FINAL REPORT

Steam Trap and Boiler Efficiency Research

MA20C05-G-STBE

October 20, 2020

Table of contents

1 EXECUTIVE SUMMARY 4
1.1 Introduction 4
1.2 Study purpose, objectives, and research questions 5
1.3 Conclusions, recommendations, considerations, and future research 7
1.4 Phase II Research 9

2 METHODOLOGY 10
2.1 Working Group 10
2.2 Interviews and best practice research 11

3 EX POST STEAM TRAP MEASURES EVALUATION 12
3.1 Methodology 12
3.2 Steam trap identification 13
3.3 Steam trap sample design 13
3.3.1 Recent evaluation realization rates 13
3.4 Conclusions, recommendations, and considerations 15

4 STEAM TRAP TOOL IMPROVEMENT 17
4.1 Methodology 17
4.1.1 Interviews and best practice research 17
4.1.2 Discrepancy analysis 18
4.2 Findings 18
4.2.1 Interviews and best practice research 18
4.2.2 Site-level discrepancies 19
4.2.3 Hours discrepancies 20
4.2.4 Pressure discrepancies 22
4.2.5 Phase II calibration 23
4.3 Conclusions, recommendations, and considerations 23

5 ANNUAL HEATING PLANT EFFICIENCY FACTOR RESEARCH 25
5.1 Methods 26
5.1.1 AHPE discrepancy analysis 26
5.1.2 Heating plant efficiency – annual variability 27
5.1.3 Annual heating plant efficiency – standby losses 28
5.1.4 Annual heating plant efficiency – alternate methods 29
5.2 Conclusions, recommendations, and considerations 29

APPENDIX A: MINUTES OF MEETINGS 1

APPENDIX B: INTERVIEW GUIDE 23

APPENDIX C: INTERVIEW FINDINGS 28

APPENDIX D: HEATING PLANT MEMO 1

List of figures

Figure 4-1. Applicant vs. evaluated hours of operation 20
Figure 4-2. Applicant vs. evaluated pressure inputs 22

List of tables

Table 1-1. Summary of Steam Trap Evaluated Performance 4
Table 1-2. Study objectives and approaches 6
Table 2-1. Working Group Members 10
Table 2-2. Working Group meetings topics 10
Table 2-3. Interview objectives 11
Table 2-4. Best practices research objectives 11
Table 3-1. Sampling options: realization rates meeting three-year precision targets 12
Table 3-2. Recent steam trap realization rates at the 80% confidence level 13
Table 3-3. Trap/no trap error ratios 14
Table 3-4. PY2018 sample design with calculated precisions @ 80% confidence 15
Table 3-5. Sample design options 16
Table 4-1. PY2017 statewide results – steam traps discrepancy impacts 19
Table 4-2. Explanation of hours discrepancies by site 21
Table 4-3. Explanation of pressure discrepancies by site 22
Table 5-1. Changes in PY2017 RR due to AHPE differences 26
Table 5-2. PY2017 results – heating plant annual efficiency sensitivity analysis 26
Table 5-3. Annual firing rate profile 27
Table 5-4. Steam boilers – efficiency profile 28
Table 5-5. Hot water boilers – efficiency profile 28

List of acronyms used in this report

AHPE	- Annual heating plant efficiency factor
BCR	- Benefit-cost ratio
EEAC	- Energy Efficiency Advisory Council
PA	- Program Administrators
RR	- Realization rate
STT	- Steam Trap Tool

1 EXECUTIVE SUMMARY

This document presents the results of the Steam Trap and Boiler Efficiency Research for the Massachusetts Program Administrators (PAs) under the guidance of the Massachusetts Energy Efficiency Advisory Council (EEAC). This study was led by ERS and included expertise from our partner firm DNV GL. PAs include Berkshire Gas Company (Berkshire), Columbia Gas of Massachusetts (Columbia), Eversource, Liberty Utilities (Liberty), National Grid, and Unitil.

The ERS and DNV GL Team completed four primary activities:

1. Researched factors associated with the evaluation of steam traps measures
2. Interviewed experts nominated by stakeholders to gather feedback on the statewide steam traps calculator (the "Steam Trap Tool," or STT) and annual heating plant efficiency methods
3. Analyzed past evaluation project data to develop findings related to steam trap discrepancies and factors associated with annual heating plant efficiency (AHPE) methods
4. Organized five Working Group sessions attended by experts nominated by stakeholders to discuss the findings of each of the three tasks presented above

Based on the findings of the research, the Team made five recommendations. Details on the study scope, findings, and recommendations are provided in the following sections.

1.1 Introduction

Steam traps constitute a large share of custom program savings (18% and 16% in PY2016 and PY2017, respectively), rely on a single statewide calculator for estimating savings, and have had poorer realization rates (RRs) than other custom measures according to recent custom gas evaluations, as shown in Table 1-1. These factors raise the question of whether steam trap measures should be evaluated as contributing to overall individual PA custom RRs as they are now, evaluated as a statewide measure, or evaluated in some other way.

Table 1-1. Summary of Steam Trap Evaluated Performance

Program Year	Steam Trap % of Custom Savings	Program RR (All Measures)	Steam Trap RR
PY2016	18%	82%	73%
PY2017	16%	87%	70%

The Team vetted and calibrated the statewide calculator using participant billing data in 2016. Users of the Steam Trap Tool (STT) have indicated additional improvements may be in order. While the poor steam trap results may not be directly due to the calculator's functionality, there may be an indirect contribution if users cannot map site conditions to the STT appropriately. Potentially, the STT could benefit from another round of revisions and a recalibration using the population of recent steam trap projects.

Measures such as hot pipe insulation and steam traps reduce the heating load served by a boiler. Converting the thermal load reductions from these measures to natural gas savings in therms requires an annual heating plant efficiency (AHPE) value. The evaluators are using site-specific but limited boiler combustion efficiency readings as a proxy for the overall heating plant annual efficiency, while the PAs have historically used a blended average value (either 80% or 75% in their tracking estimates, accounting for distribution losses) that they believe best represents the annual efficiency. The evaluators and PAs have discussed how this factor should be derived and have reached a consensus on applying a deemed combustion efficiency-based factor to heating load reduction projects going forward.

1.2 Study purpose, objectives, and research questions

This study consists of a number of unique, partially overlapping research objectives related to estimating savings for steam traps and the annual heating plant efficiency factor, which is a primary steam trap savings input (as well as for other measures like pipe insulation). The report has been broken out into three chapters (3 – Sampling, 4 – STT, 5 – AHPE) which feature their own methodology, findings, conclusions and recommendation sections for the purpose of providing a succinct narrative.

The research is divided in two phases, where Phase I leverages readily available information and interviews of stakeholders to resolve identified issues and define a more in-depth research scope for Phase II. This report explicitly details the work associated with the Phase I study and defines a proposed research scope for Phase II. Table 1-2 summarizes the study objectives, conclusions, and their associated proposed Phase II tasks.

Table 1-2. Study objectives and approaches

Study Element Objectives	Phase I Tasks	Phase I Conclusions & Phase II Tasks
Estimating and Evaluating Steam Traps		
Determine the appropriate ex post evaluation treatment of steam trap measures.	- Considered feasibility of alternate custom gas impact evaluation sampling strategies with respect to steam traps.	- The Working Group considered three sample design options and concluded the current design -- of treating steam trap projects as a separate segment within the custom program contributing to individual PA custom RRs -- should be continued. - Phase II will examine the desk reviews from the calibration pool to identify PA trends.
Identify and implement tool revisions that will improve the accuracy of the results, using reasonably available data.	- Analyzed site-by-site discrepancies and PA-to-PA variations in steam trap savings (differences in operating hours, system pressure, traps failed at the time of site visit, and boiler combustion efficiency). - Re-engaged PA engineering staff on issues of tool accuracy. Resolved technical issues. - Requested stakeholder input on usability and potential improvements to the current calculator. - Conducted best practices research in estimating steam trap savings.	- The Working Group concluded the tool should be revised to further clarify inputs and add QC checks. - Phase II will implement these improvements.
Calibrate the STT with additional pre/post billing data analysis.	- Confirmed the value of calibrating the STT. - Identified a number of sites that could likely be used for calibration.	- The Working Group concluded another round of calibration of the STT is appropriate. - Phase II includes model calibration.
Annual Heating Plant Efficiency Factor		
Determine best practice to account for AHPE for load reduction measures, using reasonably available data.	- Documented implementer concerns and technical arguments through interviews. - Analyzed past project boiler meter data to estimate variance in boiler plant efficiency across the year and between sites. - Conducted sensitivity analysis of changes in annual efficiency on program savings. - Considered the role of plant losses on calculations.	- The Working Group concluded that the STT and other load reduction measures like insulation should use a deemed AHPE value from a set of average typical boiler plant efficiencies based on

Study Element Objectives	Phase I Tasks	Phase I Conclusions & Phase II Tasks
		certain boiler operation characteristics. - Phase II includes developing deemed AHPEs.

The research discussed in this report investigated the following questions:

- Are steam traps, as a measure, feasible to evaluate in a separate segment?
- Is there any indication of gaming or unintentional systematic manipulation of the STT?
- Is user-driven error by a specific group (be it vendors or PA implementers) contributing to poor evaluation performance?
- Are there organizational or algorithmic changes that could be made to the tool to improve its accuracy?
- How should the annual heating plant efficiency factor be treated by implementers and evaluators moving forward?

1.3 Conclusions, recommendations, considerations, and future research

The Team and the Working Group reviewed the results of the research and made five recommendations.

R1 – Continue Current Evaluation Sample Design Practices

Custom steam trap projects can be easily identified using benefit-cost ratio (BCR) model measure IDs, which are shared among all MA PAs. A steam-trap-specific RR can be integrated into the tracking system for all MA PAs if segmented from the rest of the custom gas population. However, the variability in PA RRs indicates that adoption of a single statewide steam trap measure RR is premature, however t breaking steam traps further into PA-specific measure RRs will increase sample size in order to meet PA specific precision targets.

The Working Group recommends continuing the current practice of reporting all custom gas measures with a single statewide or PA-specific RR. The sample design, however, should continue to include a steam trap subsegment meeting a statewide three-year precision target of ±20% at the 80% confidence level (80/20) to allow for continued monitoring of steam trap performance.

The Working Group recommends reviewing the desk review results from the new calibration projects to further assess PA variability and to consider the role of desk reviews as a tool in tracking discrepancies in field observations like hours of operation.

R2 – Modify the STT to Clarify Inputs and Streamline QC Review Process

STT inputs interpreted from field observations (hours, pressure, and boiler efficiency) contributed to a net measure level difference of 17% (of a 30% total difference) in the realization rate for steam trap projects from the most recent custom gas evaluation. While there is no direct evidence of systematic overstatement of specific input parameters, tool users indicated a need for additional guidance on the application of these parameters along with more capabilities to detect erroneous and unreasonable inputs in their review.

The Working Group recommends updating the STT to improve consistency among various users and bolster the review process with a number of automated QC checks. The Group recommends modifying and/or adding the following fields:

- Overhaul the tool's Read Me section with more guidance on parameter application.
- Require additional input fields from the implementers including monthly billing data, boiler characterization, and condensate tank temperature measurements along with a means of identifying the portion of all the traps at the site being inspected and fixed for each project.
- Incorporate additional default picklist values with user override capabilities for hours, trap model numbers (to inform orifice size), and trap application type.
- Incorporate additional QC functionality including automatic flags for unreasonable parameter entries, trap count and failure mode distributions among the inventory, and the proposed savings fraction compared to the annual billed usage.

Phase II activities for this recommendation include revising the existing STT to include the improvements noted above with stakeholder input and review.

R3 – Update STT Parameters via Billing Analysis Calibration

Stakeholders and tool users are generally satisfied with the algorithmic approach, but there are concerns that the results of the previous calibration understated the tool-generated savings. While calibration is useful in estimating more uncertain parameters, it may also indirectly overcorrect other key input values such as hours of operation. There are also likely systematic effects such as back pressure that were not investigated during the previous steam trap study that should be considered further.

The Working Group recommends recalibrating the STT using projects from PY2017 and PY2018. The calibration procedure would involve conducting customer surveys to verify details about each facility and its operation, investigating projects with the highest rates of trap failure, performing billing analyses for the projects deemed suitable for the calibration pool and then empirically deriving values for the most uncertain parameters in the custom savings equation using the accepted billing analysis results.

Phase II activities for this recommendation include conducting project screening including customer surveys, recalibrating analytics, and revising steam trap deemed values used in the prescriptive program.

R4 – Improve Access to the STT

Although version control has not been a widespread issue with the STT, some PAs have indicated an inability to access an unlocked version of the tool while other users have suggested that they have erroneously used outdated versions of the tool in the past. Both implementers and reviewers have highlighted a need for knowing when changes or updates are made to both specific project parameters and the tool itself.

The Working Group recommends posting the Steam Trap Tool (STT) online with a summary of revisions and communicating these updates to vendors and reviewers when the revised STT is rolled out.

R5 – Use Deemed AHPE in the STT

The Team found that, while AHPE values vary within only a narrow range with respect to firing rate at any site, efficiency values vary more significantly from site to site based on the boiler type (steam, condensing, non-condensing, and potential controls). Combustion efficiency measurements are a readily available and reasonably accurate proxy for estimating the AHPE.

Using a deemed AHPE value in the STT has merit. First, it is preferred by the implementer since the insulation and steam trap vendors do not have ready access to combustion measurements. Second, using a deemed value in the STT calibration process avoids potential double-counting of efficiency discrepancies. Estimating a deemed value combustion efficiency is a function of a few readily available parameters such as boiler type (steam, hot water, hot water condensing) and control types (e.g., linkage, parallel positioning, O_2 trim).

The Working Group recommends that measures that reduce the boiler output use deemed AHPE values that are based on average combustion efficiency measurements representative of typical boilers serving steam trap and insulation measures. Implementers and evaluators should use the same deemed AHPE in calculating savings. Evaluators should continue to gather combustion efficiency measurements to build a repository of boiler efficiency measurements for potential future updates to the deemed AHPE.

Phase II activities will include compiling available spot combustion measurements from other ex post evaluation sites and from the PAs to develop average combustion efficiency values by boiler type and controls.

1.4 Phase II Research

The proposed activities for Phase II include the following:

- Review the new calibration pool with desk reviews to identify PA specific discrepancy trends
- Compile available spot combustion measurements from previous evaluations and implementer data to develop average combustion efficiency values by boiler type.
- Revise the STT with a number of modifications aimed at improving the consistency of parameter application among users and further supporting the review process.
- Recalibrate the most uncertain parameters (leak factors, condensate return factors) within the tool using a billing analysis-based calibration procedure.
- Produce a new set of deemed values using the PY2017 and PY2018 program population data along with the revised equation parameters.

2 METHODOLOGY

Stakeholders designed the study so that results of each research task were discussed by a group of experts (Working Group) who made recommendations.

2.1 Working Group

The Working Group is a body made of representatives from PA evaluation, engineering, and implementation EM&V teams, the EEAC, and evaluation consultants. The Working Group is tasked with reviewing and approving the study approach, discussing findings, and drafting recommendations to the full MA C&I stakeholder group. The full list of Working Group members for this study is below.

Table 2-1. Working Group Members

Organization	Designated Member
PA Representation	Matt Siska, GDS Associates rep. Berkshire Gas / Columbia Gas / Liberty Utilities / Unitil Glen Eigo, Berkshire Gas Monica Cohen, Columbia Gas Alex Bothner, Eversource Sharon Jones, Lexicon Energy rep. Eversource Mike Mills, Eversource Aakanksha Dubey, National Grid Dave Jacobson, Jacobson Energy Research rep. National Grid Mark Dipetrillo, National Grid
EEAC	Ralph Prahl Jennifer Chiodo
Evaluation Consultants	Sue Haselhorst, ERS Jon Maxwell, ERS Cameron Kinney, ERS George Sorin Ioan, ERS Chad Telarico, DNV GL Ryan Brown, DNV GL

Working Group members attended biweekly meetings throughout the duration of the project to comment and deliberate on proceedings and findings and develop recommendations based on the topics discussed. A summary of the topics covered in each Working Group meeting can be found in the table below.

Table 2-2. Working Group meetings topics

Session	Topics
1	• Introduce the objectives • Discuss interview objectives and questions and the target populations for a) steam trap users, b) technical experts regarding steam trap saving calculations and boiler efficiency, and c) PA tracking system experts • Confirm the best practice research sources • Determine initial schedule of subsequent meetings
2	• Discuss findings and recommendations of PA tracking interviews and steam trap sample designs

3	• Discuss findings and recommendations of remaining tasks under ex post evaluation treatment of steam trap measures (discrepancy analyses and STT interviews)
4	• Discuss findings and recommendations regarding steam trap improvements and AHPE factor recommendations
5	• Review the conclusions and recommendations of the previous three meetings • Finalize the recommendations to be included in the draft report to be sent to the MA C&I stakeholder group

Minutes of each Working Group session can be found in Appendix A, while the findings and recommendations discussed during each session can be found in the task-specific sections of this report.

2.2 Interviews and best practice research

The Team conducted interviews with PA staff, vendors, and other industry experts to collect first-hand accounts on a number of topics including feedback on the current STT, methods related to calculating AHPE, and general best practices related to implementing, reviewing and evaluating steam trap projects. Research objective summaries are shared in the tables below.

Table 2-3. Interview objectives

Objective	Data Collected	Interview Description
Improve the steam trap tool	Solicit feedback on the current steam trap tool including usability and potential algorithmic improvements.	Guided interviews of PA technical staff and designated vendors. Targeting "power users" of the tool.
Determine AHPE best practices	Solicit details on feasible methods to calculate boiler plant annual efficiency.	Guided interviews of PA technical staff and designated consultants.
Optimize sample design	Specify how each PA tracks steam traps and any limitations on identifying steam trap projects.	Telephone discussions with PA representatives preceded by an explanatory email. One representative from each PA was interviewed.

The Team attempted to gather additional details on steam trap calculators and boiler plant efficiency estimation methods by conducting best practices research, as presented in Table 2-4.

Table 2-4. Best practices research objectives

Objective	Data Collected and Sources	Notes
Improve the steam trap tool	Review of TRMs and relevant white papers to identify methods, algorithms, and assumptions used for steam traps savings calculations.	This is a refresh of the research conducted as part of the P59 study. References include updated versions of state TRMs and any relevant white papers.
Determine AHPE best practices	Review references to identify methods used to calculate boiler plant efficiency to be used to convert thermal load savings into natural gas savings.	References included most up-to-date versions of state TRMs, relevant white papers, and resources from industry leaders such as the American Boiler Manufacturers Association and ASME PTC 4 standard.

3 EX POST STEAM TRAP MEASURES EVALUATION

Steam traps constitute a large share of custom program savings yet have historically underperformed compared with the other custom gas measures. Currently, both implementers and evaluators use a single statewide calculator – the STT – for estimating savings. Unlike other custom measures, steam trap savings estimates largely depend on the observation of unit installation as well as a few key inputs that characterize the heating plant including operating hours, steam pressure, and boiler efficiency. Considering the current nature of the measure, the question was raised whether the steam trap measure should be treated as a separate segment within the custom program contributing to individual PA custom RRs, or evaluated and reported as a statewide measure with a statewide RR, or evaluated in another manner.

For this task, the Team examined alternate sample design options that would support standalone reporting of steam trap measures with a measure-specific RR in annual reporting (referred to as "reported" in Table 3-1). In the current design, steam traps are sampled to meet a specific precision target (referred to as "sampled"); however, a steam trap RR is not used in annual reporting. The alternate designs considered different segmentation and precision target schemes, as summarized in Table 3-1.

Table 3-1. Sampling options: realization rates meeting three-year precision targets

All Precisions at the 80% Confidence Level	Steam Trap Measures	Non-Steam Trap Measures	All Measures
Option 1 – current approach	Sampled, but separate RR not calculated or reported: Statewide (20%)	Sampled but not reported: Statewide (20% or better)	Sampled and reported: Statewide (10%) and by PA (20%)
Option 2	Sampled and reported: Statewide (20%)	Statewide and reported: By PA (20% or better)	Not reported, but must meet 10% or better statewide
Option 3	Sampled and reported: Statewide (20%) and by PA (20%)	Sampled and reported: Statewide (20%) and by PA (20%)	Not reported, but must meet 10% or better statewide

The current sample design (Option 1) targets the statewide and PA-specific RR at 10% and 20% relative precision at 80% confidence, respectively, over the three-year rolling evaluation period. This design further stratifies each PA into steam trap/non-steam trap subsegments; however, while there are design precision targets for these subsegments, the annual report uses the same single custom RR to report adjusted gross savings.

Option 2 and Option 3 consider samples that are designed to yield a steam trap RR with sufficient precision to meet reporting standards for natural gas reporting (nominally 80/20) and to be applied to steam trap measures to report adjusted gross savings in the annual report.

3.1 Methodology

This task had two components. The first component determined whether steam traps could be easily segmented from the rest of the custom gas population. To support the first component, the Team conducted a brief survey with PA representatives from each MA PA to discuss the nuances of tracking methodology and how the measure is typically documented internally. For the second component, the Team calculated steam trap RRs and precisions under various scenarios using data from previous evaluation efforts (PY2016 and 2017).

3.2 Steam trap identification

The first step in this task was to determine if the steam trap measure could be identified within the custom population sample, and if this methodology was shared among all MA PAs. The Team found that measures are typically identified under BCR measure IDs. These IDs were implemented in 2016 and went into effect consistently in 2017. Measure IDs are updated with each 3-year planning process, providing a simple way to identify the measure within the custom population sample. The Team confirmed that two BCR IDs (G19C2a011: Whole initiative custom – custom traps under the C&I existing building retrofit program, and G19C2a035: Turnkey custom – custom traps under the turnkey/small business program) reliably identifies custom steam trap measures. There are other BCR measure IDs associated with steam traps, however, these are used to identify prescriptive and projects and are not associated with the STT discussed in this study.

3.3 Steam trap sample design

Given the results of the surveys, the confirmation that steam traps can be easily identified within tracking, and a measure-specific RR can be applied, the Team investigated previous evaluation results to test alternate sampling scenarios.

3.3.1 Recent evaluation realization rates

Table 3-5 displays the steam trap results for the overall project as well as PA-specific results for PY2016 (P79) and PY2017 (P89) data.

Table 3-2. Recent steam trap realization rates at the 80% confidence level

Study	PA	Trap / No Trap	Sample	Realization Rate	Absolute Precision+/-	Relative Precision
PY2016 (P79)	Statewide	Overall	53	82%	7%	9%
		Traps	24	73%	12%	16%
		Non-Traps	39	84%	8%	10%
	Columbia Gas	Traps	9	79%	8%	10%
	Eversource	Traps	4	47%	22%	47%
	Liberty Utilities	Traps	1	97%	N/A	N/A
	National Grid	Traps	10	86%	13%	15%
PY2017 (P89)	Statewide	Overall	31	87%	8%	9%
		Traps	13	70%	16%	23%
		Non-Traps	18	91%	9%	10%
	Berkshire Gas	Traps	1	78%	N/A	N/A
	Columbia Gas	Traps	3	38%	22%	58%
	Eversource	Traps	5	66%	28%	42%
	Liberty Utilities	Traps	1	92%	N/A	N/A
	National Grid	Traps	3	89%	14%	16%

N/A = not applicable

The first three rows for each study in Table 3-2. Recent steam trap realization rates at the 80% confidence level list the overall statewide custom gas trap/non-trap specific RRs for each evaluation. Traps have clearly been a lower-performing measure in the previous evaluations. The data presented in this table suggest that steam traps have been consistently overestimated in recent years, and there is some variability across MA PAs and across years for each PA.

Table 3-3 displays the assumed and final calculated error ratios from the previous evaluation efforts. The assumed error ratios for each study are the average of the assumed and actual error ratio from the previous study, rounded to the nearest .05, with a hard floor of 0.25 in accordance with evaluation protocols.

Table 3-3. Trap/no trap error ratios

Study	Group	Assumed	Result
PY2016 (P79)	Statewide no trap	0.60	0.47
	Statewide trap	0.60	0.59
	Columbia, all measures	0.60	0.22
	Eversource, all measures	0.60	0.61
	National Grid, all measures	0.60	0.31
PY2017 (P89) (theory, 0.6 was actually used)	Statewide no trap	0.55	0.36
	Statewide trap	0.60	0.63
	Columbia, all measures	0.45	0.05
	Eversource, all measures	0.65	0.43
	National Grid, all measures	0.50	0.29
PY2018	Statewide no trap	0.50	N/A
	Statewide trap	0.65	N/A
	Columbia, all measures	0.25	N/A
	Eversource, all measures	0.55	N/A
	National Grid, all measures	0.40	N/A

Table 3-3. Trap/no trap error ratios above, suggests there is more variability within the steam trap measure compared to the rest of the custom gas sample. The last three rows for each program year display error ratios for the three largest PAs for custom gas overall. Row 1 for each program year displays error ratios for the statewide population without the steam trap measure, while row 2 displays error ratios solely for the steam trap measure. Compared to other custom gas measures, the RRs for steam traps are much more varied. Sample size does not change drastically by segmenting traps from the rest of custom gas. Segmenting traps by PA may increase the required sample size, however.

The approved sample design for the current custom gas evaluation (PY2018) can be observed in Table 3-4. Note that the sample was designed to meet precision by PA (80/20) and across the state (80/10) over the three-year rolling evaluation period.

Table 3-4. PY2018 sample design with calculated precisions @ 80% confidence

PA	3-Year Precision Target	RP 1 Year	Sample	Pop.	Therms
Statewide traps	20%	38%	6	119	1,238,856
Berkshire Gas – no trap	10%	97%	1	22	141,263
Columbia Gas – no trap	20%	17%	5	120	2,093,402
Eversource – no trap	20%	35%	6	95	3,284,332
Liberty Gas – no trap	10%	78%	1	5	96,762
National Grid – no trap	20%	24%	6	174	2,913,956
Unitil – no trap	10%	86%	1	7	262,910
Overall	**10%**	**15%**	**26**	**542**	**10,031,481**

The Team sampled six sites specifically for steam traps and support estimation of separate steam trap and non-steam trap RRs. As the table above shows, the sample design for PY2018 was split between PA non-trap measures and statewide steam trap sites aggregated into two buckets. If a site has only a steam trap project, it was placed into the statewide traps bin. If a site has steam traps and other measures, it was placed in the "no trap" PA bin. This sample design was run to target 80/35 precision (where the 3-year target is 80/20) for the steam trap/no trap buckets, and 80/17 (3-year target of 80/10) for statewide custom gas as a whole.

During planning, the sample design for PY2018 was run for three different scenarios:

1. By PA statewide custom gas with an error ratio of 0.6: sample size of 27

2. By PA with statewide steam traps subsegment with a targeted precision of ±20% and an error ratio of 0.6: sample size of 33

3. By PA with statewide steam traps subsegment with a targeted precision of ±20% and adjusted error ratios (Table 3-3) sample size of 26

Scenario 3 was the final design for the current custom gas evaluation. Segmenting steam traps from the custom gas sample lowers the error ratio for the no trap by PA subsegments, so the effect on sample size was the same or lower compared with random selection within each PA. Keeping the measure separate from the rest of the custom gas population ensures that they do not represent a larger proportion than necessary.

From the above data, it is possible to produce both steam trap and non-steam trap RRs statewide, as demonstrated in the current evaluation sample design. This process does not drastically impact the required sampled size. However, if steam trap RRs were to be further segmented by PA-specific rates, the sample size would increase to meet precision targets.

3.4 Conclusions, recommendations, and considerations

Discussion with the PAs confirmed that steam trap measures can be identified within the custom population through BCR measure IDs. This is true even when sites consist of projects with multiple measures, including steam traps, as each measure is labeled with a measure identifier. Each MA PA confirmed that a steam-trap-specific RR can be integrated into tracking if segmented from the rest of the custom gas measures.

Steam trap performance, as expressed by PA and over time, shows more variability and a lower RR than other measures. Currently both implementors and evaluators use the STT to calculate ex ante and evaluated steam trap savings, which should result in low variation between PAs. However, the nature of the steam trap measures requires observation of unit installations as well as other key metrics such as operating hours, heating plant efficiency, and pressure. Based on other research paths in this study, the Team found variation among the PAs in how these variables are being assessed and input into the tool.

R1 – Continue Current Evaluation Sample Design Practices

The Team considered three sample design options, including the current design. Segmenting steam traps into a single statewide grouping separate from the PA-specific non-steam trap groupings has a small impact on sample size. Using the current sample design as an example, the sample size is the same or lower compared with an overall statewide approach, as segmenting steam traps from custom gas lowers the error ratio for the non-steam trap strata.

Table 3-5. Sample design options

	Precision Targets - 3 Year Combined	Reported RR	Conclusions
Option 1 – current approach	Custom Gas, Statewide - 80/10	Yes – Small PAs	Likely minimum sample sizes. Allows steam trap performance to be monitored.
	Custom Gas, by PA - 80/20	Yes – Large PAs	
	Steam traps, Statewide - 80/20	No	
Option 2	Non-trap, Statewide - 80/20	Yes – Small PAs	Similar sample sizes to Option 1. Variance from PA to PA argues against a single statewide RR.
	Non-trap, by PA - 80/20	Yes – Large PAs	
	Steam traps, Statewide - 80/20	Yes – All PAs	
Option 3	Non-trap, Statewide - 80/20	Yes – Small PAs	Likely a significant increase in sample sizes.
	Non-trap, by PA – 80/20	Yes – Large PAs	
	Steam traps, Statewide - 80/20	Yes – Small PAs	
	Steam traps, Statewide - 80/20	Yes – Large PAs	

If the steam trap measure was to be segmented by PA and by steam trap/non-steam trap designed to meet a specific RR precision target (Option 3), sample sizes will likely increase as a result. Discussions during a Working Group session also indicated that some PA staff are opposed to adopting a statewide steam trap RR (Option 2) while there is evidence of inconsistent tool usage amongst different PAs. Neither of these options are recommended at this time.

The Working Group recommends keeping the current sample design practice in place, where custom gas is sampled and reported to provide both statewide and PA specific RRs, which meets regulatory-defined precision targets. This sampling approach is also designed to efficiently provide informational results for statewide steam trap and non-steam trap subsegments, meeting the precision levels for steam traps. However, the Group does not recommend using this for reporting purposes. Instead, the Working Group recommends to continue the practice of allocating the PA's sample points between steam trap/non-steam trap subsegments to support possible future sample design revisions, which is dependent upon the variance in PA performance to converge. Considering the uncertainty surrounding the STT, it is recommended to document results using the first option (statewide and PA-specific custom gas RRs). This method will allow the PAs to be better represented by their respective custom gas performance until further analysis can provide insight into how to better tailor the sample design surrounding steam traps.

The Working Group recommends leveraging desk reviews with a telephone survey component of PY2017 and PY2018 steam trap projects planned for STT calibration to investigate trends in PA variations of inputs with a larger sample of steam trap measures. However, desk reviews have limitations in capturing hours of operation, which is the variable with the most discrepancies. The Team should also examine whether desk reviews can capture discrepancy trends or if metering should be required for more conclusive results on PA variability.

4 STEAM TRAP TOOL IMPROVEMENT

Both the implementers and the evaluators use the Steam Trap Tool (STT) to calculate savings for all steam trap repair/replacement projects processed through the custom program. The STT was last revised in 2017 based on recommendations from the P59 Steam Trap Evaluation study.[1] Notable revisions to the STT at that time included simplifying the energy savings equation by reducing the number of potential input variables and using utility billing analyses to calibrate the most uncertain values (that are not easily obtainable through measurement).

Since these revisions, steam trap projects have experienced poor RRs due to differences between the implementer and evaluator observed site conditions. The PY2016 and PY2017 custom gas evaluations found measure-specific RRs of 73% and 70%, respectively. In the PY2017 evaluation, 27% out of the overall measure level difference of 30% was attributed to three variables: hours (20%), steam pressure (4%), and boiler efficiency (3%).

Implementers suggested tool improvements to the Team to reduce the potential source of error related to operating hours, which contributes the largest overall impact. Implementers also suggested updating the tool so an improved quality control mechanism would catch errors associated with reporting operating hours and other variables subject to uncertainty.

This section summarizes the methods used to identify potential improvements, an overview of these findings, and the Working Group's recommendations.

4.1 Methodology

4.1.1 Interviews and best practice research

The Team conducted interviews with Steam Trap Tool (STT) users including PA staff and implementation vendors to solicit feedback on the current iteration of the STT including usability and potential technical improvements. The Evaluation Team also attempted to better understand how users interact with the tool to establish if the calculator itself is contributing to the low RR or if certain PAs use or interpret the calculator differently than others.

During interviews with the initial group of PA staff and implementation vendors and a number of industry experts working outside of MA, the Team inquired about steam trap savings methodologies in other jurisdictions and their associated best practices. The Team also conducted a review of relevant TRMs and white papers to identify methods, algorithms, assumptions, and best practices used for calculating steam trap savings outside of MA.

The Team developed an interview guide that focused on a range of topics related to the STT's usability, accuracy, and areas of improvement as well as best practices used throughout the industry. The interview

[1] http://ma-eeac.org/wordpress/wp-content/uploads/Steam-Trap-Evaluation-Phase-II.pdf

guide was developed as a means of driving the conversation but featured open-ended questions to avoid restricting the range of possible responses and to capture a variety of potentially useful improvements. The interview guide is included in Appendix C.

The Team performed best practice research by reviewing the documented approaches for calculating steam trap savings in other jurisdictions. Data sources for this task included an assortment of statewide technical reference manuals (TRMs), relevant white papers, and resources from established industry leaders such as the Department of Energy and American Boiler Manufacturers Association.

4.1.2 Discrepancy analysis

Steam trap repair/replacement projects have a lower realization rate than non-steam trap projects due to differences between program reported and evaluated operating parameters of hours, pressure, and annual heating plant efficiency (of which hours was the largest contributor to the low RR).

For this task, the Team examined the sources and distribution of differences in steam trap project savings from the PY2017 custom gas impact evaluation to determine if there are characteristic patterns by PA. The Team then reviewed evaluation reports from the sites with the largest sources of discrepancies to understand the information that was available at the site and how those details were used by implementers to identify patterns of potential misinterpretation.

4.2 Findings

4.2.1 Interviews and best practice research

The Team conducted a total of 10 interviews with a variety of personnel including PA representatives, MA steam trap vendors, and various industry experts (including both MA and non-MA based respondents). Questions related to the current iteration of the STT were directed at personnel who have regularly interacted with the tool (7 respondents), while questions related to other methods for estimating savings and general best practices were directed at all 10 interviewees.

The Team also reviewed a number of statewide TRMs and other publications from industry leaders to identify various methods, algorithms, and assumptions used for calculating steam trap savings. This exercise was largely a refresh of the best practice research conducted during the P59 study, and the Team found few changes to approaches that warrant consideration of the Working Group. The majority of best practice findings were revealed during interviews with non-MA based industry experts.

High-level findings are shared below:

- Stakeholders and tool users are satisfied with the general algorithmic approach.
    - There was a general consensus among interviewees that the tool is user friendly but could use more guidance on the application of specific parameters.
    - While the accuracy of the tool results are subject to the uncertainty of its inputs, many respondents highlighted the importance of sticking to a single method and making further refinements using the available information.
- PA staff do not believe vendors are systematically overstating hours of operation but have indicated the need for more guidance in the application of parameters, specifically hours.

- Respondents asked for additional clarification on what specific values are representative of in addition to indicating a need for pre-populated default dropdown values with an option to enter an alternate values or automatic QC checks to limit the variation in potential values.

- Implementing limits and caps in the tool may prevent gross overstatement but could create other issues.
  - There is a general need for the tool to be flexible but tamperproof. Although limiting the possible number of inputs isn't feasible, automatic QC flags to highlight potentially erroneous or outlier values would be helpful.
- Application reviewers could benefit from built-in tool checks.
  - Using additional fields to track revisions or color-coding updated values would help reviewers visualize changes to the project throughout its lifetime.
- Calibration of the steam trap tool is useful for approximating values of parameters that cannot be measured directly (such as leak factor and condensate return factor), but increases the risk of double-counting some effects like hours, boiler efficiency, or orifice size.
  - Some respondents suspect that the calibration results of the recent P59 study yielded values that understate savings and expressed their concern about another round of calibration without properly accounting for demographics of the calibration pool relative to the larger population of steam trap projects.
  - There are also likely systematic effects like back pressure[2] that occur as a result of high steam trap failure rates in a facility that were not accounted for during the previous round of calibration that should be explored further in Phase II.

A detailed summary of findings can be found in Appendix D.

4.2.2 Site-level discrepancies

The Team performed a review of the categorical discrepancies from the recent PY2017 custom gas evaluation to identify potential characteristic patterns that may suggest user-driven error within the tool contributing to the low RR. The discrepancies quantified in Table 4-1 are sourced from 13 steam trap projects drawn in the PY2017 evaluation sample.

Table 4-1. PY2017 statewide results – steam traps discrepancy impacts

Discrepancy Category	Error Count	Berkshire	Columbia	Eversource	Liberty	National Grid	Statewide
Hours	10	-2.5%	-9.7%	-4.7%	-0.4%	-2.5%	-19.9%
Pressure	6	0%	1.0%	-4.3%	0%	-0.3%	-3.6%
Efficiency	11	-0.2%	-1.0%	-1.3%	-0.3%	-0.4%	-3.2%
Failed traps	2	0%	-1.8%	0%	0%	0%	-1.8%
Removed traps	1	0%	0%	-1.6%	0%	0%	-1.6%
Quantity update	2	0%	-0.1%	-0.9%	0%	0%	-1.0%
Tracking savings	1	0%	0%	0%	0%	0.9%	0.9%
Statewide	**33**	**-2.8%**	**-11.6%**	**-12.8%**	**-0.7%**	**-2.3%**	**-30.2%**

[2] The presence of steam in condensate lines reduces the amount of steam that can flow through a trap and theoretically diminishes the savings potential of a trap when a high volume of traps are failed in a given system.

Parameter updates based on field observations were the most common and significant contributors to the low RR, with hours of operation having the most significant impact.

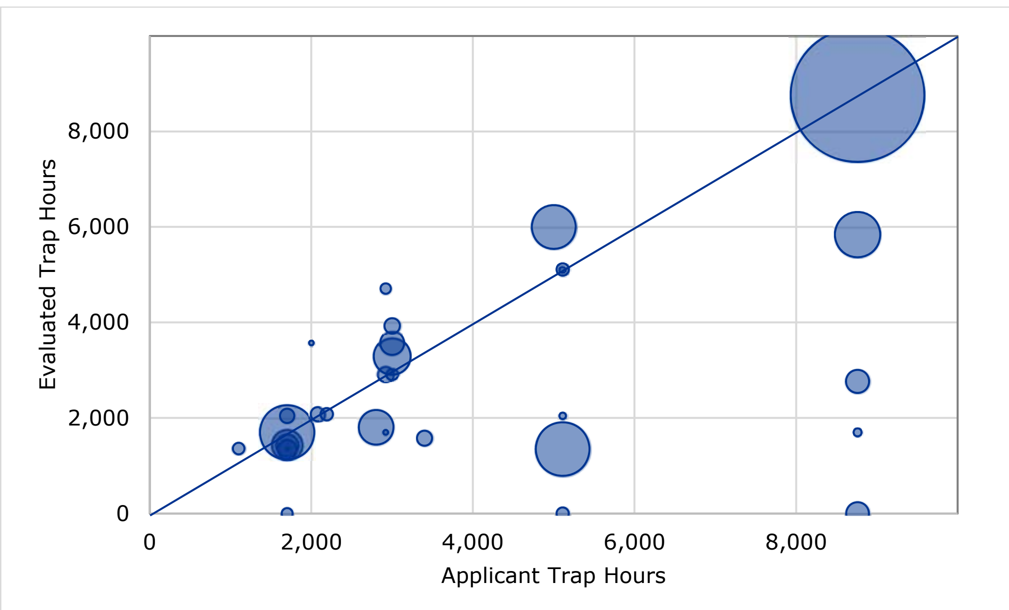
The Evaluation Team found that approximately one-third of the total measure impact was attributable to one prescriptive project run through the custom program. Discussions during the Working Group session indicated that PA staff for this project had at one point been instructed to file some prescriptive projects through the custom program but were no longer doing so.

4.2.3 Hours discrepancies

Differences between the program and evaluated operating hours had an impact of approximately 2/3 or 20% (of the 30% discrepancy) in the statewide difference between the program and evaluation results for the steam traps measure. The Team analyzed the applicant-submitted and evaluated hours values for each trap from every project in the PY2017 evaluation sample along with site reports to identify information that should have been available to vendors and reviewers at the time of implementation.

Figure 4-1 illustrates the relationship between the program and evaluated hours of operation. Each circle corresponds to a single trap; the size of the circle is proportional to the magnitude of tracked savings for each trap. Similar to an RR plot, traps that fall below the line represent instances where the hours assigned by the evaluator were lower than the program value.

Figure 4-1. Applicant vs. evaluated hours of operation



While the majority of steam traps fell along the unity line, there were a number of traps with both higher and lower hours of operation than anticipated. The two buckets that were most often updated to lower values are 5,100 and 8,760 hours, which represent a full heating season of operation and year-round operation, respectively. This finding could suggest that tool users are often misinterpreting the difference between boiler run time and pressurized full load hours. for the specific traps themselves, a point which was raised during multiple interviews. The majority of traps appearing above the line are found at lower hours values, which typically correspond with heating equipment that operates slightly differently than the implementer assumptions.

The Team reviewed the evaluation site reports of 10 PY2017 sample projects with hours discrepancies. In most instances, the evaluators had used metered data or staff-reported boiler plant schedules to update the hours of operation while implementers were limited to a dropdown list of 6 options based on trap application type. Table 4-2 lists each site that featured a difference in hours along with the tracked and evaluated weighted values and impact of difference at each site. The Team also reviewed the site reports for each listed project to identify the information used in the evaluation that would have been available to implementers during the project installation.

Table 4-2. Explanation of hours discrepancies by site

Site	Applicant Hours	Evaluated Hours	% Difference	Overall Discrepancy Contribution	Available Site Data Used by Applicant?	Reason for Difference
2017C0015	4,675	1,350	-71%	-10%	Yes	Misinterpretation of tool hour inputs
2017EV0087	2,802	1,806	-36%	-5%	No	Deemed application
2017N0194	2,078	1,467	-29%	-4%	Yes	N/A
2017B0009	4,752	4,132	-13%	-3%	No	Site contact interview
2017EV0076	7,347	6,633	-10%	-2%	No	Site contact interview
2017L0002	4,800	4,658	-3%	0%	Yes	N/A
2017N0129	4,464	4,387	-2%	0%	Yes	Misclassified end use
2017EV0057	5,892	5,826	-1%	0%	Yes	N/A
2017N0075	1,532	1,689	10%	1%	Yes	N/A
2017EV0081	3,593	4,259	19%	2%	No	Site contact interview

The majority of sites from the PY2017 evaluation were found to have hours of operation values less than 15% off from the original tracked value, most of which were updated using metered data captured during the evaluation. Note that 7 out of 10 sites listed in Table 4-2 contributed to a net discrepancy of -2%. The most significant site-level impact was attributed to misinterpretation of hourly input values in the tool (2017C0015) while the second most significant impact was due to the prescriptive site being run through the custom program (2017EV0087).

While reviewing these findings, the Working Group indicated that the hours discrepancy appeared to be predominantly due to QC review issues and expressed an interest in developing a procedure for validating the hours reported by the program participant and catching potential errors.

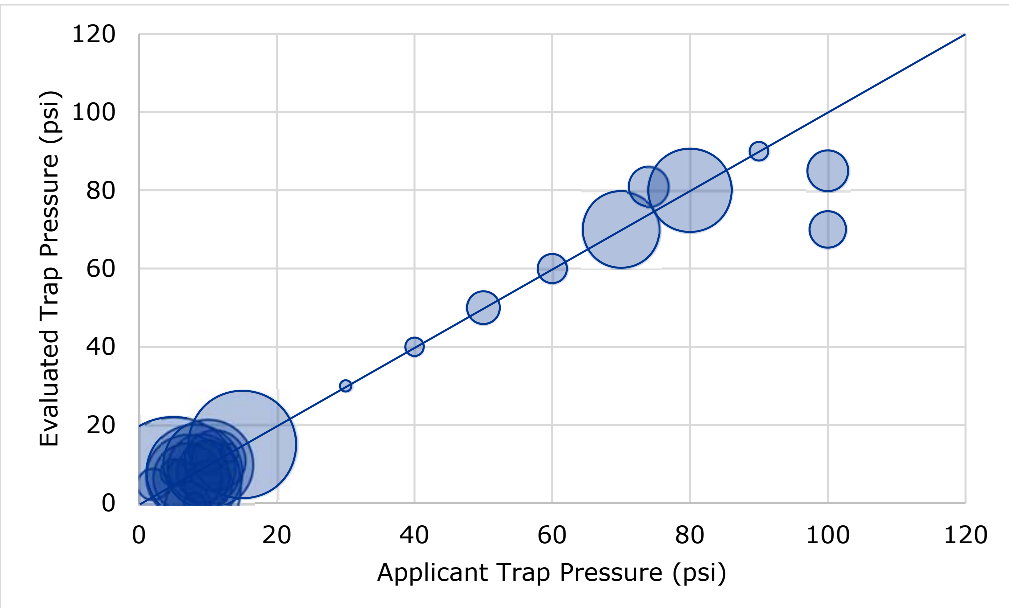
Interviews with PA staff and vendors indicated that while the application of hours of operation was fairly common sense, there is a need for additional clarification and guidance on how to apply them properly. One example given that highlights this need is the difference between boiler plant runtime and the hours in which a steam trap is pressurized. In some instances, respondents asked for a better explanation of each of the pick-list choices for hours and what they are representative of. Many respondents also indicated a need for pre-populated dropdown values or automated QC checks based on specific parameters such as facility type, trap application, or boiler control scheme, while also having flexibility to enter in a custom value. The intent of this suggestion is to improve consistency and reduce tampering by either restricting the number of potential input options or flagging values that appear to be erroneous.

4.2.4 Pressure discrepancies

Differences between the tracking and evaluated operating steam pressure had an impact of approximately 4% for the steam traps measure statewide discrepancy. The Team analyzed the program and evaluated steam pressure values to establish which steam pressure values reported by the program were different from the evaluated values and identified information readily available to the program that could have provided additional insights on the steam pressure and, potentially, could have help the program update the steam pressure.

Figure 4-2 illustrates the relationship between applicant-submitted and evaluated pressure values. Similar to Figure 4-1, each circle represents one trap and the size reflects the tracked savings.

Figure 4-2. Applicant vs. evaluated pressure inputs



With the exception of a few traps in the high-pressure range, evaluators found nearly all steam traps were reported with the correct operating pressure.

The Team reviewed the evaluation site reports for the six PY17 sample projects with pressure discrepancies. Steam pressure was updated largely using observed pressure gauges at each distribution line equipped with steam traps impacted by the evaluated project. Table 4-3 lists each site that featured a difference in steam pressure, their tracked and evaluated weighted values, impact of difference at each site, and the information that would have been available to evaluators by contacting the site.

Table 4-3. Explanation of pressure discrepancies by site

Site	Tracked Pressure	Information Available to Implementers	Evaluated Pressure	% Diff	% Traps Updated
2017N0194	10	Boiler plant operation	5	-50%	100%
2017EV0057	17	Site contact interview	15	-11%	91%
2017N0129	9	Boiler plant operation	9	-3%	11%

2017B0009	14	Boiler plant operation	14	3%	100%
2017N0075	2	Boiler plant operation	5	150%	100%
2017EV0087	N.D.	N/A – deemed application	6	N/A	100%

Despite the larger number of occurrences, the overall change to operating pressure and resulting impact on statewide results for PY2017 was approximately 4%.

While reviewing these findings, members of the Working Group suggested that the pressure discrepancy was a fairly minor issue relative to other topics being addressed in this study.

Interviews with PA staff and vendors suggested that, similar to the hours parameter (but to a lesser degree), pressure could benefit from additional guidance when entering values within the tool.

Interviews with non-MA steam trap experts indicated that including a dropdown with a limited selection of end uses being served by the trap is useful when validating both pressure and hours of operation. The current iteration of the tool does provide a field for the trap application but it is an open text field, not pre-populated with corresponding suggested pressures.

4.2.5 Phase II calibration

During the P59 study, the Team utilized a billing analysis-based calibration to derive values for the most uncertain variables used in the custom savings equation. The empirical calibration method employed in that study may have unintentionally overcorrected parameter values that were not fully examined during the study. Interview findings and discussions during the Working Group meetings suggest that the tool may benefit from another round of calibration. The revised approach can account for both changes in parameter values at the site level as well as the larger trap population demographics that were not fully considered during the initial calibration. Calibrating these parameters successfully will depend largely on having a sufficient number of sites that can pass the screening process while also being representative of the larger trap project population.

The Team conducted initial assessment of PY2016 and PY2017 projects to identify a number of candidate sites that may qualify for the calibration pool. The Team reviewed project tracking data from the two years of interest and identified any sites for which only one project was implemented (e.g., if multiple gas projects occurred in the same time span, they were removed from the candidate pool). The Team then used utility billing data records to generate average annual usage at each facility with a steam trap project. The reported project savings were then compared against the average annual usage at each facility. Any site with tracked savings greater than 5% of the average annual usage were identified as Phase II candidates. Of the 227 projects reviewed, approximately 97 projects met the criteria for the next steps in Phase II calibration, which would entail a phone interview to verify the feasibility of using billing analysis results to inform steam trap savings. Of those 97 projects, 43 exceeded 30% of the annual billed usage, suggesting the need for further review of the facility’s billing consumption prior to Phase II. The Team will conduct further review on the data used to identify Phase II candidates prior to executing customer phone surveys.

4.3 Conclusions, recommendations, and considerations

This section presents the key recommendations related to improving the STT.

R2 – Modify the STT to Clarify Inputs and Streamline QC Review Process

The hours and pressure input parameters contributed to a 24% difference (out of 30%) in realized savings in the PY2017 custom gas evaluation. In respect to hours of operation (which was the largest source of discrepancy on a count and value basis), the majority of sites showed evaluated values within ±15% of the original tracked values. Pressure discrepancies were found to be minimal both in the number of instances and overall impact to savings. The Working Group found these issues to be largely attributable to QC review and deduced that there was no clear indication of tool abuse or evidence of systematic overstatement of specific input parameters. Tool users indicated a need for additional guidance on the application of these inputs along with more capabilities to detect erroneous and unreasonable values.

The Working Group recommends updating the STT to improve consistency among various users and bolster the review process with a number of automated QC checks. The Group recommends modifying and/or adding the following fields:

- Overhaul the tool's Read Me section with more guidance on parameter application.
- Require additional input fields from the implementers including monthly billing data, boiler characterization, and condensate tank temperature measurements along with a means of identifying the portion of all the facility traps being inspected and fixed for each project.
- Incorporate additional default picklist values with user override capabilities for hours, trap model numbers (to inform orifice size), and trap application type.
- Incorporate additional QC functionality including automatic flags for unreasonable parameter entries, trap count and failure mode distributions among the inventory, and the proposed savings fraction compared to the annual billed usage.

Phase II activities for this recommendation include revising the existing STT to include the improvements noted above with stakeholder input and review.

Monthly billing data may not be immediately available to vendors when filling out the STT. PAs should consider implementing a process to populate this information internally during the review phase rather than requiring vendors to submit this information with the application.

R3 – Update STT Parameters via Billing Analysis Calibration

Stakeholders and tool users are generally satisfied with the algorithmic approach, but there are concerns that the results of the previous calibration understated the tool-generated savings. While calibration is useful in estimating more uncertain parameters, it may also indirectly overcorrect other key input values such as hours of operation. There are also likely systematic effects such as back pressure that were not investigated during the previous steam trap study that should be investigated further.

The Working Group recommends recalibrating the STT using projects from PY2017 and PY2018. The calibration procedure will include customer surveys to verify details about each facility and its operation, investigating projects with the highest rates of trap failure, and then empirically deriving values for the most uncertain parameters using a billing-analysis-driven calibration.

The calibration work needs to properly account for both site-level field input differences and the greater project population being represented. Recalibration must also account for changes to the revised annual heating plant efficiency values if accepted as a recommendation.

The previous study produced a per trap deemed savings value using the revised equation, empirically derived input values, and trap population averages from PY14 and 15. Newly recalibrated deemed values should be produced as part of the Phase II scope.

R4 – Improve Access to the STT

Although version control has not been a widespread issue with the STT, some PAs have indicated an inability to access an unlocked version of the tool while other users have suggested that they have erroneously used outdated versions of the tool in the past. Both implementers and reviewers have highlighted a need for knowing when changes or updates are made to both project-specific STT inputs and the tool itself.

The Working Group recommends posting a revised version of the tool online with a summary of revisions and communicating these updates to vendors and reviewers when initially rolled out. The PAs should consider developing a process to update deemed or fixed values in the tool based on evaluation M&V data.

5 ANNUAL HEATING PLANT EFFICIENCY FACTOR RESEARCH

The savings calculations for steam traps and pipe insulation measures use an annual heating plant efficiency (AHPE) factor to convert heat loss reduction (due to measures' installations) into natural gas savings. Instantaneous heating plant efficiency from the perspective of downstream load reductions can be measured using a combustion efficiency measurement and changes as a function of boiler loading. The AHPE is the load-weighted average combustion efficiency of the plant. The program implementers have been using an AHPE of 80% regardless of the site-specific conditions[3], while the evaluators have used site-specific spot measurements of the boiler combustion efficiency as a proxy for AHPE, when available.

PAs expressed concerns about whether a spot combustion efficiency measurement was a reasonable proxy for the AHPE. The concerns included whether a spot measurement adequately captured the yearly variation in load, whether the combustion efficiency measurement procedures accurately captured true combustion efficiency, and, finally, whether boiler standby losses should also be included in the AHPE. PAs requested that evaluators provide more details on how the boiler efficiency differences influence the differences between the tracking and evaluation results, how the boiler efficiency varies with boilers' firing rate, and what factors should be included in the AHPE value. The Team's research was designed to address four issues:

1. What is the impact of differences between tracking and evaluated AHPE values in the program results?
2. Does the firing rate vary throughout the year, requiring a load-weighted estimate of heating plant efficiency?
3. Is a spot combustion measurement a valid indicator of AHPE?
4. Should stand-by and purge losses be included in the AHPE, thus discounting the combustion efficiency to account for those inefficiencies?
5. Is an alternate option to define the AHPE without measuring efficiency appropriate for projects that reduce the output of the heating plant?

[3] The AHPE input within the current iteration of the STT is located in a locked cell within a hidden tab which does not provide all users with the opportunity adjust at will.

5.1 Methods

To determine the impact of the AHPE differences in the PY2017 results, the Team analyzed the steam traps and pipe insulation measures results of the PY2017 impact evaluation. This task provided details on the magnitude of impacts caused by differences in efficiency by PA and statewide. Next, the Team analyzed a sample of sites at which the Team conducted long-term metering of boilers. This task provided details on the boilers' firing rate profile over their annual operation and on the range of efficiency values as a function of the firing rate profile. The Working Group discussed the option of calculating a site-specific AHPE using metered data or using deemed AHPE values.

5.1.1 AHPE discrepancy analysis

The Team reviewed the site reports from the PY2017 impact evaluation for projects with steam trap and/or pipe insulation measures. Of the 21 evaluated steam traps and pipe insulation measures, 19 of them recorded differences between the applicant and evaluated AHPE values. Table 5-1 presents a summary of changes in PA and statewide RRs due to the efficiency differences.

Table 5-1. Changes in PY2017 RR due to AHPE differences

Measure Count	Differences Count	Berkshire	Columbia	Eversource	Liberty	National Grid	Statewide
Steam Traps – RR = 70%							
13	11	-0.2%	-1.0%	-1.3%	-0.3%	-0.4%	-3.2%
Pipe Insulation – RR = 82%							
8	8	N/A	2.8%	-0.7%	N/A	-4.0%	-1.9%
PY 2017 Statewide (All Measures) – RR = 87.3%							
							-1.1%

The overall statewide realization rate (87.3%) accounts for all measures and the -1.1% impact due to AHPE is the contribution of differences from steam traps (-3.2%) and pipe insulation (-1.9%) measures to the overall results.

To better understand how a change in AHPE impacts the program savings, the Team also conducted a sensitivity analysis on the results of the impact evaluation study for PY2017. To simulate how results would change because of a change in AHPE, the Team used an AHPE of 85% as a starting value and discounted that value by ±1% (e.g., 85%-1%=84% and 85%+1%=86%). Then, the Team applied the change in AHPE to the PY2017 results. Table 5-2 presents the results of the sensitivity analysis.

Table 5-2. PY2017 results – heating plant annual efficiency sensitivity analysis

Scenario	Steam Traps	Pipe Insulation	Program
PY2017 realization rate	**69.9%**	**82.1%**	**87.3%**
Evaluated AHPE increases from 85% to 86%	69.8%	81.7%	86.8%
Evaluated AHPE decreases from 85% to 84%	70.0%	82.5%	87.9%

Table 5-2 shows that a variation of ±1% in the AHPE translates to a ±0.55% in the PY2017 impact evaluation statewide results. For a difference from the value used by PAs (80% in average) and an 85% AHPE value, the impact on PY2017 results would be ±2.75%.

5.1.2 Heating plant efficiency – annual variability

To determine how the firing rate and boiler efficiency vary over the annual operation of boilers, the Team analyzed a sample of eight projects where the evaluators conducted long-term metering. For each of the sites included in the sample, hourly firing rate and combustion efficiency values were calculated using the following approach:

- Long-term metering of combustion fan amperage was used as proxy for firing rate. The combustion fan amperage is proportional to firing rate.
- Spot-measurements of combustion efficiency at various firing rates was used to derive a relationship between firing rate, combustion fan amperage, and efficiency.
- Hourly firing rate and combustion efficiency were modelled using an independent variable (OAT, or production).
- The energy model was calibrated to billed consumption over the metering period.

It is important to note that these boilers received extensive metering because they were boiler efficiency projects, typically with added controls. The combustion efficiencies may not, therefore, be typical for a boiler serving an insulation or steam trap measure; however, the variability in firing rate and efficiency is expected to be typical.

The Team calculated the percent of time each boiler operates at each of the six firing rate intervals (e.g., 13% of the time a boiler operates at firing rates between 21% and 40%). Table 5-3 presents percent time at each firing rate interval for each boiler metered at the sites included in the analysis.

Table 5-3. Annual firing rate profile

	Firing Rate					
ID	**Off**	**(1%-20%)**	**(21%-40%)**	**(41%-60%)**	**(61%-80%)**	**(81%-100%)**
122	0%	43%	0%	0%	0%	57%
185	51%	13%	36%	0%	0%	0%
306	34%	39%	23%	4%	0%	0%
353	39%	34%	28%	0%	0%	0%
109-B1	44%	0%	56%	0%	0%	0%
109-B2	44%	0%	0%	0%	43%	13%
S_NST557	0%	0%	0%	94%	6%	0%
S-NGR373	0%	0%	26%	36%	37%	0%
Overall	**27%**	**16%**	**21%**	**17%**	**11%**	**9%**

The results presented above show the firing rate varies considerably over the annual operation of the boiler as well as between sites.

Next, the Team analyzed the profile of the boiler efficiency at various firing rates. The Team calculated the boiler efficiency at each of the six firing rate intervals (e.g., 85% efficiency when the boiler operates at firing rates between 21% and 40%). Table 5-4 and Table 5-5 show the results for steam and hot water boilers separately because steam boilers are typically less efficient than hot water boilers.

Table 5-4. Steam boilers – efficiency profile

	Firing Rate				
ID	**(1%-20%)**	**(21%-40%)**	**(41%-60%)**	**(61%-80%)**	**(81%-100%)**
353	86.2%	85.6%	0.0%	0.0%	0.0%
109-B2	0.0%	0.0%	0.0%	85.1%	84.5%
S_NST557	0.0%	0.0%	84.1%	83.7%	0.0%
S-NGR373	85.6%	85.5%	85.4%	85.3%	0.0%
Overall	**85.9%**	**85.5%**	**84.7%**	**84.7%**	**84.5%**

The efficiency of the four steam boilers listed in Table 5-4, above, show an average of 1.4% variation across the range of firing rate.

Table 5-5. Hot water boilers – efficiency profile

	Firing Rate				
ID	**(1%-20%)**	**(21%-40%)**	**(41%-60%)**	**(61%-80%)**	**(81%-100%)**
122*	88.0%	0.0%	0.0%	0.0%	86.7%
185	84.1%	84.1%	0.0%	0.0%	0.0%
306*	90.0%	88.1%	87.6%	87.0%	0.0%
353	0.0%	87.6%	0.0%	0.0%	0.0%
Overall	**87.4%**	**86.6%**	**87.6%**	**87.0%**	**86.7%**

* Condensing boilers

The efficiency of the four hot water boilers listed in Table 5-5, above, show an average of 0.3% variation across the range of firing rate.

The results presented above show there is a narrow range within which efficiency values vary across the firing rate values. Although the firing rate varies significantly during the annual operation of the boiler, its efficiency does not vary significantly over its annual operation. The efficiency values vary between different boilers depending on their type. It is notable that the average efficiency of the steam versus hot water boilers differ by approximately 2%.

The PAs raised concerns associated with the combustion efficiency measurements taken by evaluators. To reduce the uncertainties associated with spot-measuring combustion efficiency, the Team used the following protocol:

- Combustion measurements are taken during the heating season when the boilers are under normal operation.
- Allow boiler to reach a steady state when taking alternate firing rate measurement.

The protocol above ensures the measured values are not altered by boiler warm-up and firing rate fluctuations.

5.1.3 Annual heating plant efficiency – standby losses

PAs requested that the Team research the factors that are included in the boiler efficiency and how they change if the boiler output is reduced by measures such as steam traps and pipe insulation. The Team determined that boiler losses that change with boiler's heating energy output should be included in the AHPE. Losses that remain constant independent of load should be excluded. Stack heat losses (SL) change approximately in proportion with boiler(s) heating energy output and therefore should be included in the AHPE. The same efficiency value can be used for both the baseline and measure condition. Combustion efficiency accounts for SLs and can be appropriately used as a proxy for AHPE.

On October 2, 2019, the Team generated a memo that describes the heat losses associated with the heating plant and the methods for calculating the heating plant efficiency for calculating gas consumption impacts due to installation of measures that reduce the output of the boilers. The memo is provided in Appendix E.

The memo concludes that standby losses remain constant before and after the implementation of a measure that reduces load (like insulation or steam trap repairs) and cancel out in calculations. The boiler combustion efficiency is therefore the key value required as input to load reduction measure calculations.

This was discussed further in the Working Group. The Group concluded that, for measures that reduce the boiler output, the combustion efficiency is an accurate proxy for AHPE.

5.1.4 Annual heating plant efficiency – alternate methods

In interviews and Working Group discussions, implementers expressed a strong preference for using a deemed heating plant annual value rather than require vendors to incorporate site-specific combustion efficiency values in the savings calculations that support the project applications.

The Team noted the STT will be calibrated using actual consumption from billing data. Using AHPE deemed values as inputs to the STT to produce the site savings estimates will calibrate the models to those deemed values, which should help compensate for uncertainty in the efficiency inputs.

The Working Group discussed this alternate approach and concluded that using deemed values for AHPE will allow implementers to focus on the primary scope of the measure (e.g., steam traps repairs, pipe insulation) to ensure that operating conditions of the measures are captured. If the evaluators use the same deemed values in computing evaluated savings, the STT estimates of savings should be accurate. However, the Working Group concluded it may be prudent for evaluators to continue to collect combustion efficiency spot-measurements and document the values for potential future updates to the deemed values. The deemed values for AHPE should be determined based on information provided by PAs and past impact evaluation studies.

5.2 Conclusions, recommendations, and considerations

The Team summarized the findings and the Working Group conclusions and formulated one recommendation associated with the AHPE.

R5: Use Deemed AHPE in the STT

The Team found that, while AHPE values vary within only a narrow range with respect to firing rate at any site, efficiency values vary more significantly from site to site based on the boiler type (steam, condensing, non-condensing, and potential controls). Combustion efficiency measurements are a readily available and reasonably accurate proxy for estimating the AHPE. Furthermore, the AHPE value used in the load reduction measure calculations should not include standby losses, since these are constant in the pre and post conditions.

Using a deemed AHPE value has merit. First, it is preferred by the implementer since the insulation and steam trap vendors do not have ready access to combustion measurements. Second, using a deemed value in the calibration process described in the previous section avoids potential double-counting of efficiency discrepancies. Estimating a deemed value combustion efficiency is a function of a few readily available parameters such as boiler type (steam, hot water, hot water condensing) and controls.

The Working Group recommends that measures that reduce the boiler output use deemed AHPE values that are based on average combustion efficiency measurements representative of typical boilers serving steam trap and insulation measures. The deemed values should be produced from a large data set of combustion measurements from both implementer and evaluator data files. Deemed values should be developed for subsets of boilers (i.e. steam vs. hot water) as supported by the data.

Implementers and evaluators should use the same deemed AHPE values to convert heat loss reductions into gas savings for steam trap and pipe insulation projects. Evaluators should continue to gather combustion efficiency measurements to build a repository of boiler efficiency measurements.

Phase II

Phase II activities will include compiling available spot combustion measurements from other ex post evaluation sites and from the PAs to develop average combustion efficiency values by boiler type and controls.

APPENDIX A: MINUTES OF MEETINGS

MA20C05-G-STBE Working Group Meeting #1 Minutes

Date: May 8, 2020 10:00-11:00 AM

Attendees: Cameron Kinney, Sorin Ioan, Sue Haselhorst, Mike Mills, Aakanksha Dubey, Alex Bothner, Chad Telarico, Glen Eigo, Jen Chiodo, Mark Dipetrillo, Matt Siska, Monica Cohen, Ryan Brown, Sharon Jones, Ralph Prahl, Dave Jacobson, Jessica Genest

Workplan

- Team has proposed to submit revised workplan following discussion of objectives and approaches with Working Group.
- Key additions to Workplan based on feedback include addressing differences among hours of operation and adding a hypothesis for corrective action
- DJ – Making major changes to the ST tool takes years to work its way through the system – changes being recommended/implemented should be significant
    - RP – Not necessarily an issue if the changes aren't major, ensure that we understand consequences ahead of changes
    - MD – Multiple tools will be confusing across the market – goal is to unify

Background & Recap

- MM – Concerned about billing analysis for recalibration (Phase 2 task). Most projects save <5% of annual usage, only selecting projects with savings >5% of annual use will bias results
- As of PY2018 all PAs are using the revised custom express tool
- Distribution losses have not been included in measure savings for recent evaluations (refers to recent memo submitted in October 2019)
- DJ – Is combustion efficiency sufficient for Heating Plant Annual Efficiency Factor? Are there better or more applicable values out there?
    - MD – Efficiency readings are often done on cold boilers, values are too high (not representative of average conditions)
- GE – When evaluating previous projects did they consider the additional traps that may have failed since initial study/repair?
    - CK – Not used in P59 study but recent evaluations had a failure rate adjust factor

Objectives

- MM – traps claimed to operate 8760 that only run during heating season need to be caught by the PA reviews (has been much more sensitive reviewing this information since the P59 study).
    - SH – We want to look at the information available and figure out if tool contributes to this issue.
    - MD – Study should yield suggestions for recommended run hours or a method for determining run hours
- SH – MM had mentioned measure life in an earlier email – not in scope of work for this effort but will forward the expected failure rate memo from previous evaluation along to the group.
- Team is planning to look at heat loss sources that should be included in boiler load reduction measures
- National Grid has combustion data that could be shared for purposes of the study

Interviews

- Three groups of interviewees, potential overlap among all groups
- Interviewees are going to be primarily PA representatives and vendors/external parties nominated by PAs. Anticipating PA assistance with recruiting these non-MA
- Potential interviewees at Enbridge Ontario – MD can provide studies and contacts
- AB – communication documents for non-PAs in MA need to be approved ahead of reaching out to those parties.
- MD – non-MA interview targets will likely be more difficult to recruit
- Interview nominees
    - MM, MD, MS from WG
    - Vendors – APM, Bumper, Greg Foote for Steam Traps; Cleaver Brooks, Spirax Sarco for Boilers
    - Adam Jacobs (EEAC team)
    - SCG?
- CK to reach out to confirm interviewees and solicit additional nominees from WG

Remaining Sessions

- Group to meet every other Friday through the end of June

MA20C05-G-STBE Working Group Meeting #2 Minutes

Date: June 5, 2020 10:00 a.m. – 11:00 a.m.

Attendees: Matt Siska, Glen Eigo, Monica Cohen, Alex Bothner, Sharon Jones, Mike Mills, Jessica Genest Aakanksha Dubey, Jaclyn Rambarran, Dave Jacobson, Mark Dipetrillo, Ralph Prahl, Jennifer Chiodo, Cameron Kinney, Sue Haselhorst, Sorin Ioan, Elana Cole, Chad Telarico, Ryan Brown, Glen Eigo, Ben Jones

Interview Guide Recap:

- So far, DNV has completed interviews for all the PA's regarding the tracking system
- Interview guide Cameron sent out - is the group okay with this guide?
  - Mike M. – has not had the chance to review the guide yet
  - Mark D. – Didn't touch on the run/heating hours
- Interview Plan:
  - Starting interviews next week, we are going start with the folks in the working group
  - Mark, Mike, and Matt – Cameron to reach out and schedule an interview with you early next week. Plan to touch on these items (i.e. run hours), cannot capture the best information here in the guide
  - ERS plans to receive the final comments from the guide today (6/5) and will send out one last version and will also use this guide for next week
  - Alex B: Communications team did approve the guide
  - After interviews are complete with the working staff, there is a slew of non-PA interviews to complete.
    - **Adam Jacobs (EEAC) –** Ralph can reach out to him and set up an interview
    - **RISE Engineering** – Mark D. can help set up that interview
    - **American Plant Maintenance -** Mike M. suggests Jonathan Davis for the interview
    - **Steam Trap Systems** – Mark D. can take the lead on an introduction email for this interview
    - **Greg Foote** – Mike M. has had conversations with Greg, he can reach out to him for the interview
    - **So Cal Gas** – Sue H. has some connections here.
    - **Enbridge -** Chad has a connection here DNV can send an introduction to
- PA Tracking Interview Session (Interview Findings) – (Ryan B, DNV GL)
  - Objective:
    - Look at sources of discrepancy and usability of tool. The interviews will help understand tool usability and tweaks to make to the tool
    - How to handle steam tramp realization rates – separate them moving forward?

    - **Conclusion:** Each PA has the ability to implement a trap specific RR in the tracking system
- Findings:
    - Steam traps identified under 4 BCR ID buckets (2019-2021 planning process BCR ID's)
    - ID's will be updated with each 3-year planning process – we will see 4 separate ID's/bins
    - Small business split out from C&I retrofit by customer consumption
        - For steam traps, doesn't matter as much for delivery and who is working with these sites
    - Historically, steam trap end uses split between process, hot water and HVAC (from 2016-2018 tracking data). Most likely due to where the steam traps they were installed.
        - 2019 the end use is consistently HVAC among PAs
        - It will be a MA effort to identify end users moving forward
    - Measure life should be 6 years
    - Custom projects may differ slightly
- Conclusions:
    - Uniformity among PAs in how traps are identified & tracked – BCR model
    - Fundamental differences between PAs may be found in the tool itself
- Mike M: Regarding the deemed savings on the prescriptive path, that would be for someone replacing all traps or is it a value for replacing failed traps?
    - Ryan: Deemed savings (prescriptive value) typically is for turnkey small business approach.
    - Mike M: this is a valid value. But on a larger scale, there are concerns if someone says a trap is failing but doesn't want a to pay for/trust an audit and accepts deemed value, this will understate the savings
    - Mark D: From National Grid's perspective, there is prescriptive limit to 75 traps (past years did vary from this). Limit is in place so people replacing specific traps on their own won't create problems with the realization rate later on. For smaller projects, anything over 75 should have an audit (NGrid tries to do this but does not always happen).
    - Dave J: Do small radiator traps have a different value? The size because they are small (and low pressure), traditionally not seeing the same amount of savings. Can't take the true value on these projects.
    - Cameron K: Should be multiple deemed values, low- and high-pressure value – 15 psi is threshold for high pressure. Do you use both values?

- Monica: Deemed value in the statewide model for this measure is the lower pressure, on a statewide basis, don't track them separately. Vast majority is low pressure
- Mark D: A lot of jobs are in schools, maybe 25% of steam traps were failed and replacing 100%. Fought against this – if you do blast replacement program the end of the day the savings would be true. NGrid tires to guide 75 or under
- Mike M: Eversource has a very aggressive radiator trap program that we ran for a couple years, agree with the thinking behind NGrid's philosophy. Threshold is 50 traps, if less than that its impossible to be cost effective to bring in a vendor and do an audit. Issue with indiscriminate replacement is then we cannot go back for 6 years, they can but choose not to because of the 6-year measure life. Year 1, customer has some losses, but by end of 5 years, they have a lot of losses, and not in the customer best interest and have backed away in the radiator offering
- Sue H: Do we want PA specific realization rate? Are differences in practices valuable?
- Mike M: If we focus on small projects (50 traps or less), deemed value is a fair value. Have had situations where people replace traps and look for incentives – but if they did not get an audit, don't have any info.
- Cameron – In the P59 study, there was no rationale for where 50% value came into play. Perhaps make the savings a function of quantity of traps or most recent previous date of steam trap work. We can address moving forward, but 50% value is something we should take a closer look at
- Sue H: How is this tied into sampling?
    - We were getting a better explanation for how prescriptive deemed savings work, but in the end, our work is focusing on the custom project of the tool
- Sue H: Sounds like we can do a PA specific realization rate for steam traps but may take more sites.
    - Mark D: for the smaller PA's could be issue with the sampling size?
    - Sue H: if you are not big enough, get the statewide number
- Sue H: If we go forward with a PA specific realization rate, we would need to add more sites to get the precision for the larger PA's
- Mike H: Doing more sites to get PA specific steam trap realization rate for using the tool– is this part of this work?
    - Chad – yes, under this work in next custom gas evaluation
    - This is all for recommendations for the next custom gas evaluation

- Recent Trap RR's (Ben Jones, DNV GL)
    - Never produced PA specific realization rates for steam traps in P79, can go back and do it, but hasn't been done yet

- **PA79 (2016 program) + PA89 (2017 program) – custom gas projects where steam traps are a just a subset, not specific to just team traps**
  - Both projects didn't discriminate by steam trap or not in sampling. Was not based on measure type
- RR's – just for steam traps in both P79 + P89 studies
  - PA specific RR are just for steam traps in custom gas in 2017
  - First two rows are RRs for steam traps only in both studies
- Traps were low performing measures in the last two evaluations
- Since the changes, RR on traps given the limited sample, has decreased
- Sue H: Differences here is not due to tool, due to differences in hours and pressure observed on site. The analysis is done using the same tool. Savings per trap is different, anticipated going down
- Lesson from this slide - Consistently overestimating steam trap impacts, and there is some variability across PAs
  - Cannot tell if it has anything to do with the tool yet
  - Hours probably have a significant impact
  - In 2016 – same issue was impacting the billing analysis results
  - Sue - Tool could be inducing drive error? We will investigate if the tool is causing people to use it incorrectly

- Trap/No Trap Error Ratios
  - More variability within traps compared to other measures. RR on traps is all over the place compared to other measures
  - When separate out traps from other PA measures, you need smaller sample sizes within PA if the statewide trap bundle is put together
  - Sample size does not change drastically whether you segment traps from custom gas. Segmenting traps by PA may increase required sample size
  - *Everything to the right of the center column - steam traps separated into their own bin. The 2016 P79 results column above the bottom line, there are no traps there. One row that has trap in it in the whole table*
  - Adding traps back in does not change error ratio that much
- C020C05 sample design – current study in the field right now
  - 2018 program year
  - This design was run with PA non traps and all statewide traps in one bin
  - Statewide traps defined:
    - If a site only has steam traps, in one bin
    - If steam trap and another measure, "no trap" bin

- Final questions:
    - Sue H: With this design, we can have statewide realization rate over time.
    - Sue H: Instead of a custom RR, do we have custom and steam traps?
        - We are on track to have a statewide trap specific RR, it is different from other custom measures. All tools will allow for this.
        - Ralph: On one hand, there is some variation in RR for steam traps across PA, argue RR at PA specific level
            - Variability is so high, take a lot of effort to consistently get PA specific/steam trap RR
            - Sharon: Does it depend on what accounts for the variability?
            - Ask PAs what they are doing differently or looking at end results
        - If we do have a steam trap RR should it be by PA and should we design the sample for that?
            - Ralph – good to weigh in on issue, this is a peer evaluation methodology issue
        - SH: Ben can you bring in the previous evaluation to look at more data points? This might help with more numbers and sense of variability.
        - Sue: We can write up findings and formulate these thoughts into a document
            - Ralph – seems clear that we are persistently overstating impacts, it looks like some PAs might be initially overstating impacts more than others, not sure what is clear if consistent state wide, but some ideas as to why (overstatement is traced to operating hours). Initial changes made is to try to narrow the gap of overestimation to change how we approach operating hours. Then see what RR we are getting and what the variability is from
            - It is an incremental process to make these decisions
    - Sue – If we do have steam trap RR, should it be by PA and do we need to design sample for this?
        - Need to look at how the tool usage varies across the PAs
        - Don’t need PA specific rate if the tool usage is consistent
- Remaining WG sessions
    - Next meeting –dive into discrepancy findings and results of ST interviews
    - Meeting #4 – scheduled for Friday before the 4th of July – Cameron to reschedule

MA20C05-G-STBE Working Group Meeting #3 Minutes

Date: June 19, 2020 10:00 a.m. – 11:00 a.m.

Attendees: Ryan Brown, Matt Siska, Sorin Ioan, Cameron Kinney, Elana Cole, Jennifer Chiodo, Mike Mills, Mark DiPetrillo, Aakanksha Dubey, Chad Telarico, Alex Bothner, Sharon Jones, Glen Eigo, Sue Haselhorst, Ralph Prahl, Dave Jacobson

---

- Action items from last call:
  - A memo is being prepared to provide details that would allow potential discussion on the steam trap sampling approach and how to evaluate it in the future. The draft memo will be sent to the group for review in the coming weeks.
- Discrepancies found from evaluation compared to what we calculated - P89 (PY2017)
  - 13 stream trap projects from most recent evaluation, 33 discrepancies identified in total
  - Discrepancies based on count, not on the impact of magnitude
  - Largest impact is due to hours of operation, the other two impacts were operating pressure and stream system efficiency
  - Smaller differences – due to less traps being installed then initially reported. Some failed traps, and one measure that was removed
  - Dave J: Can see a similar chart for 2016?
    - Sorin I: There were issues in 2016 on the calculations, not initially using the new custom express tool to calculate impacts. We can create a summary if helpful and show in the next WG session.
    - Dave J: Asked to see results from other years, to see the same information.
    - ERS: To show results from another year side by side for next WG Session
  - Discrepancy broken down by PA – The table identifies a range of changes, but the table shows 1 site that accounted for a lot of Columbia discrepancies for hours. Eversource had 1 site that had a prescriptive approach accounted for most of the negative differences in Eversource.
    - Sharon: Can we confirm - most of these sites, except for the sites where removed traps, all categories are things that could have been improved at the TA. The values were found to be different
    - Sorin: We have some more details. For failed traps you will need to guess, for removed traps and cannot tell if the trap was removed
    - Alex: Confirm – 1 prescriptive site – this is for Eversource
      - Sorin: Eversource column includes everything, meaning it includes the prescriptive site values. Found instances where hours were slightly higher on certain sites
  - **Trap level discrepancies - hours**
    - Graph on the right – plot of every trap from these 10 sites that had discrepancies.

- There are 3 sites from the previous evaluation that are missing (no discrepancies in hours or pressure).
- Each bubble represents one trap, size is the claimed savings. X axis is hours submitted originally; Y axis is evaluated hours. The line represents unity with hours, if the bubble is below line, evaluated with fewer hours, if above the line, evaluated with higher hours.
- Color code – reason for each trap being updated
- In the 2,000 range in traps on the graph – they are laid over each other so you cannot see all of the traps. (CK note: *there are approximately 500 traps on plot, a clearer plot will be submitted in final report).*

▪ Glen E: All below the line, are they from 1 or 2 sites, or are they from multiple sites? Is there any commonality between these?

- CK: Mix of multiple sites
- Sorin: Here there are only impacts by site, not weighted, just difference by savings per site by taking evaluated results/program results
- List tracked hours and sites here and provide info on if what we think is available to implementers at time of application, show evaluated hours, difference per site, how many traps/line items we updated from the inventory for the hours (right most column)
- Differentiate here site by magnitude, sorted from largest negative to positive difference. Looked at sites that have difference larger than 15% +/-. Overall the program/implementers did a good job estimating hours.
- Top 3 sites, these had large impact on Columbia's results because there were some issues with the way the implementer entered the hours in the calculator.
- Dave J: Are we consistent between asking about number of hours that the boiler can run per year vs. what we think the steam trap themselves have steam going through them?
    - o Sorin: In the large majority of sites the evaluators are going to, and allowed to installed meters at, the limit is 20 traps. We install temperature probes in the vicinity of the traps to use temp as a proxy for operating hours.
- Dave J: Is there a difference between the ones with actual measurements vs. those based on interview? Do we usually have metered data to back up interview?
    - o For the site in the table, yes, we did. On the plot, color coding shows the traps which we didn't meter traps, just went by boiler plant schedule (mainly interviews with site contact)
- Dave J: If there is just an interview with a site, are the hours adjusted on what the interview says they are operating?

        - Sorin: This information in the table is available to implementers.
      - Sue H: Is there information available to the implementers that could have improved their savings?
      - Aakanksha D: Not available means not available to the application
      - Sue H: There is a negative discrepancy in the first three, one was tool interpretation error, the second was the deemed and didn't use parameters to calculate, and the third, made a reasonable estimate but the hours were lower.
    - We corroborate metered data with site interviews to determine the operating hours. Never meter an entire heating season, or year.
    - Matt: QC issue for the Columbia site. Issue due to user error.
    - Mike M: If you have the ability to pull a bill, can see if you are using gas in the summer. The vendors frequently have the tendency to list 8760 and not appropriate to do so. Biggest component of errors, but sometimes reviewing engineer doesn't know. 8760 can be legit at times. Can't rely on boiler usage – certainly can tighten this up
      - If had more options, or ability to override options that are kicked out, can improve upon this.
    - Mark D: with zoning, might know when boiler is enabled, with zoning though, valve might be off the whole wing and no way from the office to know that
  - **Trap level discrepancies – Pressure**
    - Methods – pressure gauge at the line traps are installed, or we measure/meter temperate and use as proxy for pressure, or confirm with boiler plant manager
    - Mike M – concern is that the steam system will have modulating values, as they are brough close. Is this a legit issue?
      - Sorin: Modulating the pressure, you don't have direct means of metering pressure and rely on spot measurements from the pressure gauge and interviews with the site contact.
- Steam Trap Tool Interviews
  - Completed interviews with 10 nominated parties. PA staff, vendors, EEAC rep, and 3 non-MA industry experts.
  - User friendliness
    - Generally easy to use, straight forward enough
    - Confusion over interpretation of certain inputs
  - Whether or not results of the tool are accurate
    - Depends on the inputs
  - Hours issue

    - General consensus – relatively straightforward, but need to provide clarification on what some hours are
    - Need for variety of drop downs and QC checks, specifically for facility type or trap application or boiler control. Number of ways to implement these checks based on some element or identifier of trap type
    - Need for flexibility, only so many options for hours you can select. Should be some wiggle room for the vendor.
    - Should evaluators be able to ding the operators with metered data if PA's are limited to a fixed value?
  - Should we be building more restrictions into the tool regarding how hours are entered (maybe based on facility type)? Adding more QC elements (maybe a billing tab?)
    - Mike M – would like to see either the ability to override hours if something is more appropriate, or additional hour options. Projects he typically look at is typically 1700 or 8760, but for a site zoned off, need a middle ground
    - Glen E – Can we add in more hours for different facility types and the ability to override them. Can put in correct hours for facility type but in those one-off issues, they can be overwritten if they need to.
    - Matt S - Agree with all feedback, most of his feedback was guidance for implementers on interpretation of hours. Using default value vs. trying to asses in the field and ask teams – greatest source of confusion. More flexibility in the tool and more guidance to the implementation team, reduce issue
    - Mark D – I like flexibility, we do need to give guidance. With industrial customers, they have different shifts. Need to add guidance to the users of the trap spreadsheet.
    - Sue H: Have a pick list, and hours and applications pop up. Or text descriptions in the tool right where you input the numbers. Would these lists be a good way to go?
      - Glen E: don't need to capture every single case but give people the guideline for facility type and give a range for potential numbers. Also give reviewers see if someone gets close to the right number
      - Cameron K: feedback from interviews that indicated 3 groups to give reviewers a general feel of what they are looking for, manufacturing/process see 8760 for most traps, institutional – steam year round with process and small muni heating bucket –
      - Aakanksha D: if someone wants to override, they can fill out a box to explain why they want to override the value
    - Mike M – accused some vendors that the sales reps doing audits were biased towards overstating hours. Learned used to be the case, but as we deal with more vendors, could be a factor. Vendors claim this is not how they compensate them, but revenue is made when deal is closed. When findings are presented, it is in their best interest to have high savings

        - Bias in vendors to push the envelope – see more latitude in the hours, but don't want the vendors to have free range
        - Jenn: feedback loop for vendors would be good, if inputs are appropriate, eligible for a bonus (or some incentive)
        - Cameron: needs to exist a level of trust with PAs and vendors, although it appears less gaming is going on now suggestion were made to make it tamper proof
- Next meeting July 17
    - Will cover steam trap tool improvements (from interviews) and boiler efficiency analysis findings

MA20C05-G-STBE Working Group Meeting #4 Minutes

Date: July 17, 2020 10:00 a.m. – 11:00 a.m.

Attendees: Cameron Kinney, Mark DiPetrillo, Ryan Brown, Mike Mills, Sorin Ioan, Sue Haselhorst, Chad Telarico, Sharon Jones, Aakanksha Dubey, Alex Brothner, Matt Siska, Adam Jacobs, Glen Eigo, Ralph Prahl, Jaclyn Rambarran

---

- Heating Plant Annual Efficiency Discrepancy Table:
  - Results with PY2017 – Results show statewide RR for entire program of 87%
    - Steam traps RR = 70%
    - Pipe insulation RR = 82%
  - Dave J: Did all have spot combustion measurements?
  - Sorin I: 19 have different values than values used by the program
- Q1: Combustion Efficiency Variability
  - Sorin I: To get more detail, went back and looked at random sites that were included in previous cycles. These sites replaced boilers or replaced pre-existing controls with new controls. We measured for longer time period (2-2.5 months in heating season).
  - Used combustion fan measurements as a proxy for firing rate
  - Mike M: Feel this is one of the issues, spot measurements are not indicative of year round site operation. Not considering other factors like varying air density. Many sites use different boilers (winter vs. summer loads). There are a variety of factors not just what is mentioned on this slide, and hard to draw solid conclusions. Stress – this is a <u>minor</u> engineering concern. Changing from a practice that was instilled in the TRM calculations. In addition, there are issues, with some results coming back with ignorance with technician – if efficiency is ~80%, not indicative of true value. We have the tools to go out and measure this, and we can measure it over longer period of time. Pick a number, and have all engineers use it.
  - Mark D: Agree there is a lot of variance on each combustion result because of different standards on the boiler; each boiler is different.
  - Sorin I: We should pick a methodology engineers can use on site if we have to do this.
  - Efficiency Findings for steam boilers
    - Large range of firing rates
    - Split boilers into steam boilers and hot water boilers, and split firing rates and associated an efficiency with each bucket
    - For each efficiency calculated, there is a range of values. Based on the research, the range is not too big for the firing rate.
    - Mark D: It make sense for efficiency to go up, but the range is surprising; but not as dramatic of a range as we thought in the past
      - Sue H: It is a smaller sample (10 boilers?)

            - Sorin I: Background – we pulled any projects with metered data and 8760 analysis and select 10 of them and see how data varies. Adding more would not make a difference if the data does not vary too much
            - Mike M: 86% is high
            - Sue H: Pick a value for the implementors to use. Might make more measurements to build out the average.
            - Sorin: Results are mostly to show the range, not the absolute values.
        - Sharon: Surprising efficiencies are the highest at the lowest rates, comforting to see all ranges
        - Mike M: Site 306 with 90% efficiency – assume this is a condensing boiler?
            - Sorin I: It is possible this is the condensing boiler.
            - Mike M: In a non-condensing boiler, lose 10% because of combustion.
            - Mark D: Can pick a single number and not worry about the variance
        - Sharon J: Where do efficiency numbers come from?
            - Sorin I: Evaluated sites from previous cycles PY2013 + PY2016. Had more boiler replacements and control measures, went back as far as we could to get measures to install boilers or controls.
            - Sorin I: Efficiency numbers are based on spot measurements, we extrapolated them based on firing rate of long-term measurements to know how much time is spent at which firing range.
    - Heating Plant Annual Efficiency – Sensitivity
        - Sorin: Figure out impact of increasing the efficiency by 1% values.
- Q2: Validity of measurements
    - Sorin I: We have protocols to apply while on site – during heating season, plan to start visit before heating season, go at the beginning of mid-January
    - Wait for boiler to reach steady state before taking spot measurement
    - Mark: Do you record return water temperature (for steam boilers)?
        - Sorin I: No, not something we do.
        - Mark D: Suggest for any combustion test going forward, for steam equipment, you should take water temperate tests – we are seeing cold water come back and changing combustion rate up
        - Sue H: Changes not built into this – not perfect, but trying to get as much uncertainty as we can
        - Mike M: thanks to contributions, taken huge strides forward to improve accuracy of reviews. Hours are big issue; this is a minor issue.
- Q3: what is included in heating plant annual efficiency?

- Sorin I: Main drive - we think the constant losses should not be included in savings calcs when we convert heat losses to gas consumption.
- Mark D: Does this include condensate loss?
  - Sorin I: Do not include this, we think this is something related to the measure itself
- Mark D: Some sites have a huge loss of condensate (like colleges and hospitals), sometimes it changes after you repair the steam traps
  - Mike M: Is this coming from steam leaks?
  - Mark D: Not all, only some are better but not all. Old universities were losing a lot
- Mike M: Right to see minor variations to this argument, overall. If steam issues, we repair are significant enough, plant has the ability decrease the pressure and pipe distribution losses should go down.
- Sue H: Can we deal with this as a custom project if we are changing the pressure? Can we build into this model, or exception?
  - Mark D: Should be an exception

■ Annual Heating Plant Efficiency Interviews

- Cameron K: High level points, combustion efficiency measurement is a good standalone point. If you can make only one measurement, this is the one. Realistically, there should be a lot of other things considered. General consensus, using fixed value is most reliable/non-controversial
- Cameron K: talk to some non-MA jurisdiction, line level efficiency inputs for different boilers, but if you don't want to apply site level efficiency, having option for line level efficiency is a good option
- Sue H: Have single, default value might make sense: develop that in this report for everyone to comment
- Dave J:X up with new single value in the tools, and not worry about site level adjustments?
  - Sue – yes, something to explore, we may continue to make measurements to build a data base. Maybe process to reset value when we do a calibration (in a few years).

■ Steam Trap Tool interview results

- Most information on this topic is in the report

■ Previous ST Tool Improvements

- Dave J - Why was there a 50% drop in the deemed?
  - Cameron K: Drop in deemed, Cameron to forward the report. A factor somewhere, updating using the current population rather than old data. Used empirical data to derive that 12.2

- Improvements for usability
  - Discussed in more depth in the report
- Technical improvements
  - Cameron K: Propose for 2017 projects, rerun input paraments and give u a better sense of actual difference in savings attributed to the tool. Table shows impact was minimal compared to other impact
  - Sue H: think about calibration process and period calibrated. Hours didn't seem as different as we see now. Going forward, make adjustment in the model for the hours through telephone call process to make sure calibration isn't going to double count hours. If calibrate for leak factors, squeeze out other things that we account for in the evaluation as well. If we settle on an efficiency average, would be a better representative value
  - Mike M: One thing I wrestle with to zero in on steam trap savings, 85% of projects (20 Projects) savings were less than 10% of gas usage. In the 2016 studies and before, rely very heavily on billing analyses. Not sure if we can get there, recognize took a hit in savings at Eversource (20%), not sure if we took a step forward in accuracy. Hours are overstated often and giving us an issue, zero-ed in on issue now, but still room to tighten up. Continue to battle with vendors, put in false numbers to inflate the savings. Reviewer needs to be diligent on the hours
  - Sue H: in other programs, info has been shared and hours are being overstated. A lot of applications so can't spend too long on them. For gas, see sites, see correlation with degree days, go 10% and can be confident in the outcomes.
  - Cameron K: don't have breakdown, conduction interviews for 55 projects/customers and using this information, there were cases where couldn't do billing analysis b/c it is largely production based usage and fluctuates, one of the reasons to remove sites from 55 to 28 was b/c there were too much uncertainty, or other changes at the facility. Look at billing analysis are indiscernible. Billing data didn't give us reliable results and tried to screen out where it didn't make sense.
  - Cameron K: as volume of traps at any facility goes up, more live steam in condensation line, savings for these traps is not going to be as significant as the first 10-30 traps. There should be diminishing returns

- Remaining sessions
  - Sue H: Draft the report and send next week (before the 31st). At next meetings, go through the recommendations and synthesize what we have heard.
  - Dave J: What is the final position on recalibration and billing analysis?
    - Sue H: Recommendation will be going to phase 2 or implement immediately. There are mixed thoughts but see this as a recommendation in the report. Will schedule ad hoc meeting if we think it is needed. Will provide the report to formulate thoughts and can have future discussions in the next week and respond to any concerns if necessary.

- Dave J: Did we get enough from the interviews to write recommendations?
    - Sue: We need to put a cost in this as well – here is the level of effort required

MA20C05-G-STBE Working Group Meeting #5 Minutes

Date: July 31, 2020 10:00 a.m. – 11:00 a.m.

Attendees: Cameron Kinney, Mark DiPetrillo, Matt Siska, Aakanksha Dubey, Ryan Brown, Mike Mills, Alex Bothner, Sorin Ioan, Jaclyn Rambarran, Sue Haselhorst, Ralph Prahl, Jennifer Chiodo, Chad Telarico, Glen Eigo, Ben Jones

---

- Sampling for Steam Traps (Ryan B)
  - Confirmation that steam trap measures can be integrated into tracking for all MA PA's.
  - Sampling – trap and non-trap segments/buckets
  - Keep measure separated ensure steam trap is not randomly made to be a larger representation than necessary
  - Segment by PA – increase sample design, but not recommended at this time
  - Aakanksha D - Is there a PA specific steam trap RR rate?
    - Ryan B - No PA specific RR rate, everyone will have the same RR. Will not have RR rate segmented for steam trap segmented from custom gas. Sampling will be segmented, per steam trap and per the rest of the custom gas population
    - Ralph– Impression is, right now a fair amount of PA variability. PA should think about being lumped with other PAs. No problem with this, but important for people to understand this.
    - Aakanksha D – Surprised by statewide RR for steam trap. Everyone is using it in a different way, and not comfortable with a statewide RR for steam trap. Are we sure everyone is using it in a similar way?
    - Sue H: Do we have available what the PAs?
    - Mike M - Eversource was a poor performer and the overall results were not a fair representation of how ever source was doing
    - Alex B - How much additional sampling is needed to get the best precision level? Statewide value might be better
    - Aakanksha D - Biggest variability was seen and a lot of uncertainty amongst the user. Steam trap is a popular measure with a decent amount of savings. If we can do something that addresses both recommendations, and make sure users are clear on how to use the tool, we can keep using PA specific RR. In future when we are sure everyone is using it in a similar way, can switch to statewide RR for steam trap. Highly user dependent

- Chad T – The sample design can handle either direction (sample by trap or no trap). If this is what the PAs want, we can continue to go with this approach, and assess how any improvements to the steam trap tool are getting implemented and over the next 1-2 years come back and see if the state wide trap result should be used
- Ralph P - Make quantitative assessment to see if some PAs are doing better than others at deploying steam traps. See how good each PA is populating fields, etc. could keep it a statewide result for now, but do an expanded sample and less intense analysis to see how much variation is between PAs
- Aakanksha D - After 2 more evaluation cycles, see how the user is using the tools within the PAs and then go for a statewide RR for steam trap. For now, PA specific is better.
- Sue H - Can we identify steam trap segment in each PAs custom sample and have 2 segments in sample, with steam trap as a separate segment and everything else is non-steam trap and that will have a PA specific. some steam trap segment will be distributed among PAs. This is a different sample design. Stay with what we have now and sample by PA first?
- Aakanksha D - Right now, we should stick with PA specific steam trap RR and in the future, assess how is the usability improving and consistent among everyone and then we can analyze if we need a statewide RR
- Ryan B - Steam trap on MA level, rest of custom gas segmented out in PA Specific. Go into phase 2 work and see variability and then assess if we need to further segment PA specific RR.
- Alex B - Keep sample size, not get PA specific steam trap RR. But get non-steam RR. Compare variability to past results. If variability isn't increasing, keep as statewide.
- Chad T – Should we stick with what we are doing right now, which is a PA only RR for trap and non-trap with everything included. Not a PA trap RR, but a PA RR with everything included
- Aakanksha D - If we don't have a sampling strategy to have a PA specific RR for steam trap, fine with looping in steam trap into other custom project like doing right now. Can we have a PA Specific RR (even if steam trap is included in custom). There is variability between PAs
- Ralph P - There are things we can do short of the relatively intense per site M&V costs of our current evaluation approach. We can keep in the current study the sample design but we can add a larger sample that is analyzed and less intense (just desk review) that aims at quantifying the performance by individual PA to get a

better handle on PA variability. Right now, cannot get too much insight, we have a small sate wide steam trap sample.

- Sue H - We are going to wrap up the gas impact evaluation, next time we have a gas impact evaluation, have an augmented steam trap sample with a desk review with calls and decide from this if its PA specific?
- Chad T - Can we pull in some additional steam trap from current group and do desk reviews and not wait until the next gas evaluation? We can wait but don't need to.
- Sue – This could fit in with the calibration that will be discussed later in this working group and proceed with this. We will go with a DR process and leverage that to get some PA specific findings that can inform this a bit better. This will be pulled from some more recent projects
- Mark D - Will need to figure out heating hour variability and then check on this
- Mike M - Biggest issue are hours and constrained with options now. We should focus on being more accurate with the hours.
- Sue H - Work through issues with single statewide RR makes sense, but there is PA variability we already observed. Stick with PA specific RR that includes steam traps until we demonstrate we can have a single separate steam trap RR?
    - Alex B - How do we demonstrate we are ready?
    - Sue H - wait for another impact cycle. Might be premature to sample other projects.
    - Alex B - what indicators would we be looking or to say yes we should have a separate steam trap or PA rate?
    - Ralph P - Establish criteria to see what constitutes a successful deployment for a given case and do a set of DR to quantify successful deployment.
- Sue H – We will craft this recommendation, and everyone will have an opportunity to review the document. See if we can come to some decision and have an ad hoc meeting if needed. Conclusion: hearing some themes, are we ready to move to a single steam trap statewide rate? What is the trigger to say we are ready to move to the statewide?
- Alex B - What the benefits and the associated costs?
    - Sue H – This will be in the report.

■ Recommendation 2: Annual heating Plant efficiency

- PY2017 – the difference in annual heating efficiencies with steam trap and pipe insulation was .6% statewide
- Recommendations: statewide (implementers and evaluators) use a deemed value to convert heat loss savings into gas savings. Proposed Phase 2 to compile available spot combustion measurements and analyze and process data and come up with an efficiency for different buckets

- Recommendation 3: Field observations
  - Steam trap recommendations are mostly tool specific
  - Steam trap RR for PY2107 was 70%
  - No direct evidence of systematic overstatement of hours, though PAs did note they were watchful. Bigger findings was that more guidance is needed on the application
  - Findings – general need for tool to be flexible to have a range of inputs and detect errors and filter out unreasonable values
  - Recommendations: Update the tool including read me section, add some fields for the implementer to input (billing data requirement, what controls are being used, etc.)
  - Mike M - Billing data is something that would be a nice addition, but trap vendors do not have access and would have to go back to the PA.
  - Sue H – With this data, its part of a good check (if we can get it). Intention is to help inform implemented, but we need to think how we can get this data. This is check that can improve QC. Know percentage of totals is important.
    - This will be in the report and can respond here.
  - Mike M - onerous to put in all the trap data, to make vendors life easier: list failed traps and total traps that were inspected and total traps on site
  - Mark D - You cannot check total usage of the system if you go this route
  - Cameron K - Recommendations - Already have a pick list for hours, can provide more guidance here. Lookup function that spits out data. Allow for custom entries (but have this be highlighted and called out to the reviewer).
  - Sue H – Will write the recommendations this way, and address this further in the report
- Recommendation 4 – Accuracy of the tool
  - Propose to redo this calibration effort and we understand more how we did it in the previous round. We will recalibrate using new sites. Consider a high-level desk review and add another element and ask for things like boiler controls, and do a soft check on the hours just to confirm, and make sure things make sense at a high level

- Alex B - Would this produce a new deemed value for prescriptive measures?
  - Cameron K - We can incorporate the results from the previous steam trap study into these findings to define deemed value.
  - Sue H – Use new population distribution with the new tool, seems possible?
  - Alex B - this study is recommending future work to produce a deemed value? In this report, this would recommend for the future work, looking at prescriptive Projects and producing a new deemed value.
  - Ralph P - How do we develop current prescriptive deemed value?
  - Camron K – This was done in the P59 study. We used revised equation and used combo of average values (pressure, hours, etc.) and with basic assumptions with boiler efficiency values to generate what was a typical trap with typical failure rate. 2013-2015 program years
  - Sue H – recommendation for calibration is a go. Will put this in the report as a recommendation and can comment back

- Recommendation 5: Best practices
  - Post tool online somewhere
  - Develop process to update deemed values on EM&V data
- Phase 2 Proposed Scope
  - Alex B - Does this include the calibration schedule?
  - Sue H – Yes
- Next steps:
  - Report will be distributed by the end of next week
  - Final draft will be sent to everyone
  - Comment deadline August 21

APPENDIX B: INTERVIEW GUIDE

MA20C05-G-STBE: "Power Users" Interview Guide

May 2020

PURPOSE

The overarching objective of these interviews is to identify tool and calculation revisions that will improve the accuracy of the results, using reasonably available data by requesting program administrator (PA) stakeholder input on usability and potential improvements to the current calculator and analysis methodology for heating plant annual efficiency.

Table 6: Research Objectives Mapped to Questions in This Instrument

Research Objectives	Survey Questions Address the Objectives
Improve the steam trap tool: Solicit feedback on the current steam trap tool including usability and potential algorithmic improvements.	See questions in "Steam Trap Tool" section below.
Best practices heating plant annual efficiency: Solicit details on feasible methods to calculate boiler plant annual efficiency.	See questions in "Heating Plant Annual Efficiency Calculations" section below.

TARGET POPULATION

Component	Survey Questions Address the Objectives
Improve the steam trap tool:	Power users of the current steam trap tool that can comment on the tool estimation methodology, the ease of use, the pick-list applicability and/or other aspects of the tool. Nominees include (* targeted for both components)
Best practices heating plant annual efficiency:	Knowledgeable individuals with a technical background in boiler plant operations, combustion efficiency, annualized plant efficiency and estimating heating fuel impacts.

ANALYSIS PLAN

Table 2 describes the research objectives and the associated analysis approaches as drawn up in the workplan. The interviews will be broken into three segments to address specific study elements: steam trap project tracking, steam trap tool improvements and heating plant annual efficiency factor methods.

Table 2: Analysis Steps

Research Objectives	Qs Examined	Analysis Description
Determine the appropriate ex-post evaluation treatment of steam trap measures.	4	-Consider feasibility of alternate sampling strategies, accounting for any tracking limitations.

Identify and implement tool revisions that will improve the accuracy of the results, using reasonably available data.	12 (8 MA-based, 4 non-MA)	-Engage PA engineering staff on issues of tool accuracy -Request input on tool usability and potential improvements to the current calculator -Conduct best practice research in estimating steam trap savings.
Determine best practice method to account for heating plant annual efficiency factor	4	-Document implementer concerns and technical arguments -Conduct best practice research for applying heating plant efficiency to boiler load reducing measures

INSTRUMENT

Hi, my name is [   ] from ERS, we are a 3rd party evaluator and I'm calling on behalf of [utility]. You have been nominated by the [utility] to complete our phone survey regarding the steam trap tool and heating plant efficiency calculation methodology. We will be speaking with people who have experience with these tools to determine usability and identify potential improvements to the current calculator and analysis methodology for heating plant annual efficiency. If you are the person in your organization most familiar with using the too and have some availability and are willing to participate, I would appreciate your input on this topic.

F1. Do you have approximately [10-15 minutes] to help complete a survey?

01 [Yes]	Continue to Steam Trap Tracking
02 [No]	F2
97 [Don't know]	F2
98 [Refused]	F3

F2. Are you available later this week to complete the survey?

01 [Yes]	F3
02 [No]	F4
97 [Don't know]	F3
98 [Refused]	F4

F3. Thank you, [propose availability for later in the week or next], and what number will be best to contact you? And may I get an email address so that I can follow up with our time as well?

1. Date & time:
2. Phone number:
3. Email address:

[Proceed to F4]

F4. Thank you for your time, have a good day and stay safe.

***Important: continue only if following conditions are met:**

Time and duration of survey is convenient given their current business status and needs	Yes
Respondent expresses that phone data collection is best mode for survey.	Yes
Respondent has not conveyed a lack of time, impatience, need to prioritize other business critical activities or safety responsibilities.	Yes

Steam Trap Tracking

All questions are for MA PA staff only.

T1. How are steam trap measures recorded/monitored within their project tracking system? Are specific measure code identifiers used?

T2. Do steam trap entries have a designated identifier shared between all projects of the same measure type?

T3. Are steam trap projects identified differently if they are one of many measures included in a project, as opposed to an individual measure submission?

T4. Do all entries use the same measure life?

Steam Trap Tool

MA-Based Interviewees

Introduction – In Massachusetts, steam trap savings are calculated using a custom tool using input parameters collected on-site during a steam trap survey. This tool was adopted by PAs across the state following a 2016 study which made recommendations intended to reduce the variability and increase the accuracy of steam trap savings estimates among the different PAs and vendors performing implementation work. I have some questions about that tool. Are you familiar with it? First, we'll ask some questions about your experience with this tool and other potential methods used to generate trap estimates. Then we will ask you about any recommended improvements.

ST1. Please describe how you have used or interacted with the current version of the tool. *[Probe for how recently used, how many times used and for what purpose]*

ST2. Do you find that the tool is difficult to use? *[Probe: What aspects are difficult?]*

ST3. Is there something specific that you would change about the current tool?

ST4. We are finding that steam trap hours of operation are one of the most commonly corrected parameters during evaluation. Are there aspects of the tool or how users populate field data into it that you think might contribute to this issue?

ST5. Do you believe the results of the tool are accurate? *[Probe: Why or why not?]*

ST6. Do you currently or have you in the past used alternate methods for calculating steam trap savings? (This could include previous version of the tool)

ST7. Do you know of a different calculator or custom savings approach used outside of MA?

ST8. Do you have any additional comments on the tool that have not been addressed?

Non-MA Interviewees

Introduction – We understand you have experience working on efficiency programs featuring steam trap offerings. We are doing research on best practices and we hope you can share with us how you estimate savings. For some background, in Massachusetts, steam trap savings are calculated using a custom tool using input parameters collected on-site during a steam trap survey. This tool was adopted by PAs across the state following a 2016 study which made recommendations intended to reduce the variability and increase the accuracy of steam trap savings estimates among the different PAs and vendors performing implementation work.

ST9. Please describe your experience with steam traps and estimating energy savings associated with their repair and replacement.

ST10. Do you use a designated method or calculator to estimate steam trap savings? *[Probe for details on method, ask for documentation of approach. Probe: Do you ever check or calibrate savings estimates with utility billing data?]*

ST11. Do you use or are you familiar with any distinct or novel methods to translate on-site observations into energy savings calculation parameters? (i.e. hours of operation, operating status, pressure measurements into pressure)

ST12. Are there any other best practices you could share with us for the purposes of our research?

Heating Plant Annual Efficiency Factor

Introduction - We are researching best practice methods for converting thermal load reductions from measures like pipe insulation or steam trap repair and replacement into annual gas savings. This factor, which is used to convert the thermal load reduction from implementing an efficiency measure into the reduction of natural gas required to fire the boiler (at an equivalent load), is being called the 'heating plant annual efficiency factor'. This factor ideally accounts for the boiler annual average combustion efficiency and potentially, distribution losses, if appropriate. *[All questions will be asked for each selected interviewee.]*

HP1. Have you estimated savings from thermal load reduction measures (like pipe insulation) and if so, how have you estimated the factor to convert annualized thermal savings to annual gas saved?

a) Probe – Do you use a fixed assumption, like 75% or some similar value?

b) Probe – Do you have a reference document for your methods that you might share?

c) Do you rely on measurements such as combustion testing?

HP2. One or more of the following parameters have been nominated as a possible basis for calculating the heating plant annual efficiency factor. We are interested in your opinion about why or why not that factor would be useful and/or practical to include.

a) Boiler plant thermal efficiency rating

b) Boiler plant combustion efficiency test results

c) Estimated hours of boiler operation

d) Estimates of plant distribution losses and boiler jacket and standby losses

e) Other

f) Better to just assume a standard value

HP3. What is your opinion of best practices for estimating this factor?

HP4. Do you have any additional comments on this topic that have not been addressed?

APPENDIX C: INTERVIEW FINDINGS

This Appendix includes additional details on the interview task including interviewee details, general interview findings and the interview guide used for questioning.

Organization	Interview Subject		
	Steam Traps	Annual Heating Plant Efficiency	Tracking System
Columbia Gas			X
Eversource			X
National Grid			X
Berkshire Gas			X
Unitil			X
Liberty			X
Eversource	X	X	
National Grid	X	X	
CBLU	X	X	
EEAC	X	X	
Rise Engineering	X	X	
American Plant Maintenance	X		
Steam Trap Systems	X		
Gestra USA	X		
Enbridge	X	X	
So Cal Gas	X	X	

Findings

The Team conducted a total of 10 interviews with a variety of personnel including PA representatives, MA steam trap vendors and various industry experts (including both MA and non-MA based respondents). Questions related to the current iteration of the STT were directed at personnel who have regularly interacted with the tool (7 respondents) while questions related to other methods for estimating savings and general best practices were directed at all 10 interviewees. A summary of the high-level findings by conversation topic can be found below.

Tool Usability & Accuracy

- All respondents indicated that the current iteration of the tool is user friendly and straightforward enough to use.
- All respondents indicated that the accuracy of results from the tool are subject to the uncertainty of data inputs translated from field observations.
- 4 out of 7 respondents indicated there was some confusion over the interpretation of specific inputs (i.e. hours of operation, boiler efficiency, condensate return factor).
    - Clarification is needed on what specific values are representative of

- No respondents indicated a belief that the tool was being manipulated, although one did offer further explanation detailing how three main inputs (leak factor, hours and orifice size) can all be potentially adjusted to meet the required payback period.
    - *"The leak factor comes directly from field measurements where there needs to be a level of trust between the vendor and PA. The tool should provide more guidance on hours to reduce interpretation issues and the orifice size should be automatically generated based on the model type listed in the inventory."*

Application of Hours of Operation

- While the application of hours was generally agreed to be common sense, most respondents indicated the need for additional instruction and guidance on how to apply them properly.
    - Some asked for clarification on what specific hourly values are representative of.
    - Most indicated a need for prepopulated dropdown values or QC checks based on specific parameters such as facility type, trap application and boiler controls.
- Interviewees representing all PAs believe that the hours discrepancy from the most recent evaluation was predominantly due to QC review issues.

Other Methods for Estimation

- Respondents indicated that the Napier approach for calculating steam flow is the most commonly used estimation method outside of MA.
    - Used by multiple vendors, DOE, NY TRM, Con Ed, SCG, Enbridge.
    - Differences between the Grashof and Napier approaches are minimal.
- While respondents noted that vendor-sourced approaches often yield higher savings than the MA approach, this is predominantly due to the fact that the MA method uses empirical data specific to its population and likely isn't representative of larger facilities (oil refineries, heavy manufacturing) that nationwide vendors design approaches for.
- While there was a consensus that all methods used to calculate steam trap savings are subject to uncertainty, some respondents highlighted the importance of sticking to one method and making further refinements using available information.

General Feedback

- Regular tool users suggested the need for additional clarification and guidance on interpretation of specific parameters.
- Not all PAs have access to the unlocked version of the tool.
- The efficiency input can only be updated using an unlocked version of the tool
- There is a need for better exchange of information between the implementation, post-inspection & evaluation teams. Respondents have indicated the need to set consistent methods for determining and verifying assumptions and field observations as well as a mechanism to provide feedback when values are updated by post-inspectors or evaluators.

Best Practices

- A common trend among established trap programs is a library of steam trap models containing manufacturer specifications including suitable pressure ranges and trap orifice size. When users are populating the trap inventory into the tool, the orifice size populates automatically based on the trap model number.
- Some established tools also distinguish each trap as being located on a main header (line) or specific application (load) to further support the review process
- Some established tools also have more intricate failure modes that account for things other that steam blowing through like rapid cycling, back pressure presence or the failure of thermostatic elements on float & thermostatic (F&T) traps.

Heating Plant

- Combustion efficiency measurements thought limited in the sense of it being just a snapshot of boiler operation, are a good standalone data point to collect when on site.
    - Interview respondents were split between shaving off a fraction of boiler efficiency for other losses or using the value as is.
    - Other respondents indicated that it is the true incremental value for load reducing measures.
- The consensus among respondents was that in order to capture the true system efficiency for converting thermal load reductions into gas savings, there are a number of things to be considered.
    - Metering gas usage and steam produced is the most direct means of capturing this value
    - Boiler operation should be considered, including the types of controls on the burner as well as the parameter the boiler is trying to maintain (temperature or pressure)
- Respondents indicated that using a fixed value is the most reliable method.
- Other jurisdictions have line level efficiency inputs for each trap for instances when there are multiple boilers serving the steam system.

Recommended Improvements

Based on the Team's findings from the interviews and best practice research, they have summarized proposed improvements to the steam trap tool which are further discussed in the Conclusions and Recommendations sections.

Bolster the review process

- Identify and provide counts of inspected, fixed and reviewed traps by adding columns to the inventory to designate as such and include totals fields in the summary area to support the review process.
- Add a tab for including billing data and utility account info for comparison of facility energy usage against savings

- Include additional documentation of previous work completed (historical survey records, inventories, O&M practices) to support review process

Enhance recordkeeping of revisions

- Add fields for keeping track of which parameters are checked and/or verified on site. Potentially duplicate tabs or color-code different values to visualize changes, support review process.
- Include a post-inspection review page to highlight any updates made and allow for easier review of the completed work.
- Add fields for documenting critical measurements like the inlet and outlet temperatures of each reviewed trap along with the condensate tank temperature before and after work is completed.

Ensure accessibility for and consistency among users

- Make the tool accessible via the web and provide all PAs with an unlocked version of the tool.
- Highlight/communicate version changes for user-awareness
- Provide clear guidance on application of parameters including hours of operation, condensate return factor and leak factors.

Make tool more tamper resistant with more restrictive inputs

- Limit inputs to specific account classes or trap application by making drop down values a function of other required fields.
- Include totals fields showing the number of traps in each failure mode and potentially cap the number of allowable instances for certain failure modes.
- Develop a steam trap model library and use an automatic lookup function to produce orifice size based on the entered model number.

Further algorithmic refinements

- Use weather normalized utility billing analysis results to calibrate the most uncertain parameters. Potentially incorporate results from previous evaluation to bolster calibration.
- Investigate the impacts of diminishing savings due to presence of back pressure within the condensate lines and potentially apply to savings algorithm

APPENDIX D: HEATING PLANT MEMO

Memo to:
Massachusetts Program Administrators Research Team and Energy Efficiency Advisory Council EM&V Consultants

From: Sorin Ioan and Jon Maxwell, ERS
Contributor: Jeremy Blanchard, GDS Associates
Date: October 02, 2019

Copied to: MA19C05-G-CUSTGAS Evaluation Team

Pipes and Fittings Insulation Measure – Impacts Calculation Methodology

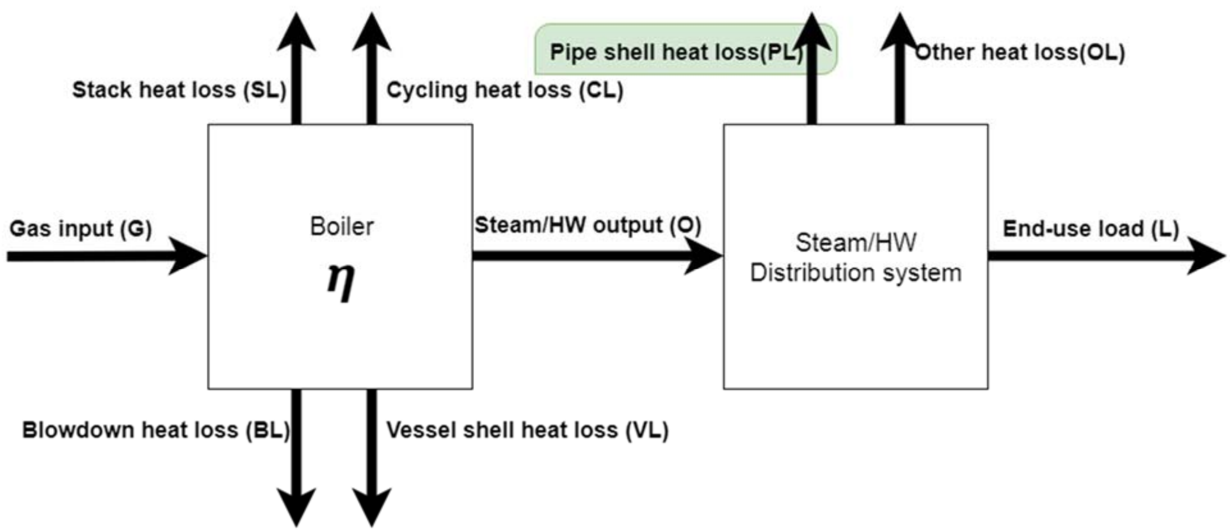
This memo provides details on the methods the evaluators use to calculate the impacts of installing insulation on pipes and fittings that are part of a steam or hot water system served by natural gas boilers.

1 MODELING PIPES AND FITTINGS INSULATION IMPACTS

This section provides details on the heat loses included in the calculations used to determine the impacts of adding the insulation.

The evaluation team modeled the heating system to identify the variables that contribute to the energy balance. The simplified heating system is presented in Figure 1-1.

Figure 1-1. Heating System - Simplified Model



Legend:

- G – gas input
- SL – heat loss through the stack of the boiler
- CL - cycling heat loss
- BL – heat loss due to blowdown (due to the purge of the boiler)
- VL - the heat loss through the shell of the boiler vessel
- O - boiler output load

PL - heat loss through the surface of the pipes

OL - other heat losses through the distribution system

L - load the system supplies to the end-use

Installing insulation on pipes/fixtures does not affect the end use load (L) or other distribution system heat losses (OL). The only impact on the boiler heating energy output (O) is due to the decrease in pipes/fixtures heat loss (PL) through their shell.

1.1 Calculations

This section provides details on the formulas the evaluation team used to calculate the following:

- Baseline boiler(s) output
- As-built boiler(s) output
- Pipes/fittings insulation impact on the boiler(s) heating energy output
- Natural gas savings impacts due to the change in the boiler(s) heating energy output

1.1.1 Baseline Boiler(s) Output:

Formula 1 below describes the baseline energy balance downstream of the boiler:

$$O_{base} = PL_{base} + OL_{base} + L_{base} \tag{1}$$

1.1.2 As-built Boiler(s) Output:

Formula 2 below describes the energy balance downstream of the boiler after the installation of the measure:

$$O_{ab} = PL_{ab} + OL_{ab} + L_{ab} \tag{2}$$

1.1.3 Impact on the Boiler(s) Output:

Formula 3 below calculates the measure impacts on the boiler(s) heating energy output by subtracting the as-built boiler(s) heating energy output from the baseline boiler(s) heating energy output.

$$O_{base} - O_{ab} = PL_{base} + OL_{base} + L_{base} - PL_{ab} + OL_{ab} + L_{ab} \tag{3}$$

where,

$PL_{base} > PL_{ab}$
$OL_{base} = OL_{ab}$
$L_{base} = L_{ab}$

Two variables impact the heating energy output of the boiler(s): baseline and as-built pipes/fittings heat loss values. All the other heating losses remain the same and do not have an impact on the result.

Formula 3 reduces to:

$$O_{base} - O_{ab} = PL_{base} - PL_{ab} = IS \tag{4}$$

where,

IS – measure impact on the boiler(s) heating energy output

1.1.4 Convert the measure savings into natural gas consumption

This section provides details on the method used to convert measure impact on the boiler(s) heating energy output to natural gas.

1.1.4.1 Hypothesis

Boiler losses that change with boiler(s) heating energy output should be included in the efficiency used to calculate the natural gas impact. Losses that remain constant independent of load should be excluded.

1.1.4.1.1 Variable heat loses

Stack heat losses (SL) change approximately in proportion with boiler(s) heating energy output and therefore should be included in the boiler efficiency used to calculate insulation impact. The same efficiency value can be used for both the baseline and measure condition[4]. To determine stack losses, either:

- Spot-measure the combustion efficiency (η_c) of the boiler that supplies heating energy to the distribution system on which the measure was installed. Attempt to take spot-measurements at various firing rates (e.g., 20%, 40%, 60%, 80%, 100%). Average the values to calculate the average combustion efficiency ($a\eta_c$) that will be used to calculate natural gas impacts.
- If combustion efficiency measurement only possible at one single firing rate, adjust the combustion efficiency using the values provided in Table 1-1 below. The adjustment factor (af) is a function of the firing rate at which η_c was spot-measured and depends on the boiler controls.

Table 1-1. Boiler Efficiency Adjustment Factors[5]

Measurement Firing Rate	Linkage	Parallel Positioning	Parallel Positioning and O_2 Trim
25%	0.6%	0.3%	-0.1%
50%	-0.3%	-0.1%	-0.1%
75%	-0.5%	-0.1%	0.2%
100%	-0.9%	-0.4%	0.2%

The adjustment factor (af) will be added to the spot-measured efficiency to calculate the average combustion efficiency ($a\eta_c$) that will be used to calculate natural gas impacts.

[4] Change in efficiency percentage due to stack loss is not worth accounting for pipes/fittings insulation measures. Combustion losses tend to decrease very slightly less than proportionally to load decreases for older less efficient boilers. For example, if a conventional boiler's combustion efficiency decreases by 5% between 80% and 20% load due to imperfect linkage controls (partially offset by increased efficiency due to greater heat exchange coil area per Btu transferred), the net change would be less than a 0.1% decrease in combustion and heat exchange efficiency for each 1% decrease in load caused by the measure. With a typical insulation measure affecting far less than 5% of the rated boiler capacity this difference is not enough to account for in the pipes/fittings insulation measure impacts calculations.

[5] The values in Tables 1-1 and 1-2 have been provided by GDS Associates, Inc and they are derived based on engineering judgement.

Example for using an adjustment factor:

- Boiler with linkage controls and η_c = 82% was spot-measured at 30% firing rate
- The adjustment factor is:

$$af = \frac{(-0.3\% - 0.6\%)}{(50\% - 25\%)} \times (30\% - 25\%) = \frac{-0.9\%}{25\%} \times 5\% = \frac{-0.9\%}{5} = -0.18\%$$

- The efficiency value that will be used in the calculation of natural gas impact is:

$$a\eta_c = af + \eta_c = -0.18\% + 82\% = 81.82\%$$

- If spot-measuring combustion efficiency (ηc) is not possible, use the $a\eta_c$ value provided in Table 1-2 below.

Table 1-2. Average Boiler Efficiency[6]

Linkage	Parallel Positioning	Parallel Positioning and O_2 Trim
83.2%	84.2%	85%

1.1.4.2 Natural Gas Impact

The impacts the measure has on the natural gas consumption will be calculated using the following formula:

$$Gas = \frac{IS}{a\eta_c} \quad (5)$$

where,

Gas – measure impact on the boiler(s) natural gas consumption
IS – measure impact on the boiler(s) heating energy output calculated using Formula 4 above
$a\eta_c$ – boiler(s) average combustion efficiency calculated using one of the three scenarios (a, b, or c) described above.

[6] The values in Tables 1-1 and 1-2 have been provided by GDS Associates, Inc and they are derived based on engineering judgement.

About DNV GL

Driven by our purpose of safeguarding life, property and the environment, DNV GL enables organizations to advance the safety and sustainability of their business. We provide classification and technical assurance along with software and independent expert advisory services to the maritime, oil and gas, and energy industries. We also provide certification services to customers across a wide range of industries. Operating in more than 100 countries, our 16,000 professionals are dedicated to helping our customers make the world safer, smarter, and greener.