumstances.

Because the concept of baseline can be hard to separate from related issues, the approach scope of the definition process needs constraints. For NYSERDA we use the following scope guidelines:

Free Ridership. In many cases what could or would have been done in the absence of the NYSERDA program is not readily apparent. Baseline definition will help evaluators determine the least efficient approach that reasonably *could* have been taken. Free ridership research will determine the difference between what *would* have happened versus what *could* have happened. To the extent that any of this interpretation is discretionary, evaluators will assess the difference as part of free ridership rather than elevating the baseline. For example, a customer could, as a matter of corporate policy, always practice a certain higher level of efficiency than some competitors. This reflects free ridership, not a high baseline.

Measure Savings Over Time. The IPE evaluation will follow NYSERDA precedent and will not evaluate measures according to "dual" baselines; that is, the baseline and post-retrofit energy use and measure savings will be estimated as being constant for all years of the measure life. In the event that a measure is removed or otherwise fails early, the life for that particular project will be reduced accordingly and remaining year savings will be zero.

Some Measures Increase Production Levels. The baseline for measures that increase production must account for alternative actions that could have been taken to otherwise increase production. Existing conditions and production methods in place should be considered when defining baseline for the added production but are not always the basis of it.

Who Defines Baseline? While the evaluators and implementer may work together for baseline determination on large complex processes, the burden for researching and defending the baseline rests with the implementation team.

Definitions

Post-Retrofit Production Level: Post-retrofit production level is the evaluator's judgment of long term expected production rates for the facility after the retrofit is completed. This level often is the production level measured in the year or two after installation, but it could be adjusted from this value based on pre-retrofit data or on forecasts of future production levels. It does not necessarily equate to production levels immediately before or after the retrofit.

Baseline System Configuration: Baseline system configuration is the least efficient system that the customer realistically could have used for the specific application to deliver the post-retrofit production level.

Baseline Energy Use: Baseline energy use is the amount of energy needed to meet the post-retrofit production level with the baseline system configuration.

Application: Application is defined by the combination of (a) the customer's industry type, (b) the particular process affected by the project, (c) the equipment itself, and (d) site. For example, adding insulation on refrigerant lines in the wine industry is a different application than adding insulation to a wine storage tank at the same site. It also is a different application than adding insulation to refrigerant lines at a refrigerated storage warehouse. NYSERDA intends to establish baseline efficiency on a site/process specific basis to the extent possible.

Minimum Commonly Used Efficiency

Minimum commonly used efficiency is the minimum efficiency that a reasonable person would choose to install for a particular application. It should be used for baseline determination.

It is never worse than minimum available efficiency but can be better, if there are minimum efficiency solutions that theoretically are possible but as a practical matter a reasonable person would not use for the particular application. For example, there may be circumstances where the least efficient technology is not the least expensive option for a customer. In such cases, it usually is unreasonable to assume that the higher cost approach is the baseline, even if it is less efficient.[1]

It is never better than standard practice but can be worse, if there are a measurable number of market actors that install less than the predominant/standard practice level of efficiency.

There should be evidence that it is an approach currently used in industry for the type of application under consideration.

---

[1] As an example: A customer has an older 70% efficient boiler that needs to be rebuilt/re-tubed or replaced. If rebuilt, the resulting efficiency will be 75%. A new minimum efficiency new boiler is 80%. The customer receives NYSERDA funding to buy a new 86% efficiency boiler. It turns out that a rebuild would cost more than replacement. In that case the baseline efficiency is an 80% efficient new boiler, not the 75% rebuilt one.

Individual customer policies and purchasing practices should be considered. Regional practices may be applicable as well.

Production Energy Use Intensity (EUI)

Energy use per unit production

Baseline EUI = Baseline Energy Use / Post-Retrofit Production

Post-Retrofit EUI = Post-Retrofit Energy Use / Post-Retrofit Production

Energy Impact

The difference in energy use between what would have been used by the baseline system configuration and the installed system at the post-retrofit production level.

Annual Energy Impact	=	Post-Retrofit Production Level	x	(	Baseline EUI	-	Post-Retrofit EUI	)
(Energy/Yr)		*(Units/Yr)*			*(Energy/Unit)*		*(Energy/Unit)*	

Section 3:

HOW TO DETERMINE BASELINE EFFICIENCY

The objective of the procedure is to guide the evaluator and program staff in determining the baseline system configuration, that is, the least efficient system that the customer realistically could have used for the specific application to deliver the baseline production level. The underlying premise is that baseline determination is unique to each project, and that decision-making is complex and customized for each.

This section explains the process of how to determine system configuration and in turn baseline efficiency and production EUI, with special consideration to measures that increase production levels.

The logic flow chart in Figure 4-1 together with the definitions in Section 3 guides the decision-making process.

Retrofit. The first consideration is the same as with traditional commercial and non-process industrial projects: Is the project a retrofit, new construction, failure replacement, process expansion, or major process change? To be a retrofit, the project must satisfy the following criteria:

1. It replaces old working equipment;
2. That old equipment was not otherwise going to be replaced in the near future;
3. It does not increase overall plant production

If these two conditions are met, the baseline is defined by the pre-retrofit conditions. Evaluators may require that the application demonstrate the applicability of the criteria. Interviews with site personnel expressing intent are necessary but not necessarily sufficient evidence. For example, if the project claims that the equipment was working and not going to be replaced in the near future and yet was very old for its type, the project may need to demonstrate why it was reasonable to assume that the equipment was expected to have a long future life and not be replaced in the near future in spite of its age. Possible ways to do this are to copy maintenance records showing that a 25-year old boiler, normally considered near the end of its life, had had its tubes replaced in the last five years and recent boiler efficiency test data that showed good combustion efficiency.

If equipment is in good operating condition but external circumstances drive its replacement, new emissions regulations, for example, the replacement will be considered new construction unless the applicant demonstrates that there was a reasonable way to comply with the new requirements through modification rather than replacement. If so, the modification alternative will be the theoretical baseline.

New Construction and Failure Replacement. If the project does not meet either criteria #1 or #2 above, then it is either new construction or change out replacement due to equipment failure. The two are treated the same in terms of baseline definition. The next step is to determine new construction/failure replacement baseline. The different bases are as follows, in order of priority:

1. Applicable code or standard
2. Minimum commonly used efficiency for the application
3. Custom-developed baseline
4. Existing Process (on-site or at other applicant owned site)

Some codes and standards apply to industrial process projects. The Energy Policy Act (EPAct) specifies selected minimum motor and refrigeration equipment efficiency standards. ASHRAE and the NY ECCC for buildings may affect insulation practices for a facility. OSHA ventilation standards may apply. An applicable code or standard is the preferred basis for defining baseline.

In the event that no code applies and the system being installed is not unique to the customer, then "minimum commonly used" efficiency standard is the preferred choice. Many refrigeration and compressed air systems, crushing, conveying, process heating, pumping, and other industrial processes are common to many facilities. If the technical reviewer can document a minimum commonly used standard for the application, that standard should be used.

The minimum commonly used basis is application-specific. For example, it may be that the market offers a low cfm/kW and throttle-modulation air compressor of the right size, but if it is not an option commonly or realistically used for the particular application, it is not the baseline.

If the minimum commonly used basis is inappropriate or impossible to determine, the next alternative is a custom-developed application-specific baseline. If the data or research already are available, or if the project itself conducts such research, a project-specific baseline may be used.

The last and default option for new construction projects is to use the existing process as the baseline. In

all cases program staff are responsible for requiring applicants to demonstrate the appropriateness of the baseline. Evaluators may verify it.

Increased Production and Major Process Change Measures. Projects that increase production and projects that change the fundamental way in which the product is made are treated the same for the purposes of baseline definition,. There are two additional key factors to consider for such projects: Energy use per unit production, and alternative approaches to meet production.

Energy use per unit production, or production energy use intensity (EUI), is the basis for measuring energy efficiency improvements in such projects. For increased production and major process change measures the applicant must compute the baseline and post-retrofit EUIs and multiply them by the annual baseline production level (as determined per the definitions) to estimate normalized energy use before and after the retrofit.

Alternative approaches to increase production may be necessary to assess instead of just pre-retrofit conditions to define baseline. The guiding principle in determining baseline EUI for productivity increasing projects is that it should be based on what the applicant otherwise could have done to increase production without the program-funded action(s). If the applicant could have increased production using existing methods, such as by increasing operating hours, by increasing the processing season, or by activating other similar equipment as already was in place, and it would not have fundamentally changed the process EUI, then pre-retrofit EUI can be the baseline EUI.

If on the other hand the plant's equipment was at capacity then the project represents a market opportunity. Baseline definition must consider how else the plant, the larger corporation, or the industry as a whole would otherwise have met production needs absent the funded project. If none of those options can be determined, the default is pre-retrofit EUI. The flowchart guides the decision-making process.

Figure 4-1 Baseline EUI Logic Flow Chart

HOW TO DETERMINE BASELINE ENERGY USE INTENSITY (ENERGY USE PER UNIT PRODUCTION) AND EFFICIENCY FOR PRODUCTION-RELATED MEASURES

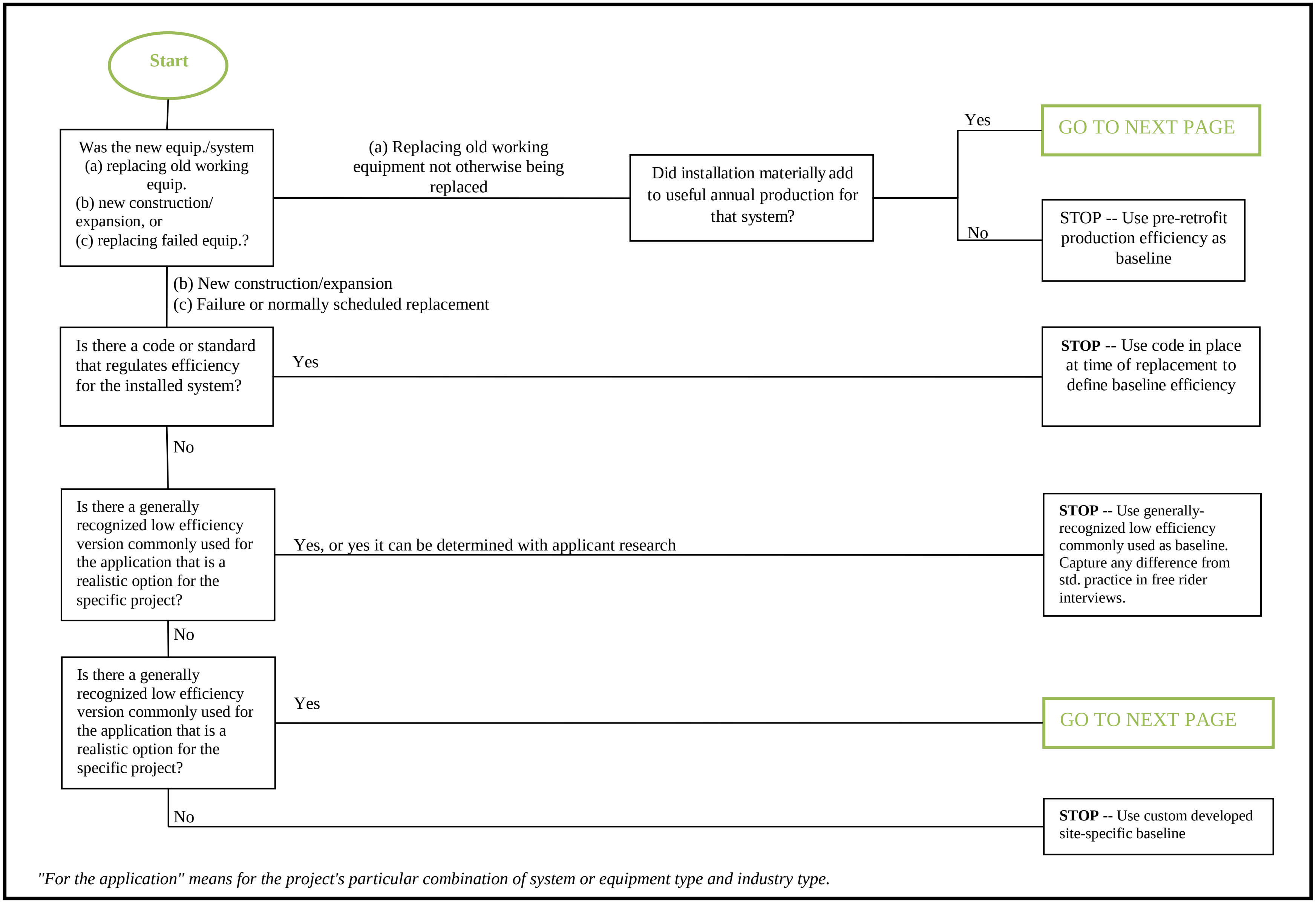
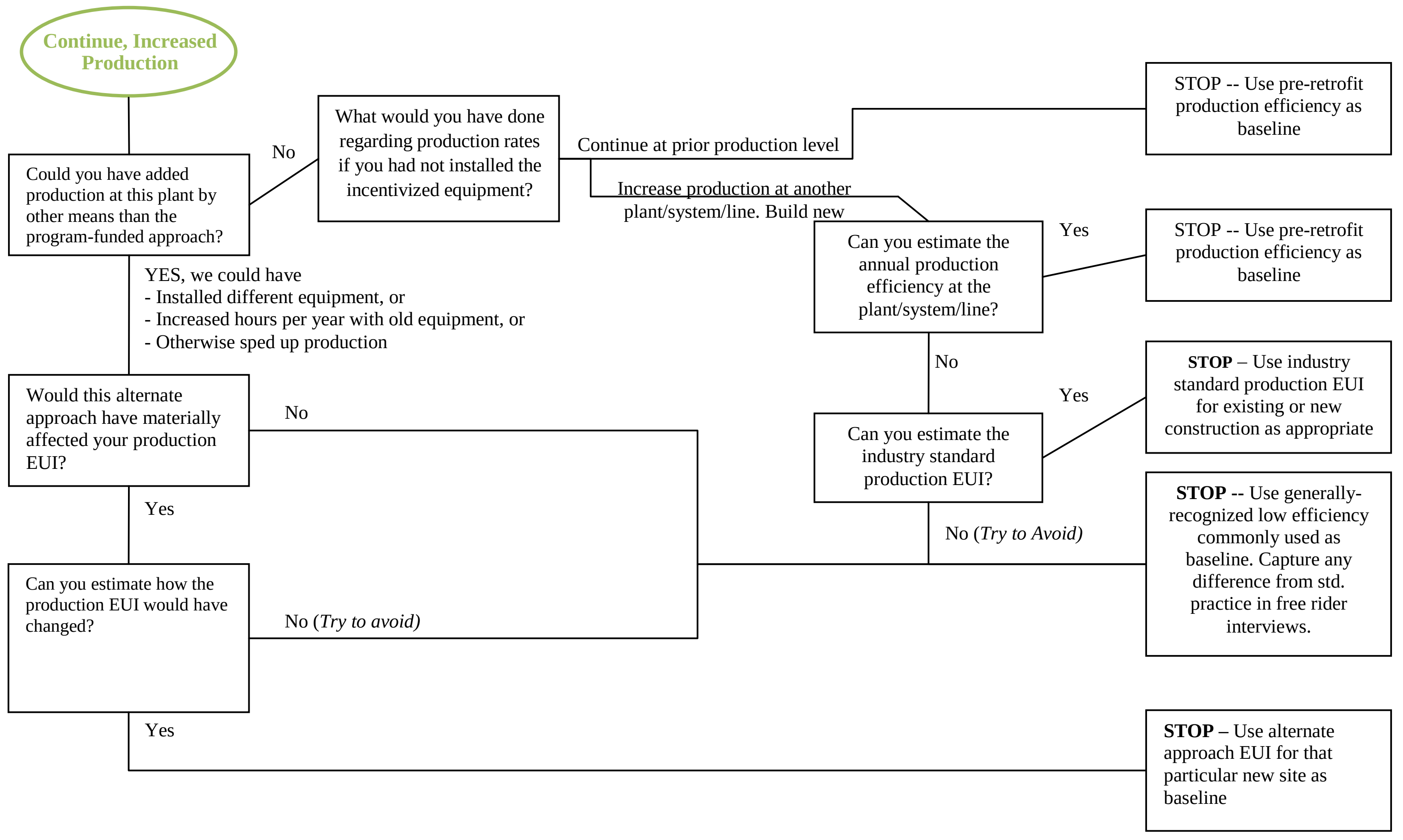


Figure 4-1 Baseline EUI Logic Flow Chart (continued)

HOW TO DETERMINE BASELINE ENERGY USE INTENSITY (ENERGY USE PER UNIT PRODUCTION) AND EFFICIENCY FOR PRODUCTION-RELATED MEASURES



"For the application" means for the project's particular combination of system or equipment type and industry type.

APPENDIX D: CONCURRENT REVIEW PROJECT SELECTION PROCEDURE AND OVERVIEW OF CONCURRENT REVIEW PROCESS

710 Park Place
College Station, TX 77840
Phone: (979) 696-5465
Fax: (978) 478-5330
Web: www.ers-inc.com

M E M O

DATE: Original submission on July 29, 2010, revised May 23, 2013

TO: Carley Murray and Judeen Byrne, NYSERDA

FROM: Betsy Ricker and Jon Maxwell, ERS

RE: PRE-RETROFIT REVIEW PROJECT SELECTION PROCEDURE AND OVERVIEW OF PRE-RETROFIT REVIEW PROCESS

The impact evaluation team has committed to performing both pre-retrofit and post-retrofit data collection for larger projects in accordance with the Industrial and Process Efficiency (IPE) Evaluation, Measurement, and Verification Plan (a.k.a the Work Plan).

The Pre-retrofit Review and Project Selection process was implemented in July 2010. This memorandum summarizes the existing process and offers enhancements to assist both program staff and the evaluation contractors in the delivery of program services and increased rigor in the impact evaluation. This pre-retrofit M&V will require the continued cooperation of IPE program staff and their contractors.

Notify the Evaluation Team About Which Projects?

The evaluation team works with program staff and collects pre-retrofit data on IPE projects. Evaluators recommend the following criteria for identifying candidate projects for pre-retrofit review:

1. Over 5,000,000 kWh/yr or 40,000 MMBtu/yr expected savings
2. Over 1,000,000 kWh/yr or 10,000 MMBtu/yr expected savings and
    a. Involve process changes, or
    b. Baseline definition requires site-specific characterization, or
    c. Program staff has concerns with baseline definitions
    d. Measure(s) includes controls upgrades (i.e. lighting controls, motor controls, VFD installation, VSD compressor installation, HVAC for process loads, etc.)

Evaluators will not become involved in every project that meets the above criteria but request the opportunity to consider all projects that meet the above criteria for inclusion in the pre-retrofit evaluation.

When Should the Process Start?

The program will benefit most if the evaluation team's involvement occurs at the time a customer is ready to develop "investment grade" energy savings estimates and prior to the program's commitment of funding based on those estimates. This will ensure that NYSERDA funding aligns with agreed-upon savings estimates and, if past history is an indicator, will increase program cost-effectiveness.

In some cases this early level of involvement may not be possible. Evaluators should be notified no later than the time an applicant commits to a project and NYSERDA encumbers funding. At a minimum the evaluation team needs to have time to request or collect pre-retrofit nameplate data and billing history data and perform short term metering for a period of up to four weeks prior to original equipment removal.

How Will the Notification Process Work? The sections below outline the process by which projects will be selected for pre-retrofit evaluation and how the evaluation will proceed after project identification.

Monthly Meetings and Project Selection

The NYSERDA Program Manager and IPE Impact Evaluation Manager will review the list of candidate projects and identify projects for inclusion in the pre-retrofit evaluation on a monthly basis. This list will be updated monthly by the evaluation contractor, based on information from NYSERDA's Buildings Portal, and will include all newly acquired or encumbered projects meeting the criteria outlined above.

Once a project is selected for inclusion in the pre-retrofit evaluation, the NYSERDA Program Manager or designee sends an email to the IPE Impact Evaluation manager notifying the impact evaluation team of the project and attaching:

- All application materials
- Contact information for key individuals including NYSERDA project lead

The evaluation team will review the initial project materials and confirm that the project should involve evaluation pre-retrofit involvement. If so, the evaluation manager will designate an evaluation lead engineer and will work with the NYSERDA pre-retrofit review process manager to set up a phone call between NYSERDA's designated project lead and the project's technical reviewer.

Pre-Retrofit Review Process

Project Kick-Off Meeting and Pre-Installation Site Visit

Next, the evaluation lead engineer will talk directly with the applicant, their consultants, and NYSERDA's third party technical consultants and/or Outreach contractors as appropriate to gather additional information on the project, further refine the implementation timeline, and request a pre-installation site visit.

The next steps in the evaluation pre-retrofit review process depend on the timeline and status of the encumbered project and may be different for each project. For example, a new construction project may not require a pre-installation site visit while a retrofit project would require such a visit to collect pre-installation data. Although the specific steps required for each project may differ, the key steps expected for a typical concurrent review project are similar and are outlined below.

- **Evaluation Review of Program M&V Plan -** For all IPE projects over 1,000,000 kWh/yr of electric savings or 10,000 MMbtu/yr of natural gas savings, the program requires that applicants perform M&V. Evaluators will review the M&V plan proposed by the applicant and will request adjustments or additional data collection based on their review. Applicants are encouraged, but not required, to adopt evaluator recommendations in their M&V plans. Should the applicant and evaluators not agree on the proposed M&V strategy or data collection requirements, evaluators will perform the metering and M&V in excess of what the program performs. After discussing the selected project with the technical consultant, evaluators will request a copy of the proposed M&V plan for the project (if not already received from NYSERDA Program Staff). Evaluators will review this M&V plan and provide feedback and additional information requests directly to the technical assistance provider with CC to the NYSERDA Program Manager.
- **Pre-Installation Site Visit -** Evaluators will then request a pre-installation site visit to review the existing equipment and its operation and perform any pre-installation metering not included in the technical consultant's M&V activities.
- **Pre-Installation Site Visit Evaluation Summary Memo -** Following the pre-installation site visit, evaluators will draft a formal memo outlining the activities undertaken during the site visit, summarize the evaluation recommendations on the technical reviewer's planned M&V, and detail any additional data collected during the pre-installation site visit. This memo will also summarize the baseline applied in Program savings calculations and, should it differ from the Program, the baseline recommended by evaluators. This memo will be delivered to the Program Manager who will distribute it to the project's technical reviewer.
- **Post-Installation Site Visit** – Post-installation evaluation site visit to verify installed measures and deploy equipment to cover any metering not included in the Program M&V.
- **Evaluator Review of Program M&V Results** – The evaluators will review the Program M&V findings and either approve savings or provide an evaluated savings value for the project.

  Post M&V Evaluation Summary Memo – The evaluation lead engineer will draft a memo summarizing the evaluator findings and indicating the evaluator's savings for the project. This memo will read similarly to a more traditional M&V report with some additional content to address the pre-retrofit review process and highlight specific areas of deviation between Program and evaluator methods and results.

- **Review of Evaluation Summary Memo with Program Staff** – The evaluation summary review memo will be submitted to program staff for review and a conference call will be scheduled as needed to discuss the evaluation findings.
- **Final Evaluation Summary Memo** – The evaluation summary memo will be finalized after conversations with program staff.

Communication

After the project kick-off meeting it will be the responsibility of the technical reviewer to keep the lead evaluation engineer up to speed on any project developments, including, but not limited to: any scheduled site visits, changes in project scope, timeline, or planned
M&V, and any other project developments that might prevent evaluators from gathering the information necessary to complete a comprehensive pre-retrofit project review.

The Evaluation and NYSERDA leads will meet monthly to update project statuses and to verify that the appropriate information is being communicated between the Program and its technical reviewers and the Evaluators.

Transitioning of Projects between Evaluators

Evaluators recognize that large projects (large in savings or capital costs) can be long, perhaps even longer than the current evaluation contract. In instances where the same evaluation lead engineer is unable to take a project from inception to post-installation saving true-up and memo finalization, an interim will be generated that summarizes the evaluation findings and recommendations and the current project status. Documenting this pre-retrofit and project selection process will enable a smoother hand-off to the next evaluation team.