Impact Evaluation of Commercial Strategic Energy Management *Final Report*

Energy Trust of Oregon

Date: October 20, 2016

(Intentional Blank Page)

.

Table of contents

List of figures

List of tables

1 EXECUTIVE SUMMARY

This report presents DNV GL's impact evaluation of Energy Trust of Oregon's Commercial Strategic Energy Management (SEM) offering, which is part of the Existing Buildings (EB) program. The impact evaluation is specific to the energy savings achieved by participants in 2013, 2014, and 2015. The evaluation concludes that over these three years the program achieved 51,541 MWh and 984,063 therms of energy efficiency savings by enabling and supporting improved energy management practices in commercial buildings. This report provides multiple performance metrics for the Energy Trust to consider for future program reporting.

1.1 Program and Evaluation Overview

Commercial SEM at Energy Trust is a program offering designed to deliver comprehensive energy services to large commercial customers, and focuses on behavioral and operational changes, as well as identifying capital projects. Energy consulting firms, known as Program Delivery Contractors (PDCs), hold workshops, help identify energy savings opportunities and provide training and technical support to participants over the course of a year. Energy Trust claims savings based on a top-down analysis of building-level energy use and pays performance-based incentives to participants.

Energy Trust enrolled its first group of participants in Commercial SEM in the fall of 2011. Now in its fifth year, Commercial SEM has evolved into a large portion of Energy Trust's EB program, with significant annual gas and electric savings. In 2014, SEM comprised 27% of the EB program's gas savings and 13% of its electric savings.

The primary results of this evaluation are the determination of achieved energy savings and associated realization rates. In addition, the evaluation provides recommendations to improve the accuracy of energy savings estimates and feedback from participants on how the SEM is impacting their energy consumption. The activities completed during this evaluation include interviews with program staff, PDCs, and participants; visits to participant locations; and analysis of participant consumption. The evaluation estimated overall savings achievements by extrapolating results from an analyzed stratified random sample.

1.2 Summary of Results

Table 1 shows the evaluated energy savings achieved for calendar years 2013, 2014, and 2015. These savings are net of capital project savings reported through other Energy Trust programs. The table compares the evaluated energy savings to the savings claimed to be occurring during the calendar year, irrespective of the program year of savings acquisition. The table shows that the accuracy of savings estimation has improved year over for both gas and electric fuels. Evaluators believe this is due to both improvements in savings estimation and the increase in the program population. Section **Error! Reference source not found.** contains further details about these results.

Table 1: Cumulative Energy Savings, 2013-2015, Claimed and Evaluated, by fuel and cohort

Resource Fuel	Savings Year	Claimed Savings	Evaluated Savings	Realization Rate	Baseline Consumption	% of Baseline Consumption
Electric (kWh)	2013	5,299,318	7,350,568	139%	175,140,706	4.2%
	2014	14,024,257	16,338,244	116%	318,777,912	5.1%
	2015	26,959,489	27,852,207	103%	394,271,281	7.1%
Gas (therms)	2013	126,942	-18,452	-15%	4,001,720	-0.5%
	2014	496,277	155,938	31%	7,340,910	2.1%
	2015	926,966	846,577	91%	13,051,759	6.5%

Table 2 compares the claimed and evaluated first year savings by program year. The total first-year savings in this table equals the 2015 calendar year program savings since the savings acquired in PY2012 and PY2013 still occur at the meter in calendar year 2015. The evaluation believes there are multiple reasons for the variance between claimed and evaluated first-year savings. One reason is the different time periods used to estimate these values and the inherent challenge faced by the program in estimating 12-months of future SEM savings based on a partial year of facility consumption.

Table 2: Incremental savings acquired in each program year

Resource Fuel	Program Year	Claimed Savings	Evaluated Savings	Realization Rate
Electric (kWh)	PY2012	5,299,318	7,350,568	139%
	PY2013	8,724,939	8,987,675	103%
	PY2014	12,935,232	11,513,963	89%
	Total	**26,959,489**	**27,852,207**	**103%**
Gas (therms)	PY2012	126,942	-18,452	-15%
	PY2013	369,335	174,390	47%
	PY2014	430,689	690,639	160%
	Total	**926,966**	**846,577**	**91%**

1.3 Recommendations

This evaluation produced the following recommendations. These recommendations are a direct result of the activities completed as part of this impact evaluation.

- DNV GL recommends changing the modeling methodology used for SEM. The majority of program savings reviewed were determined using models based on the average temperature during the program period. DNV GL recommends utilizing degree-day estimates as independent variables: heating, cooling, or both. DNV GL also recommends avoiding the use of a polynomial term in any regressions; use of linear degree-day estimates more closely aligns with the inherent energy consumption processes in commercial buildings and provides results that are more intuitive. Use of a polynomial regression term may provide a better fit to the data, but is risky when extreme conditions occur. However, the evaluation results show that the methodology used by the program did reasonably estimate savings for the program has a whole so this change is not critical. DNV GL shared this recommendation with Energy Trust during the evaluation, and Energy Trust is currently working with a separate contractor to develop a standard savings estimation methodology covering all of its SEM programs. DNV GL believes the results of this impact evaluation will influence the proposed standard estimation methodology.
- DNV GL recommends changing the savings calculation method. Program practice is to forecast the energy savings that will be achieved in the next calendar year, based on consumption during the current year. This forecast has typically occurred in the fall, with the first calculation occurring at the end of the initial engagement year. This initial forecast often relied on trends in consumption over only a few months. DNV GL recommends that Energy Trust consider calculating savings for the program year at the end of each program year, based on the consumption and weather that occurred within that year. This calculation would take about the same time as the current process, but be completed a year later. The benefit to this methodology is that it reduces and, in many cases, eliminates the risk of over- or underestimating savings. The challenge to this methodology is that the program must operate longer

without claiming a full calendar-year savings for a site. However, DNV GL recommends that the program also calculate achieved savings at the end of the initial engagement year. DNV GL observed many cases in which savings began occurring during the initial engagement year, but these savings went unclaimed.

- DNV GL recommends enhancing program tracking records. The program data we received did not include any identifiers – other than a participant name and site name – to connect SEM participants with participation records in Energy Trust's traditional energy efficiency programs. As a result, identifying capital projects that occurred at participant locations required substantial evaluation effort. Given the relative size of the SEM program, DNV GL recommends updating the program tracking records to include a site identifier that aligns with other program databases. This will reduce the risk that the program and future evaluations incorrectly account for savings occurring at each site.
- DNV GL recommends enhancing program tracking data. Utility meter data supporting the top down estimation of savings for each site was stored in unique Microsoft Excel files for each site. However, program tracking data did not include a comprehensive list of the utility accounts and meters associated with program sites. As a result, compiling utility meter data and associated account numbers required substantial additional evaluation efforts. DNV GL recommends that Energy Trust consider updating the program tracking data to include the account and meter numbers associated with each participating site. The program also should validate these numbers against the utility data received by Energy Trust before claiming program savings. This change will reduce the evaluation burden and risk on the SEM program.

Energy Trust of Oregon

MEMO

Date: February 28, 2017
To: Board of Directors
From: Kathleen Belkhayat, Commercial Sector Program Manager
Dan Rubado, Evaluation Project Manager
Subject: Staff Response to 2012-2014 Commercial SEM Impact Evaluation

The impact evaluation of Energy Trust's Commercial Strategic Energy Management program found that from 2012-2014, the program realized 104 percent of its claimed electric savings and 91 percent of its claimed gas savings. Although the realization rates varied over time as the program changed and refined its energy modelling and forecasting methods, the overall program realization rates are very good. Of note, gas realization rates were very low in the first two years of the program, but appear to have improved over time. The program has directly addressed several of the factors that likely contributed to these low initial realization rates. On average, participating SEM sites achieved evaluated electric savings of 7.1 percent and gas savings of 6.5 percent. In general, participants were highly engaged and continued to hold energy team meetings, track energy use, and prioritize energy projects after the initial SEM engagement ended. This evaluation was not able to draw any conclusions about the persistence of savings.

The evaluator provided valuable feedback on the program's modelling and forecasting methods. These recommendations influenced and enhanced the SEM modelling guidelines that Energy Trust was developing at the same time. These guidelines have now been finalized and are being implemented with 2017 SEM participants. They include the evaluator's recommendations for establishing the baseline period and selecting model parameters. Energy Trust significantly changed the way the program forecasts and claims savings in 2016 and this was strongly reinforced as a good decision in the evaluator's recommendations.

The evaluator noted that some SEM sites were not individually recorded in Energy Trust's project tracking database. The evaluator recommended recording each individual site and its identifier in the database to improve project tracking and to simplify the process of connecting SEM sites to capital efficiency projects. This change was made prior to the evaluation, but the decision to do so was reinforced by the evaluator.

Commercial SEM is becoming a mature program in Oregon with 8 cohorts completed to date, including about 60 participants and around 500 buildings. Two more cohorts are getting started in 2017. One of the largest administrative changes to the program is the Existing Buildings PMC taking over management of SEM and the implementation contractors in 2017. In addition, the SEM continuation cohorts will be combined by geographic region, and may include first year participants. This will provide some delivery efficiencies, make it easier for participants to connect, and provide them with more customized topic areas. The next SEM impact evaluation will be conducted as part of the Existing Buildings program impact evaluation. A single evaluation will allow Energy Trust to better analyze the interactions between capital project and SEM energy savings.

2 INTRODUCTION

Energy Trust performs evaluations of its programs on a regular basis. DNV GL was selected to conduct an impact evaluation of Energy Trust's Commercial SEM offering. This program offering is designed to deliver comprehensive energy services to large commercial customers, and is focused on behavioral and operational changes, but also supports capital projects. This evaluation covers program years 2012 through 2014.

2.1 Study Objectives

The impact evaluation of the SEM program had the following objectives:

- Develop reliable estimates of actual gas and electric savings realized by SEM participants over several years. This information will be incorporated into Energy Trust's annual true-up of program savings and used for program budgeting and forecasting.
- Account for capital project savings at SEM participant sites and evaluate them separately, if necessary.
- Report observations and make recommendations to improve the way SEM energy savings are estimated.
- Determine how engaged participants are and how well they have implemented and maintained changes in their organizations and buildings over time.

2.2 Background

Energy Trust performs process and impact evaluations on all of its programs on a regular basis. Although Commercial SEM is a component of Energy Trust's Existing Buildings (EB) program, it is structured and implemented separately and differently from the rest of the overarching program. Commercial SEM is also relatively new compared to the rest of the EB program. Therefore, Energy Trust chose to evaluate SEM separately. Past Commercial SEM evaluation efforts have been primarily process evaluations, with some engineering review of the methods used for facility regression modeling and for calculating energy savings. This impact evaluation is the first investigation of the energy savings achieved by the Energy Trust's Commercial SEM participants.

2.3 Program Description

Energy Trust enrolled the first participants in Commercial SEM in the fall of 2011. Commercial SEM is now in its fifth year and has evolved into a large portion of Energy Trust's EB program savings, with significant annual gas and electric savings. In 2014, SEM made up 27% of the EB program's claimed gas savings and 13% of its claimed electric savings. Delivery of SEM has evolved over time as well, with an increasing focus on providing support to participants on an ongoing basis and recruiting organizations that operate large portfolios of commercial buildings.

To acquire energy savings, the program provides support and resources to large commercial customers so they can build a formalized energy program and capabilities in their organizations. The objectives include not only direct energy savings from operational changes, but also indirect savings through changes in the energy management culture of the organization. Key elements of the program include:

- **Employee engagement:** Participants form an energy team made up of representatives throughout the organization. Within each team, an energy champion is identified. The energy team completes activities throughout each year to engage employees in the program and energy savings.

- **Executive support:** Each participant must identify an executive sponsor who provides support and resources to the energy team and holds them accountable for reaching their objectives.
- **Energy use tracking:** Tracking and monitoring energy use allows participants to see how their activities are affecting their energy consumption.

Commercial SEM requires an initial one-year commitment from participants. Over this time, the participants are guided through a structured process to build the awareness of energy usage at their facilities, identify energy-saving opportunities, set up a system for monitoring and tracking their energy use and savings over time, and engage employees and building occupants in the efforts to save energy. Most participating customers operate a portfolio of buildings and select a subset to engage.

Program delivery contractors (PDCs) implement SEM.[1] The PDCs are energy experts who run the participants through the SEM curriculum and provide a wide range of ad hoc support. The PDCs are responsible for training participant staff, providing technical assistance to find energy waste, and assisting participants as they develop and implement an energy management plan. The PDCs also set up a tracking tool for each participant and site (typically spreadsheet-based) for participants to record changes in their energy consumption. Most commonly, the PDCs developed a monitoring, targeting and reporting spreadsheet (MT&R). Section 2.3.2 provides more information on this process.

After their initial year of participation, participants can opt to participate in SEM Continuation, in which the PDCs provide support to the energy teams on an ongoing basis. During Continuation, participants may elect to have new sites from their organization participate.

Participants are eligible to receive an incentive based on the estimated energy savings achieved during their first year of participation. The incentive is provided for the Operations and Maintenance (O&M) energy savings only; savings from any capital projects is removed from the estimated energy savings, though participants may receive incentives for these through other Energy Trust programs. Each site’s opportunity register documents both the identified energy-saving opportunities and when any SEM actions occurred.

2.3.1 Delivery Approaches

The program used two different approaches to implementation during the study period: the cohort approach and the corporate approach.

Cohort: Through the cohort approach, the PDC delivers the program services to participants in a group environment. The activities consist of a combination of group workshops and one-on-one onsite meetings. The benefits of the cohort approach are that participants are able to share their experiences and learn from the other members’ activities. Strategic Energy Group (SEG) was selected as the PDC to implement the cohort approach.

The majority of SEM participants opted for the cohort approach. Four cohorts were formed from 2012-2014. The cohort participants included organizations that operated office buildings, colleges, correctional facilities, hospitals, and other business and government facilities. Each cohort has a unique engagement timeline. Table 3 lists dates specific to the engagement year for each cohort. The program continues to deliver the cohort approach at present.

---

[1] Two PDCs delivered SEM in parallel during the years included in the evaluation. Only one PDC is currently delivering commercial SEM.

Table 3: Engagement year timelines, by Cohort

SEM Activity	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Kickoff Workshop	Nov - 2011	Jan - 2013	Oct - 2013	Jan - 2014
Energy Accounting & Benchmarking	Dec - 2011	Feb - 2013	Nov - 2013	Mar - 2014
Effective Energy Team Workshop	Apr - 2012	Mar - 2013	Jan - 2014	Apr - 2014
Performance Tracking & Reporting Workshop	Feb - 2012	Apr - 2013	Feb - 2014	May - 2014
Energy Analysis & Audits Workshop	Jan - 2012	Jun - 2013	Apr - 2014	Jun - 2014
Employee & Occupant Engagement Workshop	Jun - 2012	Jul - 2013	May - 2014	Sep - 2014
Energy Management Planning Workshop	Jul - 2012	Sep - 2013	Jul - 2014	Oct - 2014
Report Out	Nov - 2012	Nov - 2013	Oct - 2014	Dec - 2014

Corporate: Through the corporate approach, participants received services on an individual basis. The benefit of the corporate approach was that participants had the flexibility to move through the process at their own pace. Ecova was the PDC selected to deliver the corporate approach. Corporate participants included colleges and restaurant chains.

Energy Trust no longer implements the corporate approach at this time. Energy Trust stopped implementing the program after determining that the cost to acquire energy savings using this approach was significantly higher than using the cohort approach. Energy Trust believes this higher acquisition cost was due to reduced recruitment success, reduced savings identification per site, the smaller size of the sites, and the lack of a group setting with a peer-to-peer element.

Even though the corporate sites did not work together, they are referred to as a fifth "corporate" cohort in this report for simplicity. Each participant followed its own unique engagement timeline. These timelines started between February 2012 and February 2013.

2.3.2 Energy Use Monitoring and Tracking

Participants used three different modeling methods to estimate energy savings over the analysis period. These building-level energy consumption models allow each facility to track its energy use over time, along with key independent factors that affect consumption. The facility can use this information to consider how the actions taken were affecting their energy use. The PDCs use the models to estimate energy savings and determine performance-based incentives to participants. Thus, the MT&R or other tracking tools served multiple purposes, providing updated savings information for the participant, the program, and the evaluation.

The PDC involved, the data available, and the participant's preference determined the methodology used for a given site. The following describes the three model types used by the program and reviewed by DNV GL:

- **MT&R** –MT&R-based models were used most often and they support the majority of savings claimed. Microsoft Excel was the modeling platform. Typically, these models used monthly consumption, weather and non-weather independent variables to develop a regression model from a baseline period. This model then predicts consumption in future periods. Site savings were determined from the difference between predicted and actual consumption during specified periods.
- **Energy Expert** – Energy Expert is a platform offered to Portland General Electric (PGE) customers. The platform is online and integrates with PGE's hourly consumption data for the facility. The software uses the baseline period to estimate consumption for each hour of the week across different temperature bins. For example, the model will estimate unique consumptions for Monday at 18:00, from 50-55°F and

Monday at 18:00, from 55-60°F. The PDC worked with the software vendor to ensure the baseline periods and savings calculations aligned with the cohort timelines.

- **Corporate SEM** – Ecova used the Georgia Tech EnPI tool for Corporate SEM sites, developed for Superior Energy Performance, to generate consumption models, and then summarized the modeled consumption in a single spreadsheet at all locations for one participant. The models typically used cooling degree days (CDD) to model electric consumption and heating degree days (HDD) to model gas consumption. Actual consumption was compared with modeled consumption to estimate program savings.

2.3.3 Claimed Program Achievements

Commercial SEM achieved savings over the analysis period at 167 unique electric participant sites and 126 unique gas participant sites, as shown in Table 4. The program assumes a measure life of three years for the annual savings at a facility. Table 4 shows only the incremental savings for a given year. These values are equal to the values recorded in the program tracking file provided to DNV GL. These values represent the initial savings claimed for a site and any adjustments made in subsequent years. These values are equivalent to the first year energy savings acquired during each program year and equal to the savings reported by Energy Trust.

Table 4: Claimed Incremental Energy Savings, PY2012-2014, by fuel and cohort

			Incremental Tracked Annual Savings		
Fuel	**Cohort**	**Site Count**	**PY2012**	**PY2013**	**PY2014**
Electric (kWh)	1	45	5,299,318	2,174,159	680,604
	2	38		6,090,749	2,432,492
	3	30			2,830,422
	4	23			6,761,472
	Corporate	31		460,031	230,242
	Electric Totals	**167**	**5,299,318**	**8,724,939**	**12,935,232**
Gas (therms)	1	30	126,942	119,518	120,380
	2	25		186,524	18,787
	3	22			178,543
	4	12			94,215
	Corporate	37		63,293	18,764
	Gas Totals	**126**	**126,942**	**369,335**	**430,689**

Table 5 shows the cumulative[2] savings, or total resource acquisition expected in the calendar years listed. These savings are equivalent to the energy savings reported to be occurring at the participant utility meters due to the SEM program. The Energy Trust assumes a three-year measure life, so any savings acquired in PY2012 and PY2013 are assumed to still be occurring in 2015. The impact evaluation models were directly compared to these values.

[2] The program tracking data only recorded the change in savings expected from one year to the next. This is termed the "incremental" savings. Therefore, the expected savings acquired from any site is equal to the sum of the savings listed in the tracking database for the year plus any years prior. This is referred to as "cumulative" savings in this evaluation report.

Table 5: Claimed Cumulative Energy Savings, 2013-2015, by fuel and cohort

			Cumulative Tracked Annual Savings		
Fuel	**Cohort**	**Site Count**	**2013**	**2014**	**2015**
Electric (kWh)	1	45	5,299,318	7,473,477	8,154,081
	2	38		6,090,749	8,523,241
	3	30			2,830,422
	4	23			6,761,472
	Corporate	31		460,031	690,273
	Electric Totals	**167**	**5,299,318**	**14,024,257**	**26,959,489**
Gas (therms)	1	30	126,942	246,460	366,840
	2	25		186,524	205,311
	3	22			178,543
	4	12			94,215
	Corporate	37		63,293	82,057
	Gas Totals	**126**	**126,942**	**496,277**	**926,966**

2.4 Report Organization

This report consists of five sections plus appendices, beginning with this introductory section, which provides background and program information. The rest of the report is organized as follows:

- **Section 3. Evaluation Activities** describes the methods and results for data collection, including the evaluation tasks, sample design, review of program data and program and participant data collection. This section includes the results of the interviews of program staff and participants and results of site visits.
- **Section** Error! Reference source not found.**. Gross Savings Methods** provides the methodologies used in the determination of evaluated savings.
- **Section** Error! Reference source not found.**. Evaluation Results** provides the site level and program level results of this impact evaluation.
- **Section 6. Conclusions and Recommendations** provides the meaning of the evaluation results and provides recommendations.
- **Appendices** provide the stratum specific results, interview guides for program staff and participants, as well as the sample and model diagnostics summary memoranda.

Table formatting

Evaluation of the Commercial SEM programs required the determination of both the total savings achieved at a location and the change in savings compared to the previous year. The total savings is what the analysis measures, but the change in savings is what the program reports. DNV GL uses specific table formatting throughout the report to help the reader understand what information is included.

Table 6: Table formatting key

Table Heading	Meaning
PY2012, PY2013, PY2014	Incremental savings **claimed** to be attributed to the program year shown. These are the savings reported by Energy Trust for the program year.
2013, 2014, 2015	Total energy savings **claimed** to be occurring in the calendar year shown. This value is the sum of incremental claims for all previous program years.
2013, 2014, 2015	**Evaluated** total energy savings for the calendar year shown.
PY2012, PY2013, PY2014	**Evaluated** incremental savings attributed to the program year shown.

3 EVALUATION ACTIVITIES

This section describes the activities completed as part of this impact evaluation including the findings that resulted from these activities. Section 4 discusses the savings estimation methods and results based on the data collected.

3.1 Evaluation Tasks

DNV GL completed the following tasks to achieve the goals of this evaluation. The subsections that follow provide details regarding these tasks, the data collected, and the associated results.

1. **Documentation Review:** DNV GL received and reviewed program documents and individual project data from Energy Trust.
2. **Site Sampling Plan:** DNV GL developed a sample design for the impact evaluation to achieve reliable estimates of actual gas and electric savings realized by SEM participants.
3. **Model Review:** DNV GL reviewed all sampled savings models to determine the validity of the methodology, identify which models required modification and/or a site visit, and summarize the data or information necessary to verify the impacts.
4. **Staff Interviews:** DNV GL conducted interviews with Energy Trust staff responsible for managing the PDCs.
5. **Participant Interviews:** DNV GL interviewed participants to: determine the status of measures and activities conducted or planned; determine the status of building changes or operations that may affect the model or energy savings; gather data to determine impacts; and recruit for site visits, where applicable.
6. **Site Visits:** DNV GL completed site visits for selected sites in order to collect additional data and information expected to improve the evaluation team's understanding of how energy is consumed and saved at the location.
7. **Impact Analysis:** DNV GL calculated the achieved savings for each sampled site as the difference between a modeled baseline consumption and actual consumption, net of claimed gross capital project savings. DNV GL used two different methods to estimate baseline consumption: the original program method (M1) and a linear degree-day method (M2). Final estimates of achieved program savings were determined based on the M2 linear degree-day methodology.

3.2 M&V sample design

SEM may target a site's electricity consumption, natural gas consumption, or both. In order to provide reliable estimates of achieved electricity and natural gas savings, DNV GL considered each fuel separately. The sampling unit and unit of analysis in this study was therefore not the site in general, but the fuel-specific savings at a site. The sample frame for this study contained 292 fuel-specific sites[3] associated with 27 distinct participating customers.

Stratified random sampling with a certainty stratum was implemented. Stratification is an important and commonly used design feature in data collection efforts. Stratification refers to the process of partitioning the sample frame into distinct groups (strata) and sampling is done independently within groups.

[3] One site was inadvertently dropped from the original population of 293 fuel-specific sites.

Stratification was used in this study to: 1) improve expected precision of the final estimates and 2) control the sample size by subgroups of interest during the analysis. Thirty unique strata were defined by the following:

- **Fuel Type**: Electric or Gas.
- **Certainty**: The sites with the largest energy savings within each participating organization were selected for the certainty stratum. A few electric and gas sites with the largest overall savings, regardless of participant, were also placed in this stratum. The certainty stratum ensured that at least one site per participant would be included in the study and that the largest sites were included. This balanced the objectives of studying all participants and minimizing the sampling error in the results.
- **Sampling Cohort**: The evaluation utilized four sampling cohorts, 2012, 2013, 2014, and Corporate SEM. Except for Corporate SEM participants, the sampling cohort was determined for all sites associated with a participant based on the first program year of SEM participation. As a result, program cohorts 3 and 4 were combined in to one 2014 sampling cohort.
- **Size**: This dimension is defined for those sites not in the certainty stratum (defined above). Size strata were created within each fuel type and cohort group by looking at the distribution of energy savings among all sites within the group. Those sites in the lower 25% were considered to have "Low" savings, those in the 25% to 75% group were considered "Medium," and those in the upper 25% were considered "Large." Size strata ensure an adequate number of low, medium and large sites (in terms of energy savings) in the respondent sample.

DNV GL allocated 30 sites (28 electric and 2 gas) to the certainty strata and the remaining 56 sample points to each non-certainty stratum roughly proportional to the total energy saved among sites in the stratum. A few minor deviations to this rule occurred when necessary. For example, at least one sample point was required in each non-certainty stratum. Table 7 shows a summary of the sample selected.

Table 7: Evaluation sample summary (number of fuel-specific sites)

Sampled Fuel & Cohort (Sampling Cohort)	Sample Frame	Certainty Sample	Random Sample	Total Sample
Electric	**166**	**28**	**30**	**58**
1 (2012)	44[4]	7	8	15
2 (2013)	38	5	9	14
3 (2014)	30	7	4	11
4 (2014)	23	6	4	10
Corporate	31	3	5	8
Gas	**126**	**2**	**26**	**28**
1 (2012)	45		9	9
2 (2013)	38	1	4	5
3 (2014)	30	1	3	4
4 (2014)	23		4	4
Corporate	31		6	6
Grand Total	**292**	**30**	**56**	**86**

DNV GL estimated a relative precision of 9.6% for total electricity saved, based on this sample. Cohort-specific electricity savings were expected to achieve a precision ranging from 13% to 23%. The estimated

[4] The dropped site was an Electric 2012 site with a single negative savings claim for the 2014 calendar year.

relative precision for total gas estimates was 13.5%, with an estimated range of 19% to 35% for the cohort level results. All relative precision estimates were calculated as two-tailed at the 90% confidence interval.

Additional details regarding the sampling methodology used are included in Appendix E.

3.3 Model Review

This section of the report summarizes DNV GL's review of MT&R and other regression models sampled for impact evaluation. DNV GL reviewed modeling files and specifications associated with all sampled sites. The model review process consisted of the following steps:

1. Collect the MT&R and other regression model files for each sampled site
2. Review a pilot group of files to assess the available information
3. Interview PDCs to understand the evolution of the regression model over the program's life
4. Review the models by type (MT&R, Energy Expert, Ecova model)
5. Compile and assess the model diagnostics
6. Summarize findings

The PDC provided information about the development of the MT&Rs and models used. They noted that the analysis template for the MT&R has evolved over time as the PDC worked to improve their process. In some cases, site specific adjustments to the baseline modeling methodology were made. The approach to independent variables also evolved over time. Cohort 1 models used only weather; the PDC started testing other independent variables in Cohort 2.

Timeline of SEM Implementation and Evaluation Savings Estimation

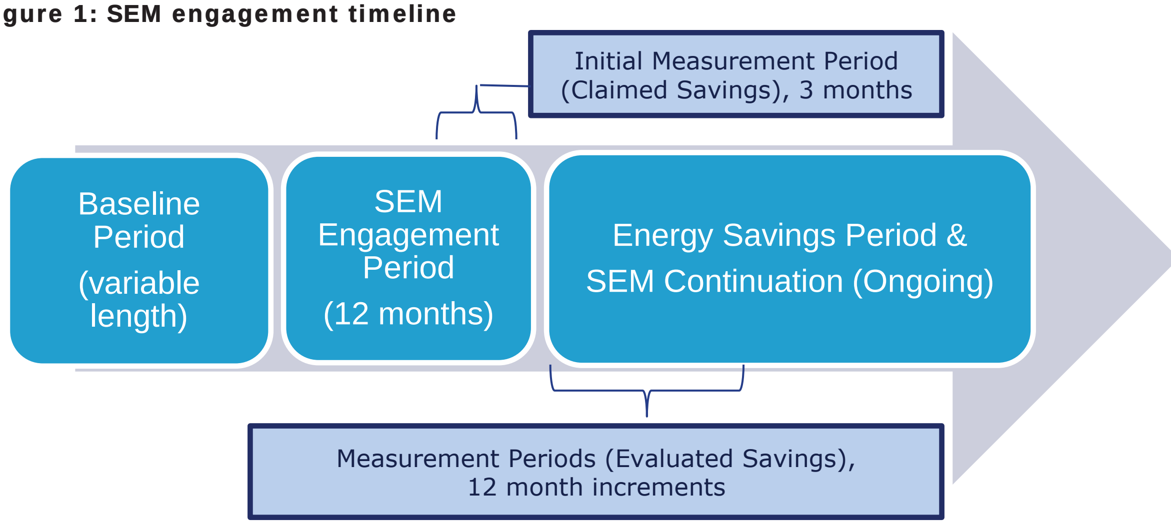
Three time periods and two measurement periods are considered in the determination of evaluated SEM savings. These periods occur over at least three years in the Commercial SEM program. DNV GL refers to these periods throughout this evaluation report. Figure 1 shows these periods for each SEM participant. These periods are:

- **Baseline Period**. This is the time directly before program intervention. Facility consumption during this period is used to develop a model that estimates baseline consumption after program intervention. The baseline period can range in length. For this evaluation, the baseline period length was set to 24 months unless insufficient data were available or a known capital project occurred 13 to 24 months prior to engagement. In these cases, a 12-month period was used.
- **SEM Engagement Period**. This is the period when a PDC works with a particular site. For the Commercial SEM program, this period starts with the kick-off meeting, ends with the report out, and is typically 12 months long. The end of the period is equivalent to the measure installation date in traditional measure based energy efficiency programs. Energy Trust claimed savings in the program year that the engagement period ended. For the Commercial SEM program years evaluated, the engagement period occurred during multiple program years for cohorts one and three.
    - **Initial Measurement Period.** Near the end of the SEM Engagement Year, the PDC compares recent consumption to the baseline model and estimates the annual energy savings acquisition reported for the program year. The Commercial SEM program typically used a three month

Initial Measurement Period and extrapolated the savings from these three months to a full year to calculate annual savings claimed by the program.

- **Energy Savings Period**. This is the period after program intervention (or measure installation) when the actual energy savings occur. The difference between actual consumption during this period and modeled baseline consumption, net of capital projects or other facility changes, is achieved program savings. Commercial SEM participants can elect to continue to engage with the program during this period. Any change in site-level savings in subsequent program years, compared to the savings estimate from the previous year, is claimed as an additional savings.
  - **Measurement Period**. This is the period used in the evaluation to determine the achieved annual savings due to SEM at each site. The evaluation used consecutive 12-month measurement periods starting immediately after the Engagement Period to calculate the savings achieved in each calendar year.

Figure 1: SEM engagement timeline



All SEM engagement activities listed in Table 3 occurred during the SEM engagement year. The evaluation used a measurement period starting immediately following the report out meeting or identified conclusion of SEM engagement. The Energy Trust assumes that the actions taken by SEM participants program would deliver energy savings over three years (i.e., assumed a three-year measure life).

Site Baseline Model

The PDC, with input from the customer and Energy Trust, built an initial site baseline regression model at the outset of the SEM engagement year based on baseline period energy consumption, weather, and other non-weather variables (e.g., occupancy). The baseline period occurred immediately before SEM engagement and was typically the 12 months prior to the start of program participation.

The evaluation team reviewed the concept of the model, the model form, and the variables considered for the models used to calculate energy savings resulting from the SEM activities. The modeling concept was to identify and collect relevant explanatory, or predictor, variables for energy consumption, and develop a regression analysis that was used to predict energy consumption as a function of these variables. The predictor variables in this analysis were temperature, or a variable such as degree-days representing weather conditions, and other predictors of energy consumption where relevant, such as occupancy. A model was fit to predict the monthly energy use for the baseline period with the following form:

- Energy = Intercept + B1 x Temperature variable + B2 x Other predictor variable
- Where:
    - Energy = monthly electricity or gas consumption by the facility
    - Intercept = constant, representing fixed electricity consumption independent of other variables
    - B1 and B2 are coefficients representing the relationship between weather and another predictor variable, respectively, with energy consumption.

The model was fit to the data for the baseline period. Assuming that the model form was appropriate in both the baseline and subsequent years, the model can be taken as a reasonable predictor of site energy consumption in the absence of SEM engagement.

The initial steps to build the regression models were:

1) Identify energy drivers: Energy use in commercial buildings is primarily weather driven. Additional variables may be appropriate; for example, occupancy was included in some models.

2) Collect consumption data: Energy meter, weather, and other pertinent data were collected for 1-2 years before the SEM program started.

With these two pieces of information, the PDC constructed a regression model that was designed to reasonably predict the energy consumption in future months based on how energy was consumed in the baseline period. The predicted energy consumption ("predicted value") was assumed to be the energy the building would have consumed if SEM was not implemented at that location. The difference between the predicted value consumption and the actual consumption was assumed to be the result of SEM, and was considered the savings due to SEM.

For Cohorts 1 through 4, the PDC typically calculated a site's first estimate of annual savings by using the last three months of billing data in the engagement year.[5] The steps used to calculate annual energy savings were:

- Calculate the difference between the actual usage and the modeled baseline usage for the last three billing periods of the SEM engagement year.
- Calculate the average savings per day across the billing periods analyzed.
- Extrapolate savings to a year by multiplying the savings per day by 365 days.
- Calculate annual SEM savings by subtracting incentivized capital project savings. Claim calculated savings if this is the first year the site is expected to acquire savings.
    - If the calculation did not show a change in consumption compared to the baseline, the program did not claim savings from SEM (positive or negative).
    - In early program years, no savings claims were made if the calculation showed an increase in consumption compared to the baseline. In later programs years, the program claimed negative savings if this calculation estimated an increase in consumption due to SEM.[6]
- Continuation sites only: For the second and third energy savings years, the program calculated the change in savings from the previous year and claimed the change in savings. If no change in

[5] It is unknown what the savings estimation process was for Corporate. However, based on the review of documentation and interviews completed, DNV GL believes the process was similar, if not exactly the same.

[6] Energy Trust recognized that excluding negative savings would inherently bias the savings claims.

consumption is calculated, the program claimed zero incremental savings. For this calculation, the PDC typically used twelve months of consumption.

Each estimated annual site savings claim was assumed to have a lifetime of three years, so the energy savings estimated for the first resource acquisition year were assumed to continue at the same level for the following two years. However, if the PDC saw a change in energy savings when updating the MT&R at the conclusion of each SEM continuation year, they reported the change in savings as the program's achievement during the continuation year. Program savings achievements in continuation years can therefore be positive and negative.

3.3.1 MT&R Review Findings

DNV GL reviewed the MT&R models produced by the SEM PDCs. The cohort MT&R-based models relied heavily on average temperature, as either a linear or squared term. In our opinion, this does not represent the underlying energy consumption process in commercial buildings. A temperature term (linear or squared) cannot account for both the energy required to heat and cool buildings. For example, in hot periods, energy used for cooling is positively correlated with temperature, but in cold periods energy used for heating is negatively correlated with temperature. The MT&R models often used the three months at the conclusion of the engagement year as the initial measurement period to estimate annual savings for the site. Extrapolating to annual savings from consumption over a partial year adds risk and uncertainty to the savings estimate.

DNV GL's conclusions from the MT&R model diagnostic exercise are as follows:

- Analyses and comparisons are difficult because of non-intuitive model coefficients (e.g., a temperature-squared term, negative linear temperature coefficients) and the variability of the units used between models.
- Based on our experience modeling consumption outside of this project, DNV GL finds the correlation coefficient (R-squared) terms for these models to be higher than expected and too closely skewed to the high end of the range. This suggests that incorrect model variables are utilized, the model is over fit, or there has been biased selection of baseline input data (perhaps through the use of an adjusted baseline).
- A high rate of non-standard baseline periods was observed in the sampled models: for both fuels, slightly less than one-quarter of models included non-standard baseline periods. The MT&R files inconsistently document the reason substantiating baseline adjustments. As a result, it was difficult for the evaluation to assess whether the baseline adjustments were necessary. Potential adjustment of the baseline period to maximize the R-squared term is a concern. The resulting model may be a good fit for the baseline data, but will not necessarily be a reliable predictor of annual consumption outside of the modeled timeframe.

3.3.2 Energy Expert Review Findings

In addition, DNV GL reviewed the Energy Expert models. Overall, the Energy Expert model appears robust and user friendly, but has little functionality for this evaluation. The evaluation was not able to directly assess the model's validity without receiving the interval data for each model. The baseline periods used cannot be verified independently without support from the software providers, which was not obtained.

3.3.3 Corporate Cohort Review Findings

Finally, DNV GL reviewed the models used in the Corporate cohort. Overall, the Corporate cohort modeling approach is similar to the MT&R files reviewed. This approach balances simplicity with a desire for accurate savings estimates. The baseline models appear to use the Georgia Tech EnPI tool[7] or another proprietary tool, which generate the regression equations. In one case, the model equations and data are clearly identified for each site. In others, a proprietary tool provides results.

The corporate PDC's models utilized CDD and/or HDD for many models. Average temperature was used for some models. For the three corporate participants, the models were completed using a spreadsheet model, such as the EnPI tool, but the original spreadsheet was not available for the evaluation. For one customer, the model equations and data are available in separate tabs of a large summary spreadsheet. It is currently unclear how three of the models were completed, but it appears they used CDD and possibly HDD as input variables in most cases. Gas models typically used HDD.

The evaluation did find items of concern in these models. One example is the lack of non-weather parameters when there is a clear need for these. A college campus' dorms were aggregated for electricity modeling. In this case, modeling was based strictly on the average temperature during the billing period. The highest consumption occurred during the colder winter months, while the lowest consumption occurs during the summer. The temperature-based model used a high intercept coefficient and a negative temperature coefficient. The result is a fairly smooth, wavelike forecast of consumption. However, actual consumption is not smooth and shows a dramatic reduction in consumption during the summer. Without any more information, the quick conclusion is that this period aligns with a reduction in occupancy for the summer session. A review of all the dorm-related consumption data in the file shows that this pattern existed every year. The appropriate variable to include in the model in this case may have been "number of weeks in session" during the billing period. Use of weeks will account for the variation in length of breaks throughout the year.

3.3.4 Model Review Conclusions & Implications

DNV GL had the following primary conclusions based on the review of the program regression models.

- DNV GL concluded that the MT&R models could not support the determination of evaluated savings. The models were created to support the program's information requirements and the participants' energy management needs. The files therefore balance energy modeling complexity with user simplicity. For example, the files are simple in their use of daily average temperature readings only. The additional step of degree day calculations would be a further sophistication. Since the models were expected to be operated by a building operator or energy champion, further sophistication did not guarantee an increase in accuracy, especially since this would place greater burden on the building operator. However, the models are complex in that the regression parameters are custom to each building based on its consumption and history. This tailoring of models created a quasi-custom analysis for each site that could not be readily verified by the evaluation.
- DNV GL concluded that a higher than expected risk of not achieving the claimed energy savings existed based on the modeling structure used. The MT&R models almost always use a linear average temperature term and often use a squared average temperature term. This approach required the least data processing and was therefore expected to result in the least user created modeling error. However,

[7] This tool, developed for Superior Energy Performance, is available at this DOE website, https://ecenter.ee.doe.gov/EM/SPM/.../Step%202.6.5%20EnPI%20Tool%20v3.02.xls

use of these terms does lead to non-intuitive regression coefficients that prevented simple review or comparison. Furthermore, extrapolating to annual savings from consumption over a partial year adds risk and uncertainty to the estimate.

The results of this modeling review supported DNV GL's decision to determine evaluated savings at sampled sites using a new model developed using a standard modeling specification (M2) that was different from the modeling specification used by the program. The M2 evaluation model provides an independent estimate of savings that is used to verify the program's achievements.

Appendix F provides additional information about the model review and associated evaluation implications.

3.4 Program Staff Interviews

DNV GL conducted interviews with Energy Trust and PDC staff responsible for delivering the program. The topics discussed during these interviews included:

1. SEM implementation: how it has changed over time and how it influences or drives the models and energy savings.
2. MT&R models and program savings estimation method.
3. Participants or buildings with unique or unusual circumstances, such as those who needed a special approach to the MT&R or savings estimates or had MT&Rs that were especially difficult to develop.
4. Participant engagement and the level of organizational commitment for each cohort and corporate participant.
5. Best methods for participant outreach and recruitment, and whether they are aware of any changes to the customer contacts.

DNV GL developed a structured guide for conducting the staff interview and data collection. The staff interview guide is available in Appendix C.

SEM Implementation

When the program was launched, it largely mirrored the industrial SEM program. The program design used the same cohort format and many of the industrial SEM forms and processes. However, the Energy Trust program manager did allow the implementation contractors to use their judgment to adapt the program to the commercial market. For instance, in the commercial market, the energy champions and teams are usually responsible for multiple buildings where in industrial it is more common for them to work with a single, large plant. The commercial program design was adjusted to facilitate participants with multiple participating sites. The program design sought to strike a balance between getting participant energy teams to work with enough managed sites to result in cost-effective savings for the program while not including so many that the team can't effectively engage with them all. Site specific MT&R development is also time consuming for the implementer, so they want to make sure the energy teams have the capacity and energy required for every building included in the program. If too many buildings are included in the program, achieved energy savings risk not being cost-effective to acquire.

Overall, Energy Trust is pleased with their investment in commercial SEM. SEM participants report high levels of satisfaction with the program and with Energy Trust. Even the participants who have dropped out of the program tend to be satisfied with their experience.

Though they generally regard the commercial SEM program as successful, program cost effectiveness is something they must manage. As mentioned earlier, striking a balance between the number of buildings each participant enrolls and the effort required to create each MT&R is important. The program implementers continue to identify ways to distill the program down to the most effective elements to make participation smoother and simpler for customers. Each change to the program is expected make it easier for participants to obtain more savings with less effort from the energy teams and the implementers.

The PDC described the following major steps in the process of SEM implementation.

Kick off meeting: The focus of the kick off meeting is to have the participants think about what they want to accomplish during their first year engagement. They establish goals and the goal metrics, and set objectives for the team and for energy policies. They establish guidelines for areas like engage employees or capital measures.

Energy management practices assessment: This assessment is held one-on-one with the PDC and each participant. They bring together the executive sponsor and other key roles within the participant organization, like finance, and the energy champion. The assessment takes about three hours to complete and is typically a challenging but important exercise. The PDC finds that it expands the participant's recognition that developing an energy saving culture goes beyond facilities management and requires engagement throughout the organization. For instance, recognizing how they budget for utilities and capital projects, and how well they keep records of usage information to provide feedback. The outcome is a near term action plan.

Effective team workshop: This group workshop covers who should be on the energy team, and how to effectively manage the team.

Building operations assessment: This is a one-on-one assessment that looks at each participant's building operations. They conduct these as soon as possible after the kick off meeting. The scan is a full day building audit during which they attempt to identify project opportunities. The PDC then spends another half day delivering training to the operators. The PDC will assess at least one of each participant's buildings; the number of assessments depends on how different or similar the participant's buildings are.

MT&R models/opportunity registers: The PDCs begin building the MT&R models in parallel with the building operations assessment. An opportunity register (now called an energy action item list) is developed in collaboration with the facilities team and energy champion through the building walks. An event log was added to the MT&R (in year 2) to create a more comprehensive list of all events completed. These event logs document the opportunities implemented and also capture other occurrences that may impact energy usage.

Operations and energy champion monthly calls: These calls are initially held every two weeks then move to once a month as the year progresses. In these calls, they review recent MT&R changes, discuss facilities events and ideas for new activities. The PDCs are available to problem solve and to provide ideas for new activities or occupant engagement.

Executive sponsor calls: These are monthly check in calls with the executive sponsor. These are important because, unlike the industrial SEM (where the exec sponsor may be a facilities person) in the Commercial SEM program, the executive sponsor is typically responsible for a much broader organization. As a result, they are able to participate in the kick off and energy management assessment but are likely not available to participate in the rest of the workshops.

Drivers of Savings

Typical energy savings to date are less than five percent of consumption. The participant activities that have driven the most energy savings have been the educational campaigns that make building occupants aware of their energy use and actions they can take to reduce it, building operator adjustments to time and temperature settings, and organizational policies such as removing space heaters and using power strips to control cubicle equipment.

SEM Continuation

During the first year of the commercial SEM program, it takes a lot of time on the part of both the energy teams and the PDCs to get the infrastructure in place. The primary objective of SEM Continuation is to make sure this infrastructure remains in place and the energy savings stick. A secondary objective is to achieve additional energy savings if they can identify the opportunities.

As first conceived, SEM Continuation was a separate agreement with a subset of six to seven participants to provide monitoring for a second year. However, these participants requested more support from the PDCs during this period, which proved to be successful at keeping the participants engaged, conducting additional activities, and adding more buildings. Based on this success, Energy Trust decided to offer SEM Continuation as a permanent program feature.

In SEM Continuation, the PDC continues to hold the monthly energy champion and building operator calls. During the calls, the participants give updates on their status and discuss specific issues the team have and opportunities they can pursue. They also hold approximately four group workshops each year. When participants add a building, the PDC will build a new MT&R workbook and conduct a facility site assessment. In some instances, the PDC will work with the participant to conduct a second facility site assessment on a participating building.

Pinpointing the specific benefits of SEM Continuation is not straightforward. While some participant practices dropped off after the first year of engagement, others are maintained because of the continued focus and motivation. The program has observed instances were competing organization priorities have been a distraction to the energy teams, but the PDC has been able to get them back on track. SEM Continuation also provides longer-term insights into which support and activities work well for the participants and which don’t work well.

Energy Champions

The selection of the energy champion is a driving factor in the success of the energy team; typically, they fall into two staff types: facilities staff or sustainability staff. Sustainability staff tended to get occupant engagement activities started sooner whereas the facilities staff tended to go after operational changes. How long they have been with their organization and the relationships they have across their organizations drives how effectively they can influence change. They also find that some energy champions are naturally more proactive and can move things forward better than less motivated people.

Through 2015, four energy champions have left their positions during the course of the program and the program team is aware of two more who have left since their program participation ended. One of these champions reappeared as an energy champion to replace another who had left their position.. It’s unclear what impact this had on participant engagement and energy savings.

MT&R Models

When developing the MT&R models, the PDC is looking to identify what drives the variability in energy use, not what drives usage. Weather was an identified driver of variability for almost every participating site; holidays and occupied square footage were identified drivers for some. Each MT&R model is expected to track site consumption and savings throughout the duration of program participation. However, the PDCs find that it's very difficult to track building performance over time, an issue compounded with SEM Continuation. The PDCs created a second MT&R model in a couple of instances where the participant had a change to their baseline.

Baseline period selection: The initial selected baseline period was the 12 months immediately preceding the start of SEM workshops. The final selected baseline period may have been shifted or adjusted until the baseline regression model had a good R-squared. While the standard baseline period length was 12 months, 18 and 24 month periods have also been used.

Weather modeling: The average temperature during the billing period is used as the independent variable for modeling and savings estimation; the MT&Rs do not use HDD or CDD. The PDC stated that the primary reason for this is that calculating HDD or CDD requires the calculation of a building balance temperature. Another issue with the weather variables is that in the early versions of the MT&R the participants input the weather data and each participant may have chosen a different source. Unless the PDC sees issues with the model, they won't scrutinize the participant's weather source.

Non-weather independent variables: The independent variables used for cohort 1 were limited to weather. For cohort 2, the PDC began including non-weather independent variables. The PDC tested an independent variable to determine if it should be included in the baseline model. The PDC tested "holidays" for most participants. Unique site specific variables used included "events" or "occupied square footage" for a meeting facility. For cohort 3 and beyond, the PDC analyzed other factors that were likely to drive variability in usage (such as daylight hours in buildings with many windows). However, weather and holidays were the most frequently used independent variables. For participants with solar PV, they try to get the system generation and add it back into the model to derive total energy used.

If a model could not be created with a good R-squared or p-value, the PDC generally used the average consumption per day during the baseline period to predict future consumption. This only happened in a small subset of models. In some cases, these were manually controlled buildings with no consistency in the building operation. The PDC observed that over the course of SEM participation, weather dependent energy consumption patterns emerged at these sites as the building operators became more engaged in energy management and implemented more automatic building controls.

Evolution of the MT&R

The MT&R models for cohort 1 were loosely structured without much explanation of the variables. After the first cohort, the PDC added a team member who worked to standardize the models and make them easier for the participants to use and maintain. Some of the improvements made by the PDC over time are:

- Added a building profile sheet with basic documentation such as meter number, square footage, building description;
- Weather can be automatically retrieved and calculated in later models. This was a significant improvement over the original laborious manual input and calculation method.

- Newer versions of the model update the graphs automatically. Original versions required manual updating;
- Standardized the function that calculates the energy savings;
- Added a dashboard feature;
- Created a single model validation tab where the user can view p-value and other variables on one tab;
- Combined gas and electric models into a single spreadsheet workbook;
- Automated the calculation to align non-weather variable values to utility billing dates;
- Added event lines to charts to make it easier to see the correlation between activities and change in performance; and
- Made changes to the scale of the model to better illustrate the level of variability in the model.

Specific MT&R Challenges

The PDC gave three examples of specific challenges they encountered while developing the MT&Rs:

- There was one waste water treatment plant that can be described as semi industrial. The challenge was that the facility had a lot of ongoing changes, so the PDC had to develop the MT&R model using a baseline of only eight to nine months.
- The PDC gathered information on capital measure projects from the Energy Trust and factored these into the MT&R model. However, there were some instances where the participant conducted significant upgrades without applying for a rebate.
- It has been difficult to create an MT&R when the participant has a capital improvement with savings that are seasonal. When this occurs, they shift the baseline instead of trying to incorporate the seasonality.

Building Aggregation

MT&Rs are developed at the building level for each fuel type. They will combine all the meters in one building, as long as the building is a single location, because it's not possible to separate out the different building systems. At campuses where they have a central plant, each building will have an MT&R along with one for the central plant.

Energy Expert

Some participants use PGE's tool, Energy Expert (developed by Northwright Inc.) Energy Expert is an hourly bin analysis tool that divides a week into 168 hours and matches each hour to the temperature. It's a good tool that calculates CUSUM saving similar to the MT&R, but there are challenges with developing the baseline and with reporting. It's also not for natural gas, only electricity. The PDC does not create an MT&R for participants who use Energy Expert.

Savings Persistence

When asked about the best predictor of persistence of the SEM practices, the Energy Trust program manager stated that it's probably a combination of factors. A strong energy champion supports persistence of savings, but not if they leave the position. Revising company policies and other documentation is another, but these aren't effective if they aren't used again. A simple MT&R tool is more likely to be maintained and will alert the team to backsliding on savings. Lastly, SEM Continuation, and the technical and organizational support it provides, supports persistence of savings.

The PDC staff gave two examples of participants who exemplified the extremes. The first participant originally thought there was limited opportunity for energy saving improvements because their organization had completed several projects before enrolling in the program. The energy champion was engaged and motivated despite this concern and ultimately realized significant savings through the program. This participant is an example of one who is very likely to maintain the savings because they were engaged at a deeper level.

In the other example, a participant with a strong and motivated energy champion engaged with the building engineers to realize significant savings during their engagement year. Unfortunately, the energy champion left the organization and, as a consequence, the team lost their motivation and their savings flattened out. Savings for this participant are not expected to persist due to lack of energy management engagement by the facility.

Future Changes

Because of the cost to build new MT&R models, the program may not encourage participants to enroll additional buildings in the future unless they are confident that the participant will see good savings. Instead, they may have them conduct their employee engagement campaigns across multiple sites. Starting with cohort 7 and 8, they will not extrapolate the energy savings from two or three months out to a year. Instead, they will consider the first full year of energy savings from the start of the SEM engagement and compute savings at the end of each individual year after that.

3.5 Participant Interviews

Energy Trust's Commercial SEM program from PY2012 – PY2014 had 27 unique organizations participate under four cohorts (1-4) and the corporate initiative. DNV GL conducted interviews with 20 of the 27 participants in the program. Of the 20 participant interviews, 12 were conducted via telephone and eight during the site visits. DNV GL attempted to interview all participants, but not all participants were successfully recruited. The SEM energy champion was typically the primary interviewee, though at two firms, two individuals were interviewed, including the energy champion.

Each participant represented an organization that typically had several sites enrolled in SEM. Some participants were interviewed more than once, because more than one site was selected for a site visit and interview. During the interview, two categories of questions were asked:

- **Organization Level:** General practices performed in the organization related to SEM.
- **Site Level:** Site-specific questions that covered the activities performed at a particular site/location.

During an interview, a maximum of four sites were discussed with each participant, regardless of the total number of individual sites operated by that participant, in order to keep the interview duration to a reasonable length. DNV GL engineers collected site-specific information for 34 sites during participant interviews. The interview topics included:

- Review of O&M and behavioral actions and capital measures, both implemented and planned at the site.
- Practices followed by the organizations to maintain their energy consumption-tracking tools (such as MT&R models).
- Recent operating conditions or changes to the facility that may affect the energy savings or the validity of the MT&R model.
- Role of energy champion and energy team to implement and maintain SEM activities.

- Persistence of energy savings, O&M activities, and SEM framework elements, including energy team/energy champion, workshops and executive sponsorship.
- Spillover of activities done at sites participating in SEM to non-participating sites.
- Feedback about the SEM offering and recommendations for improvement.

DNV GL developed a structured guide for conducting the interview and data collection. The guide can be found in Appendix D.

3.5.1 Organizational Level Findings

This section summarizes the findings from the participant interviews about the general practices that are adopted by organizations as a result of their participation in Energy Trust's SEM offering.

Maintaining the MT&R or energy consumption tracking tool

PDCs developed either an MT&R or another energy consumption-tracking tool, such as Energy Expert, to track energy use and calculate energy savings. Based on the interviews, all but one participant stated they were regularly tracking their energy use. The majority of the participants interviewed maintain their MT&R (15) or other tracking tool (4) on a regular basis; nearly half (10) maintain their MT&R monthly, 3 sites maintain them every 2 months, 1 site updates it every week and 1 site does it daily. This indicates that the participants recognize value in tracking their energy use. Some participants (7) acknowledged MT&Rs are a valuable tool to track energy savings at their sites, and that they helped promote activities listed in the opportunity register at other non-SEM participating sites within the same organization. Some participants (4) were using monitoring methods other than MT&R models; these sites are listed in Table 8.

Table 8: Activities adopted by participants who are not using MT&R models

Participant	Last MT&R Update	Tool used to track energy usage over time
1	2012	Use EnergyCap software[8] to populate ENERGY STAR portfolio manager
2	2012	Energy Star Portfolio Manager
3	2015	Energy Star Portfolio Manager and Ecova tracking system
4	Don't know	Participant specific software (monitors site inputs, outputs and reviewed monthly), spreadsheet based whole building metering

Twelve participants provided updated models for 23 sites with the latest energy consumption data and other variables to track energy use and monitor energy savings patterns over time. DNV GL extracted non-weather dependent variables from these models and used them for evaluated regression models.

Review of SEM activities

SEM activities that impact the continuation of SEM include engagement of the executive sponsor with the energy team and setting energy policy and energy use reduction goals to motivate the organization. Most of the participants (17 of 20) reported their executive sponsor has remained actively engaged with the SEM program after the SEM activities in the opportunity register were completed at their respective organization. One participant reported that they are in-between executive sponsors (their previous sponsor left the organization and the new one is coming up to speed) and two others reported that their energy team does not have an executive sponsor. Executive sponsor involvement after the program engagement period

---

[8] EnergyCAP is a family of energy management and energy accounting software products, used for tracking, managing, processing, reporting, benchmarking, and analyzing utility bills and energy and sustainability information. EnergyCAP software also interfaces with the EPA Energy Star Portfolio Manager.

generally involves regular meetings with the energy team (monthly to bi-monthly) to approve SEM activities and budgets, and review energy savings and savings metrics.

Most of the participants (16 of 20) have established an energy policy as a result of their participation in the program. Only four participants report that they did not set an official energy policy, but indicated that they strive for continuous improvement in energy savings. Most of the participants (13 of 16) who have set an energy policy due to SEM participation also defined numeric energy savings goals. Only two participants did not set any numeric energy use reduction goals. One participant already had an energy policy before participating in SEM; however, their reduction goal was revised due to SEM participation to achieve a 20% reduction by 2015. The participant informed DNV GL during the interview that there is no new official target set for the near future (as of March 2016).

The numeric energy-reduction goals that were set by some participants (13) are summarized in Table 9. One out of these 13 participants defined energy reduction targets in kBtu/ft^2 and did not provide any numeric goal to the interviewer. This suggests that most of the participating organizations (13 of 20) are actively engaged to implement and follow-up the energy savings delivered by SEM .

Table 9: Summary of Participant's Energy Savings Targets

Participant #	Reduction in Energy Use Goal (%)	Notes
1	3%	year over year
2	20%	until 2030 (2014 baseline energy)
3	15%	until 2016 (2011 baseline energy)
4	10%	(2010 baseline)
5	5%	year over year
6	3%	every year
7	25%	Greenhouse Gases Reduction
8	25%	in 10 years (2015 baseline year)
9	3%	every year
10	20%	by 2023 (Only for office spaces)
11	10%	
12	20%	

One of the participants had already achieved their target goal of a 3% energy reduction, while the other 11 are on their way to achieving their respective goals.

Role of Energy Champion and Energy Team

Fourteen of 19 participants maintain an organizational energy team led by an energy champion (response to question was not obtained at one site). Of the five participants that reported to not have an energy team, one participant reported that their energy team is composed solely of the energy champion, three participants informed that they do not have an energy team, one participant does not have a formal energy team and the energy champion meets with the executives, R&D and operation staff. Eleven of 14 energy champions have energy management as a documented part of their job responsibilities. The structure and responsibilities of the energy teams vary between the participants. In some cases, team members execute activities and track progress, while the energy champion's main function is to clear any obstacles. In another case, team members implement the actions while the energy champion's main function is maintaining the tracking and other documentation. Others split the different functions between team members.

The majority (9) of the energy teams meet with their energy champion every month, three organizations meet weekly, one meets biweekly and one meets quarterly.

Energy champions have changed at four of 20 organizations; three of these four new energy champions were new employees to the organization. Though a change in the energy champion can be disruptive to the energy team, there can be positive outcomes as well. In one case, the previous energy champion was replaced with someone who proved to be more passionate about the activities and resultant energy savings and showed more care towards the program. Another participant noted that SEM-related activities were put on hold during the transition, but now that the new person is on board, they can be resumed without further delay. The third participant indicated that their change in energy champion had little impact on the rest of the team's progress. However, the final participant's new energy champion stated he found the program challenging "because it is very administrative".

Program Spillover

Half of the interviewed participants (10 of 20) indicated that they have adopted SEM practices at other, non-participating sites. This demonstrates that the engagement of the energy teams and the value of SEM programs extended beyond the participating sites. Table 10 lists the additional buildings at which the participant energy teams have expanded their energy management practices.

Table 10: Summary of SEM Spillover Sites

Participant #	Buildings outside the program using strategic energy management practices
1	1 remote university campus in Oregon
2	9 facilities: 4 fire stations, 1 fire training center, 1 warehouse, 1 police department, 1 parking structure and 2 large offices, all located in Oregon.
3	2 facilities: 1 auditorium and 1 research lab & animal program center
4	11 buildings: 9 office buildings, 1 state police building, 1 juvenile detention department
5	3 prisons in Oregon
6	32 office facilities, inside and outside Oregon
7	5 hospitals in Oregon
8	61 restaurants
9	2 admin buildings and 2 hospitals
10	numerous facilities all across the state

Feedback about the program

Most of the participants (12 of 20) provided feedback about the SEM program. All but one participant considers this program to be very helpful in saving energy and providing valuable knowledge and tools to track energy use. Two participants praised the Cohort 1-4 implementer in helping them identify the energy savings opportunities. The only participant who was not very positive in regards to the SEM program said they "felt participation was tedious and time consuming because the program had never handled restaurants." With the exception of this comment, all remaining customers consider the SEM program to be a valuable asset to their organizations.

Organizational Level Interview Summary

1. Nearly all (19 of 20) interviewed participants reported continuing to use either MT&R models or some other means to track their energy use and savings resulting from SEM activities. This shows most of the organizations find the MT&R spreadsheets or using other tools to track energy savings helpful and valuable. This finding supports the assertion that energy savings persist even after direct engagement concludes.
2. Most of the participants currently have an executive sponsor involved in energy savings efforts and are striving to reduce energy use as a result of participation in SEM. Most participants have a defined energy policy and set a numeric energy use reduction goal because of participation in SEM.
3. A designated energy champion led most energy teams. They provided proper direction and accountability toward SEM activities. Interview results show that organizations with energy teams were more engaged and motivated to do SEM activities. Four organizations have new energy champions since beginning SEM, and in two organizations turnover negatively affected the implementation of the SEM.
4. Half (10 of 20) of the interviewed participants implemented SEM activities at other sites that were not participating in SEM. This shows that the engagement and the value of SEM programs extended beyond participating sites.
5. Most of the participants (11 out of 12) considered the SEM offering very valuable and acknowledged it resulted in energy savings in their organizations.

3.5.2 Site Level Findings

This section summarizes the information about specific sites that were discussed during participant interviews. The objective of the site level questions was to understand the activities performed during the SEM implementation and other practices that impacted energy consumption at the site. During the 20 participant interviews, 34 unique sites were discussed. The overall findings were:

1. Most of the sites completed the activities they defined under their opportunity registers and the associated activities can be reasonably expected to result in reductions in energy consumption.
2. Building energy use has changed at some sites due to factors outside of the identified SEM activities and weather. When possible, the associated data about these factors was collected and added to the evaluation energy model. However, not all changes in consumption, either increases or decreases, can be accounted for through the modeling process.

Actions Taken as part of SEM

Nearly all the sites (29 of 34) completed at least some of the actions they listed in their opportunity registers. Four out of five sites stated to have not completed any SEM actions and one site had a major renovation occur leaving no measures in effect.

Six out of 29 sites (Table 11) either added or removed the number of activities they listed in the opportunity register. Three of those six sites removed some activities and replaced them with other activities; two of the remaining three sites completed all but one of their action items listed in their opportunity register; and the remaining site did one more action item than what was on the list.

Table 11: Addition / Removal of SEM Activities from Opportunity Registers

Participant #	Site #	Activities removed at a specific site	Additional Activities completed at a specific site
A	1	Kitchen Variable Speed Drives (VSD)	New boiler heat exchanger installed
B	2	Awareness program	None
C	3	Investigated power strips, but couldn't find equipment to meet needs	Data center HVAC sequencing
D	4	Cooling tower measures	Chilled Water control valves (butterfly valve with Building Automation System integration) in 2013
			VSD on exhaust fans in 2015
E	5	Operable Window Closure	
F	6	None	Added EMS to one site as a demonstration/ pilot in 2015

Capital Projects performed during / after SEM participation

Almost half of the sites (16 of 34) completed capital projects after engaging in the SEM program. Fifteen out of those 16 sites are continuing the operation of those capital measures. One site did not complete the planned capital projects because they were discontinuing operation of the building. Four sites completed capital projects not identified in the 2015 program documentation. These are listed in Table 12.

Table 12: Additional Capital Projects performed at sites

Participant #	Site #	Additional Capital Projects performed
A	1	Lighting controls, HVAC fan motor VSDs
B	2	Building was completely renovated during SEM period.
C	3	Lighting upgrade, HVAC upgrade, HVAC Controls upgrade (completed March 2016
D	4	Solar PV Installation

The site that completed the solar PV project provided the most up-to-date information about the energy generated from the solar panels. DNV GL used the information about solar generation and completed capital projects in the evaluation models.

Understanding site-specific energy consumption and changes

Most of the sites (32) stated that energy use patterns have not changed since the time they participated in SEM. Only two sites reported increased energy use due to increased resident population and/or significant changes in occupancy during renovation.

About one-third of the sites (12 of 34) reported that factors other than outside air temperature driving energy consumption at their respective site. The other factors include occupancy, holiday schedules, wastewater treatment flow, and other site-specific factors. In all cases, the data supporting these factors were updated and included in the evaluation models. In most cases, these parameters were already included in the MT&R model and the data only required an update for evaluation. In one case, the evaluation created a new variable reflecting the effect of the parameter on energy consumption at the site.

3.6 On-Site Data Collection

DNV GL selected a subsample of sites to recruit for sites visits. The site visits provided an opportunity to assess the reasonableness of the regression models used to estimate savings, and the appropriateness of the independent variables on which the model is based. Initially, DNV GL planned to estimate savings for energy-saving actions and other measures using engineering analysis as a check on the regression models. However, this was not realistic in practice because the program did not collect sufficient information about measures completed. For example, information was not collected to sufficiently characterize the pre-existing and post-SEM conditions, such as equipment size, capacity, schedule, and load.

The goals for these visits were to:

- Acquire a better understanding of energy consumption drivers and the associated site model
- Confirm that capital and SEM actions were performed and are still in place
- Verify the completion and current operational status of large capital projects during SEM participation

The site visit subsample was designed to reduce uncertainty in the savings results from the evaluation. Sites were selected based on three criteria after review of the original models: 1) savings were large as a percentage of total annual consumption, 2) savings comprised a large percentage of total program savings, and/or 3) large capital projects were completed. Based on these criteria, 25 sites were selected and 19 were visited, based on the site contact’s response and timeline of the impact evaluation.

DNV GL engineers engaged with the energy team at each site to understand the energy consumption patterns. These on-site visits allowed for the verification of both capital project and SEM-driven actions and measures, as well as gaining an improved understanding of building use, operations and constraints. Additional data gathered from the site visits included verification of building system set point reset schedules, acquisition of on-site electrical generation, event occupancy, and EMS trend data. These findings later informed the development of the savings regression model and enabled improved characterization of the SEM-driven actions and measures performed at the sites. Some examples of site visit findings that explain the energy savings at the site are described below.

3.6.1 Site Visit Results

The information collected was site specific and covered a large range. Overall, the site visits provided useful insights to verify the SEM activities and capital projects beyond what was achievable in telephone interviews. The site’s energy team provided more information during an on-site visit than during the phone interviews, and DNV GL staff had the opportunity to better observe building systems and occupant practices and behaviors during onsite visits. In some cases, the data collected supported the inclusion of non-weather independent variables in the evaluation regression model. DNV GL also used the data collected to understand inconsistencies observed in the program models. This work helped the evaluation team understand the reasons behind significant changes in energy use and/or the cause of the energy savings. The following four examples illustrate the range of findings.

1. One of the visited sites claimed natural gas savings of 42% of their annual consumption. According to the MT&R event log, gas energy savings were driven by removing the wastewater heating system from service. During the visit, the site contact reported that the heating system was not removed (as noted in the MT&R), but that its operation is limited to a few days in winter months. Previously, the heating system was used all the time in the winter season, but now it is used only under extreme weather conditions to unfreeze the water pipes. The site contact confirmed the company

has reduced its natural gas consumption significantly, and this information collected during the site visit clearly explains the large savings observed.

2. At another site, the claimed natural gas savings resulted from renovations done in the dining hall and the closing of operable windows, thereby reducing heating load during winter months. During the site visit, the contact confirmed that renovations were done at the dining hall. The documentation provided by the implementers did not clearly state what kind of renovations were done, so it was not verifiable. It was found that the operable windows in the dining hall were open during the site visit and when the DNV GL engineer asked about the general trend, the site contact mentioned they are open most of the time. He also explained the increased usage of the dining hall since 2014 and change in building operation practices, such as shutting down heating during the winter break months. This indicated that the claimed savings were likely too high.
3. During another site visit, DNV GL confirmed that the building underwent a large renovation during the program period that resulted in a substantial reduction in building occupancy. This change in occupancy is identified in the MT&R but not accounted for in the savings analysis. The DNV GL engineer developed an occupancy level schedule with the site's energy team that was used as the non-weather input variable in the evaluation's estimate of savings. This information resulted in improving the regression model by taking into account a major driver for energy use (occupancy).
4. At another visited site, the program claimed electrical energy savings only. Capital projects implemented during this facility's involvement in the SEM program consisted of replacing the existing roofing and adding a 70 kW thin film solar PV system. The new roof has an increased R-value and reflectance compared to the previous one, which would reduce heating and cooling loads on the building. The solar electric generation system began operating in April of 2012 and offsets on-site electrical consumption. The tracked monthly PV system output was collected during the site visit. This information was incorporated into the evaluation model alongside the utility consumption data to calculate SEM energy savings.

4 GROSS SAVINGS METHODS

This section presents the methods used in this evaluation to develop gross savings, followed by a discussion of the results.

The gross savings analysis relied on statistical energy consumption modeling using available historic energy consumption, weather data, and non-weather dependent variables expected to influence consumption at a sampled site. DNV GL primarily copied monthly facility energy consumption from the MT&R files for the analysis. In some cases, Energy Trust or the PDC provided the monthly consumption directly.

DNV GL applied two methodologies to develop savings estimates for comparison with the claimed program achievements. The first method (Method 1, M1) used the same regression equation the program used to estimate savings, updated with data collected from the evaluation. The second method (Method 2, M2) follows common approaches to evaluate savings from regression models, incorporating the same weather data, but using different predictor variables than the program models.

Modeling vs. Fitting

One significant risk in statistical modeling is the trap of "over-fitting" to the available data when developing regression models. Curve-fitting tries to find an equation that fits well with the present data, while modeling tries to find an equation that represents the underlying data generator. Curve-fitting can be misleading and can lead to over-fitting in the sense that the fitted curve may not accurately represent periods of time outside of what was used to create the curve; the classic example is always being able to fit an (n-1)th-degree polynomial to n data points. For these regression models, the energy consumption should be directly correlated with what actually drives usage. As stated in Section3.3 Model Review, DNV GL concluded that the program models leaned too close toward over-fitting due to modeling methodology and parameter differences across the different sites. This conclusion led to the use of two methodologies to estimate program period baseline consumption, one that aligns with the original program (M1) and one that is independent of any curve-fitting (M2).

Modeling Criteria

DNV GL considered statistical criteria and the appropriateness of the model when developing models for use in this evaluation. In general, the strength of a model follows from its ability to tell a concise, consistent, and compelling story.

- *Concise* models are able to explain the appropriate amount of variation in the dependent variable under conditions experienced most frequently. There can be a large amount of variation in factors outside of weather that drive energy consumption. The intent of the energy consumption model is to best explain energy consumption as a function of weather and other predictor variables when those values are in the most common regions of their respective ranges.
- *Consistent* models have coefficient values with logical relationships. For example, a model should typically yield higher estimates of energy consumption as weather conditions become extreme or building occupancy or activity levels increase.
- *Compelling* models have a strong statistical fit. The probability that the coefficients are different than zero should generally be greater than 90%. Further, the overall model should account for a large amount of the observed variation in energy consumption. The adjusted R-squared statistic captures how much variation in the dependent variable (energy consumption) the model explains. Values greater than

0.8 denote a very strong statistical fit. Models that have an adjusted R-squared less than 0.5 are unable to explain half the variation in energy consumption.

To assess whether the models are consistent and concise, DNV GL assessed the available data on the drivers of energy consumption at SEM sites. Often we did not have sufficient visibility into the energy drivers to assess if the models were well defined. For example, hospitals likely have factors other than weather that drive energy consumption. However, we did consider if the models made sense overall, adapting appropriately to the known variables:

- Was energy consumption predicted to change appropriately in response to the weather conditions?
- Were the predicted savings reasonable for the actions and measures implemented?

For this evaluation, DNV GL used adjusted R-squared values to assess the statistical fit. Adjusted R-squared is reduced when the model includes too many predictor variables. Increasing the number of variables may lead to a high R-squared value, but also can lead to interpretation issues, especially when the predictor variable is seemingly unrelated to energy consumption. The evaluation therefore limited the independent variables to weather based variables and one non-weather variable.

4.1 Determination of Capital Project Savings

In both M1 and M2 it was necessary to subtract savings achieved through capital projects and accounted for in other Energy Trust programs. DNV GL received a database of capital projects supported by Energy Trust through the New Buildings and Existing Buildings programs. No numeric identifier existed to connect SEM program sites to the sites listed in this database. DNV GL reviewed the database visually and used the customer name, site name, and address listed to identify capital projects that occurred at sites in the SEM program. Projects were classified as before baseline, during the baseline period, or after baseline. Projects occurring before the baseline were removed from the list of capital projects occurring at SEM sites.

DNV GL identified large capital projects during the potential 24-month baseline period at three sites. Since these projects all occurred during the first 12 months of the 24-month baseline period, DNV GL adjusted the baseline period to the 12 months immediately preceding the engagement year. The resulting baseline models or outputs therefore did not require any adjustment for these projects.

DNV GL identified capital projects reported to occur after the baseline period at 28 of 86 sites. DNV GL assumed each project was installed on the date shown in the tracking database ("installeddt"). DNV GL assumed project savings started to exist on the day after the date listed. Project savings for the year the project was installed are prorated as:

- Prorated Annual Savings = (Days Left in Year) / 365 * Annual Savings.
- Where: "Days Left in Year" = Days between the install date and 12/31 of the same year.

Prorating is necessary to ensure that savings are correctly attributed to the SEM initiative during the year of capital project completion.

DNV GL calculated the annual capital savings from the "reported" savings listed in the Energy Trust database. DNV GL removed the transmission and distribution losses from the electric project savings and added back the savings assigned to freeridership. The final annual capital savings calculated is equivalent to the evaluated gross savings reported to have occurred at the site. Table 13 shows the total capital project savings subtracted from the modeling results by year and fuel.

Table 13: Total capital project energy savings subtracted from modeling results by evaluation measurement period

Fuel	2013	2014	2015
Electric Savings (kWh)	1,984,021	4,688,103	7,685,375
Gas Savings (therms)	21,135	29,448	30,254

4.2 Method 1: Program parameters approach

For Method 1 (M1), the objective was to simulate what the program models would have estimated for savings had the program calculated savings during the Energy Savings Period using consumption over a 12-month Measurement Period instead of forecasting savings based on consumption in the program year (Initial Measurement Period). DNV GL applied the same regression coefficients developed by the program to the data (weather and non-weather) associated with 12-month intervals in the Energy Savings Period. Capital project savings were subtracted from the modeling result to estimate final M1 savings for each Energy Savings Period. Savings were estimated for every year since the site's engagement year, even if the participant or site was not considered an active participant during that year. Calculating savings in subsequent years provides estimates of the energy savings achieved over the life of the measure.[9]

DNV GL was only able to complete the M1 analysis for sampled sites in which original program regression parameters were available, a single regression was used across all temperatures, the required independent non-weather variables were collected, and the analysis interval was monthly.

4.3 Method 2: Degree Day approach

For Method 2 (M2), DNV GL created a standardized regression modeling approach for gas and electric usage to estimate annual energy consumption for each sampled site (or associated meter if multiple meters serve one site). The development of the M2 approach was a direct result of the program model review. DNV GL utilized HDD and/or CDD, rather than average temperature as used in the MT&Rs, to capture the underlying physical heating and cooling processes. This standardized M2 modeling approach serves to independently verify the claimed program savings. For M2, DNV GL developed the best model for each site based on a standard modeling criteria. In order to find the best model for each site, DNV GL tested several different models using various reference temperatures:

- Heating only - uses HDD term only. This model was used for all gas models.
- Cooling only – uses CDD term only.
- Single reference temperature – uses HDD and CDD calculated using the same reference temperature.
- Dual reference temperatures – uses HDD and CDD, where unique reference temperatures are calculated separately for cooling and heating.

Model Selection & Development

DNV GL developed the models using site-specific data from the baseline period (consumption prior to the start of the program). DNV GL used 24 months as the baseline period unless sufficient data was unavailable or a large capital project occurred during the first 12 months of the 24-month period. If 24 months was not feasible, the baseline period was set to 12 months. Model development for each site occurred in two stages:

[9] In this case, the "measure" is all SEM related activities completed in one calendar year. The program currently assumes a three-year measure life.

Stage 1, Determination of optimal model type reference temperatures: The first stage determines the optimal reference temperature for each potential site model type. The temperature value that produced the highest adjusted R-squared value for a type was chosen to represent that type.

Stage 2, Model type selection: The best site model type of the four types listed above was the model type with the highest adjusted R-squared value. Table 14 shows the model types used for the 92 M2 models developed. Only four electric models selected the dual reference temperature model. This suggests that future model specifications do not need to include this modeling option.

Table 14: Selected evaluation model types

Fuel	Temperature Response Model Type	Model Count
Electric	CDD Only	14
Electric	CDD & HDD, Single Reference Temperature	35
Electric	CDD & HDD, Dual Reference Temperature	4
Electric	HDD Only	10
Electric	Subtotal	63
Gas	HDD Only	29
All	**Total**	**92**

Table 15 provides the average optimized degree-day reference temperatures used the selected by building type, fuel, and model type. If more than one model existed, the table also shows the minimum and maximum values used. These values can be used to inform future SEM modeling guidelines and as comparison for future program developed models. The minimum and maximum degree-day reference temperatures allowed were 50°F and 84°F, respectively. Constraints on the reference temperatures prevent the model from optimizing to values that prevent the model from modeling the energy consumption process.

Table 15: Method 2 Optimal reference temperatures by building type, fuel, and model type

				CDD Reference			HDD Reference		
Building Type	Fuel	Count	Model Type	Min	Max	Avg.	Min	Max	Avg.
Education	Electric	1	CDD & HDD, Dual			77.0			56.0
		3	CDD & HDD, Single	54	69	63.7	54	69	63.7
		2	CDD Only	53	73	63.0			
		1	HDD Only						62.0
	Gas	3	HDD Only				58	66	62.7
Jail	Electric	2	CDD & HDD, Single	50	52	51.0	50	52	51.0
		1	CDD Only			52.0			
		1	HDD Only						74.0
	Gas	6	HDD Only				57	84	71.5
Medical	Electric	2	CDD & HDD, Single	50	54	52.0	50	54	52.0
		2	CDD Only	50	50	50.0			
		1	HDD Only						84.0
	Gas	1	HDD Only						84.0
Meeting Space	Electric	1	CDD & HDD, Single			61.0			61.0
		1	CDD Only			51.0			
	Gas	1	HDD Only						63.0
Office	Electric	1	CDD & HDD, Dual			77.0			53.0
		21	CDD & HDD, Single	50	82	58.8	50	82	58.8
		6	CDD Only	51	70	57.5			

Building Type	Fuel	Count	Model Type	CDD Reference			HDD Reference		
				Min	Max	Avg.	Min	Max	Avg.
		6	HDD Only				50	84	78.3
	Gas	9	HDD Only				50	84	61.4
WWT Plant	Electric	1	CDD Only			76.0			
	Gas	1	HDD Only						69.0
Residence	Electric	1	CDD & HDD, Dual			80.0			58.0
		1	HDD Only						68.0
	Gas	2	HDD Only				61	64	62.5
Restaurant	Electric	1	CDD & HDD, Dual			67.0			50.0
		6	CDD & HDD, Single	50	62	54.7	50	62	54.7
		1	CDD Only			60.0			
	Gas	6	HDD Only				50	84	63.3

Savings Calculation Methodology

DNV GL used the model selected to calculate evaluated energy savings for each site for each 12-month Energy Savings Period that followed the site's engagement year. The following steps were taken to complete this calculation:

1. DNV GL calculated M2 meter-level monthly energy savings as the difference between the estimated baseline consumption (using the regression model) and actual meter consumption (residual for the read period). All calculations used monthly utility meter reads and daily weather data aggregated to each utility meter read period.
2. DNV GL calculated partial annual savings as the sum of monthly differences with read dates in the calendar year.[10]
3. DNV GL calculated full annual savings by calculating the average savings per day across all days represented in the year's metering periods and then multiplying this value by 365 days. This adjustment was not significant but was required since the associated metering periods often represented slightly more or slightly less than 365 days. The savings calculated in this step represent the annual savings observed at the meter for the model. However, these savings are not adjusted for capital measures known to have been installed at the facility.
4. If multiple models exist for one site, then annual savings are summed before capital project savings are subtracted.
5. DNV GL calculated evaluated site-level energy savings as the difference between the savings calculated in step 3 above and capital project savings accounted for in other Energy Trust programs, based on the reported program savings for the capital project measures. These savings represent the energy efficiency savings at the site during the calendar year attributable to SEM.

[10] This means that the calculated result does not perfectly align with the calendar year. DNV GL considers the associated error to be insignificant. This methodology also improves the alignment of savings to the program delivery schedule.

5 EVALUATION RESULTS

This section describes the savings calculated using the methodologies discussed above and the associated estimates of program savings acquisition based on the sampling methodology used.

5.1 Site Results

Method 1

DNV GL completed a Method 1 analysis for 46 of the 86 sampled sites, 30 electric and 16 gas. In general, the M1 and M2 site results are aligned, but some significant differences do exist. The similarities and differences in the results are discussed further below. DNV GL did not use these results to estimate final evaluated savings, but the results suggest how much of the variance in savings may have been due to the measurement period used versus the change in modeling methodology.

DNV GL did not complete an M1 analysis for 40 sites: 16 due MT&Rs with change point models, 23 due different model types (Energy Expert sites and Corporate cohort sites), and one due to lack of data (use of weekly consumption in MT&R).

Method 2

DNV GL completed a Method 2 analysis for all 86 sites using 92 unique models. DNV GL calculated achieved savings for all years after the site's engagement year through 2015. Enough variability exists in the site-level results that it is impossible to show examples that are representative. This methodology estimated savings that were both higher and lower than the program's claims on a site-by-site basis. Results for 26 sites included an estimate of negative savings in at least one energy savings period. The report discusses the results at a program level after the next section. All conclusions are drawn from savings comparisons at the program level.

Each site-level savings estimate uses a baseline model determined to be the optimal model type using the optimal reference temperature.

Example Site Result

Table 16 shows an example of M1 and M2 measurement period results for one site. This table shows the energy savings achieved in the three measurement periods following the engagement year. The increase in claimed savings from 2013 to 2014 means that incremental savings were reported at the conclusion of PY2013 that were expected to be measurable in 2014.

Table 16: Site 2501, cumulative kWh savings by model and energy savings period

Model	2013	2014	2015	Three Year Total
Claimed Savings	381,078	544,892	544,892	1,470,862
M1 Savings	442,805	403,205	486,874	1,332,884
M2 Savings	464,092	391,102	485,207	1,340,402
Realization Rates				
M1/Claimed	116%	74%	89%	91%
M2/Claimed	122%	72%	89%	91%

Table 17 shows the savings reported by program year for the same site and the change in evaluated savings associated with each program year. The PY2013 column shown in this table is the difference between the 2014 and 2013 columns in the table above.

Table 17: Site 2501, measured incremental kWh savings by model and associated program year

Model	PY2012	PY2013	PY2014
Claimed Savings	381,078	163,814	
M1 Savings	442,805	(-39,599)	83,669
M2 Savings	464,092	(-72,990)	94,105
Realization Rates			
M1/Claimed	116%	-24%	N/A
M2/Claimed	122%	-45%	N/A

This site demonstrates some of the challenges faced by the program and evaluation when estimating and determine savings for Commercial SEM. Opportunities exist to overcome these challenges or at least reduce their impact, but none of the potential changes are simple and all require careful coordination across Energy Trust programs and stakeholders.

- **Claimed and evaluated savings are expected to be different.** The program estimated the savings that will be observable in a future energy savings period based on previous consumption and weather while the evaluation estimated savings for the same period based on actual consumption and weather. The PY2012 claimed savings for 381,078 kWh was determined based on the average daily savings observed during a 4-month Initial Measurement Period from 6/14/2012 - 10/12/2012 during the Engagement Period. In this case, the extrapolation underestimated the savings that the evaluation observed in the 12-month Measurement Period following this claim. During the first continuation year, the program reported an additional 163,814 kWh in savings based on the average daily savings observed during a 7-month Measurement Period from 4/12/2013 – 11/12/2013 minus the installed capital project savings of 32,109 kWh. The total 544,892 kWh reported across the two program years was expected to be measurable in 2014. However, both the M1 and M2 analyses estimated that achieved savings dropped from 2013 to 2014. The program did not report an additional savings claim in PY2014 based on 2014 consumption, but the evaluation models estimated that savings rebounded to above the first year values, but still below the cumulative savings reported for the site. This challenge will always exist if program and evaluation reporting are structured this way. However, the variance could be minimized if both estimates were normalized to typical conditions instead of actual conditions. Normalizing would introduce new challenges associated with developing a model of post engagement period consumption instead of relying on actual consumption during the period. DNV GL recommends that the Energy Trust carefully consider the benefits and challenges that normalization would bring.
- **Separate reporting of capital project savings may have resulted in incorrect estimates of Commercial SEM savings.** This site completed a lighting project through the Existing Buildings program in late 2012. The savings reported for this project through the Existing Buildings program was removed from the program and evaluation estimates of SEM savings to avoid double counting. However, if the capital project savings were incorrect, the variance impacts the SEM program's achievements. If the capital project savings were overestimated, then the calculated SEM savings underestimate SEM's impact. Similarly, if the capital savings are underestimated, SEM savings are an overestimate of SEM's impact. This challenge can be overcome if all savings at SEM sites are reported and evaluated together.
- **The evaluation produces multiple results that require careful interpretation.** The underestimate of PY2012 savings produces a realization rate that could be used to adjust first year claims to savings delivered in the first year. This result and its potential application are intuitive, but do not adjust savings to the average annual savings observed across multiple years only to the first year.

- The program estimated that savings went up due to activities in PY2013, while the evaluation concluded that savings went down. This is in part since the program estimate is on 2013 consumption and the evaluation estimate uses 2014 consumption. The incremental realization rate is negative because the direction of the change is different between the two estimates, not because negative savings occured. If the program estimated a reduction in savings, but evaluation found an increase in savings, the simple ratio would also be negative. This result is strictly a realization rate on the additional savings acquired, not the actual savings achieved.
- Finally, no realization rate can be calculated for the change in savings measured by the evaluation from 2014 to 2015 since the program did not make an incremental savings claim for this time period in PY2014. For sites where the program did not report a change in savings in a particular year, but the evaluation observed a change in savings, no site level realization rate can be reported for that program year.

5.2 Method 1: Program Results

DNV GL summed the savings estimates for sites with M1 results instead of extrapolating the results based on the sample design. The number of models completed and the sample design used prevent extrapolation. Table 18 shows the measured savings and ratios calculated for each program year by fuel. These are the total cumulative savings observed in the calendar year, not the incremental savings.

Table 18: Method 1 Total Sample Results by measurement period

Fuel	Estimate	2013	2014	2015
Electric	n - # of models	4	12	30
	Claimed Savings	1,269,154	2,951,021	10,397,582
	M1 Savings	456,347	4,353,865	7,980,351
	M2 Savings	-863,039	3,100,207	7,626,332
	M1/Claimed	36%	148%	77%
	M2/Claimed	-68%	105%	73%
Gas	n - # of models	7	11	16
	Claimed Savings	65,879	215,436	320,896
	M1 Savings	318,675	428,175	542,998
	M2 Savings	-6,900	106,110	216,008
	M1/Claimed	484%	199%	169%
	M2/Claimed	-10%	49%	67%

The most informative ratios are the 2015 ratios since DNV GL completed a 2015 savings estimate for all 46 sites. The PY2012 and PY2013 ratios are based on only a few results so the error in the estimate is large. Furthermore, 2015 is the most recent program year analyzed, so these ratios best represent recent or current program practice and the current mix of new and continuation sites. For 2015, the table shows:

1. If the program estimated electric savings based on a 12-month measurement period following the engagement period instead of using shorter initial measurement periods, estimated energy savings would have been 77% of the actual claims and only 5% greater than the evaluation claims. In this case, it appears that the timing of the analysis had the largest impact on savings versus the modeling method used.
2. However, if the program used 12-month measurement periods to estimate gas savings, cumulative claimed savings would have been 69% higher and over double the evaluation's estimate of savings

for the year. For gas savings, both the timing and methodology appear to be driving the variance between claimed and evaluated savings.

DNV GL suggests that Energy Trust consider calculating savings for SEM at the conclusion of a defined Energy Savings Period that occurs after the conclusion of the Engagement Period. Changing the program reporting process will produce different estimates of savings that: 1) are better aligned with actual consumption during the same reporting year, and 2) can accurately account for capital projects that occur in each year. The analysis also shows that there is potential to improve the realization rate and reduce the evaluation risk to the program. However, comparison of the estimated cumulative gas savings achieved in 2015, demonstrates that the timing of the analysis is not the only reason for the variance between claimed and evaluated results.

5.3 Method 2: Program Results

DNV GL estimated the total energy efficiency resource acquired in each calendar year by extrapolating the sampled site results to the total sampling frame based on the sample design used.

Extrapolation to the program

DNV GL used a separate ratio estimator to obtain unbiased estimates of the total modeled savings (either kWh or therms) for any group of interest. This estimator will yield, by design, unbiased estimates of some outcome measure, and is particularly beneficial when the outcome measure is correlated with something known for all members of the sample frame. In this case, the modeled savings are logically correlated with claimed savings as listed in the tracking database. In general, the separate ratio estimator works as follows.

Suppose the indices:

g = Sample design strata. This is defined by certainty indicator, cohort group, size group and fuel type (kWh or therms). For some outcome measures and some reporting domains of interest, these strata had to be collapsed with one another during the estimation process. This occurred with $Y_g \neq 0$ but $\sum_{i \in Sample} w_{ig} y_{ig} = 0$ (these terms are defined below).

i = Site.

And suppose:

x_{ig} = Modeled savings for site i in group g.

y_{ig} = Claimed savings for site i in group g.

w_{ig} = Sample weight for site i in group g. This reflects from the sample selection process that was used at the beginning of the study to select the 86 sites.

Y_g = Population total claimed savings in group g. So $Y_g = \sum_{i \in Frame} y_{ig}$

$$\hat{R}_g = \frac{\sum_{i \in Sample} w_{ig} x_{ig}}{\sum_{i \in Sample} w_{ig} y_{ig}}$$ is the Ratio estimate for group g.

Then the separate ratio estimator that will yield the total model savings is:

$$\hat{T} = \sum_g \left(Y_g \cdot \hat{R}_g \right)$$

And the ratio estimate of total modeled savings to total claimed savings is:

$$\hat{R} = \frac{\hat{T}}{\sum_g Y_g}$$

The reported standard errors of $\hat{T}$ and $\hat{R}$ were computed using the Taylor Series Linearization method. This method was first suggested by Tepping (1968) and has been discussed in numerous articles and books since then [see, for example, Binder (1983) and Wolter (1985)].[11] The Taylor Series Linearization process for estimating variances accounts for the complex design features of a study including stratification, clustering, multistage sampling and/or unequal weighting. So this variance estimator is appropriate for the estimates produced from this effort. Note that this variance estimator does not account for any uncertainties or additional variances that might be associated with each site-level model predictions x_{ig} or y_{ig} .

Evaluation Results by Calendar Year Achievement

DNV GL extrapolated the M2 site results to determine the annual energy savings achieved by the program. Overall, the program acquired energy savings each year, but our confidence in the final estimates is less than expected. The table also shows the impact of negative savings. Negative savings, or consumption increases, will occur, and since it is unlikely that any increase is attributable to SEM, the changes are assumed to be random. DNV GL also assumed that the identified positive savings include a random amount of decreases in consumption not attributable to the program. The evaluation assumes these random consumption changes cancel each other in the population. However, due to the small sample sizes within some strata, an increase in consumption can affect the result for a specific year and fuel.

Table 19 shows the total energy savings achieved by all participants in each calendar year by fuel. These values are the total savings achieved in the calendar year compared to the baseline, net of capital projects. A table of results by program year cohort is included in Appendix A and a table of detailed results is included in Appendix B. The table shows that the accuracy of the savings claims has improved as the program matured. This is likely due to multiple factors, including improved program delivery, improved understanding of building consumption at continuation sites, and adjusting site savings claims at the end of each program year.

[11] Binder, D.A. (1983). "On the variances of asymptotically normal estimators from complex surveys." International Statistical Review, 51, 279-292.
Tepping, B.J. (1968). "The estimation of variance in complex surveys." Proceedings of the Social Statistics Section of the American Statistical Society, pp 11-18.
Wolter, K.M. (1985). Introduction to Variance Estimation. New York: Springer-Verlag

Table 19: Annual savings acquired across all strata by fuel

Fuel	Savings Year	Claimed Savings	Evaluated Savings	Realization Rate	Baseline Consumption	% of Baseline Consumption
Electric (kWh)	2013	5,299,318	7,350,568	139%	175,140,706	4.2%
	2014	14,024,257	16,338,244	116%	318,777,912	5.1%
	2015	26,959,489	27,852,207	104%	394,271,281	7.1%
Gas (therms)	2013	126,942	-18,452	-15%†	4,001,720	-0.5%
	2014	496,277	155,938	31%	7,340,910	2.1%
	2015	926,966	846,577	91%	13,051,759	6.5%

†=Not statistically significant from zero

Table 20 shows the number of sites claimed to be achieving savings in each year, the number of sites evaluated, the realization rate, and the relative precision of the realization rate at 90% confidence. The relative precision of the results improves with each year primarily due to the increase in participation and associated increase in sample. Except for 2013 gas savings, the evaluated savings are statistically significant from zero. Unfortunately, the variance observed in the evaluated projects resulted in larger relative precisions than originally estimated. Only the realization rates for gas in 2013 and 2014 are significantly different from 100%. This means that a 100% realization rate is within the 90% confidence interval for all other results. The realization rates shown are the results of the study, but the true savings achieved could be equal to the savings originally claimed, except for gas in 2013 and 2014.

Table 20: Relative precision of the realization rates

Fuel	Savings Year	Participating Sites	Evaluated Sites	Realization Rate	RP of Realization Rate
Electric (kWh)	2013	27	12	139%	66%
	2014	98	36	116%	30%
	2015	167	58	104%	21%
Gas (therms)	2013	13	7	-15%*	700%
	2014	78	20	31%*	73%
	2015	126	28	91%	26%

RP=relative precision; CI=confidence interval; *=Statistically different from 100%

Table 21 compares the claimed and evaluated first year savings by program year. The total first-year savings in this table equals the 2015 calendar year program savings since the savings acquired in all programs years still occur at the meter in calendar year 2015. The participating site counts shown in this table are the number of unique sites that included an incremental savings claim for the program year. The evaluation believes there are multiple reasons for the variance between claimed and evaluated first-year savings. One reason is the different time periods used to estimate these values and the inherent challenge faced by the program in estimating 12-months of future SEM savings based on a partial year of facility consumption.

Table 21: Incremental savings acquired in each program year

Resource Fuel	Program Year	Participating Sites	Claimed Savings	Evaluated Sites	Evaluated Savings	Realization Rate
Electric (kWh)	PY2012	27	5,299,318	12	7,350,568	139%
	PY2013	81	8,724,939	36	8,987,675	103%
	PY2014	97	12,935,232	58	11,513,963	90%
	Total	**167**	**26,959,489**	**58**	**27,852,207**	**104%**
Gas (therms)	PY2012	13	126,942	7	(18,452)	-15%
	PY2013	72	369,335	20	174,390	47%
	PY2014	65	430,689	28	690,639	160%
	Total	**126**	**926,966**	**28**	**846,577**	**91%**

5.3.1 Initial Savings Determination

DNV GL also compared the M2 energy savings estimates with the claimed savings for a site's first year. This comparison assesses the accuracy of the program's initial savings claims at the end of the engagement year and if the accuracy of initial claims has improved over time. These initial estimates, non-zero or zero, were often based on a fraction of meter reads that occurred during the engagement year requiring the program to extrapolate from a few meter reads to an entire year. Extrapolation from a partial year to a full year introduces a savings risk to the program. The evaluation created the following terms for this analysis:

- **First Eligible:** First eligible refers to the first year a site was eligible to deliver savings. This is equal to the year immediately following the SEM engagement year. In some cases, the program did not claim savings for a site after the initial engagement year, but did claim savings in subsequent years.
- **First Claimed:** First claimed is the first year the program claimed savings for the site, irrespective of fuel. In some years, the program claimed savings for one fuel, but decided the engagement year consumption pattern for the other fuel did not merit a savings claim.

Table 22 provides the results of this comparison. The program underestimated initial cohort 1 and 2 savings for electric, but overestimated gas savings. This demonstrates the risk inherent in estimating annual savings based on consumption over a fraction of the year. While not known, it is possible that a reliance on end of year consumption values during the heating season contributed to an overestimation of annual gas consumption and savings. The table also shows that more recent initial claims (cohorts 3 & 4) have overestimated electric savings and underestimated gas savings. DNV GL's recommendation to calculate savings at the end of each savings acquisition year would eliminate the risk associated with extrapolating savings to a full year from a partial year.

Table 22: Realization Rates, First Eligible and First Claimed Year, by fuel and sampling cohort

Sampling	Electric (kWh)		Gas (therms)	
Cohort	First Eligible Year	First Claimed Year	First Eligible Year	First Claimed Year
2012	153%	105%	-14%	42%
2013	129%	141%	21%	33%
2014	75%	75%	177%	177%

5.3.2 Indications of Measure Persistence

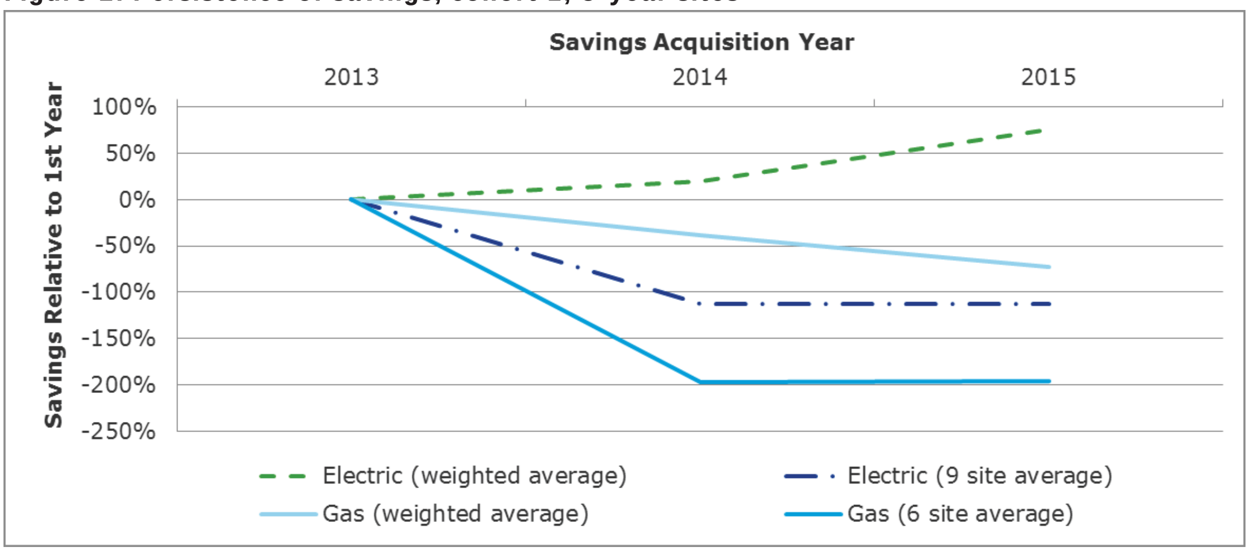
DNV GL assessed savings persistence by reviewing the energy savings calculated for 15 of the 16 cohort 1 sites with three completed savings acquisition years, 9 electric and 6 gas. Results for the one electric site not included are abnormally large within this small sample. DNV GL did not include the site to prevent it from completely driving the resulting weighted averages calculated. The remaining 15 sites are the only other sites in the evaluation sample with three savings acquisition years.

The persistence metric must indicate if savings are going up or down and what that change is relative to the first year. DNV GL calculated a persistence rate for each site as the ratio of the difference between evaluated savings in each year and the evaluated savings in 2013 to the absolute value of the evaluated savings in 2013. For example, the persistence rate for 2014 is calculated:

2014 Persistence Rate = (2014 Savings – 2013 Savings) / Absolute Value (2013 Savings)

This provides the percentage change in acquired savings normalized to the first eligible year. If the percentage is positive, then savings increased. If the percentage is negative, then savings decreased. If the percentage is less than -100%, then savings no longer occurred and consumption increased. DNV GL calculated persistence rates for each site and for the total savings for each fuel. Figure 2 shows the average values by fuel and the weighted average values by fuel.

Figure 2: Persistence of savings, cohort 1, 3-year sites



There is substantial variation in the values calculated, so the relative precision of the calculated values is high. However, the trends are informative to the program.

- 6 of 9 (67%) of electric sites had reduced savings in the second year, but only 3 of 9 (33%) of electric sites' savings reduced from the second to third year. While the weighted average of electric savings shows an increase in savings in years two (20%) and three (75%), on average a site's savings declined in subsequent years, -112% in year two and -113% in year three. The difference between the weighted and non-weighted persistence rates was due to a large increase in savings at one large participant. These results suggest that while the program may have achieved more savings in the second year, a significant number of sites will achieve less savings in the second year and some will show increases in consumption.

- 5 of 6 (83%) of gas sites had reduced savings in the second year and 5 of 6 (83%) of gas sites' savings reduced from the second to the third year. The average site lost 200% of initial savings in 2014. This means that the average site's consumption increased compared to the baseline by as much as was saved in the first year. The small sample and the estimation of negative savings for some sites in certain years prevent strong conclusions from being drawn based on these results.

DNV GL completed a similar analysis for cohort 2 sites, which first claimed savings in 2013. These sites only have two years of savings results to analyze and the results are different than cohort 1.

- 9 of 13 (69%) electric cohort 2 sites had increased savings in the second year. The weighted persistence rate was 39% and the non-weighted average persistence rate was 38% (one electric cohort 2 site was removed from this analysis since it was an extreme outlier).
- 4 of 5 (80%) gas sites had increased savings in the second year. The weighted persistence rate was 33% and the non-weighted average persistence rate was 31%.

DNV GL concludes that the persistence of SEM program savings requires more research before conclusions can be made. The evaluation cannot easily explain the difference in second year persistence rates between cohorts 1 and 2. Future evaluations will have a larger sample of participants in their third and fourth year of acquisition and should be able to provide confident estimates of persistence. DNV GL recommends that the Energy Trust complete additional research to better understand the persistence of SEM savings.

6 CONCLUSIONS AND RECOMMENDATIONS

DNV GL concludes that energy efficiency resource savings were acquired by Energy Trust's Commercial SEM offering during the analysis period. The savings are observable at the site and statistically different than zero in all years and fuels except for gas savings in 2013. This conclusion is based on the results of an analyzed sample of 86 unique participant sites extrapolated to the program population. The following items address the specific objectives of this evaluation.

Develop reliable estimates of gas and electric savings realized by the program

DNV GL estimated savings using a different method from the program, providing an independent estimate of savings. Further, the evaluation confirmed savings were occurring through interviews and site visits with participants. The evaluation concludes that from January 2013 through December 2015 the program achieved 51,541 MWh and 984,063 therms of energy efficiency savings. These values are 111% and 63% of the energy savings claimed by Energy Trust to have occurred over the same three-year period. Only the gas realization rate is statistically different from 100%.

Individual site savings estimates from the evaluation did vary from the tracking data. This result is not unexpected, given that the implementer and evaluation models for energy consumption did not attempt to not account for the wide range of possibilities that may affect energy consumption at a commercial facility. This observed variation in the modeled saving estimates was higher than expected resulting in higher uncertainty due to sampling than originally anticipated. The evaluation can conclude that savings are occurring, but for most fuels and years the evaluated savings estimates are not statistically different from the tracking estimates. Future evaluations of SEM participants should expect high variation in the results when designing the study's sample.

The savings represent 5.8% of the modeled electric baseline consumption and 4.0% of the modeled gas baseline consumption over the same three-year period. The most recent year, however, achieved 7.0% of electric baseline consumption and 6.5% of gas baseline consumption.

Account for capital project savings at SEM participant sites

Using the data provided by Energy Trust, DNV GL identified 14,357 MWh and 80,837 therms of capital projects that occurred after the conclusion of the baseline period at sampled SEM sites. These savings were removed from the initial savings estimates involving the difference between the baseline modeled consumption and the actual consumption. The savings removed were equivalent to the adjusted gross savings expected to be achieved by each project. Electricity savings were also reduced by 10% to remove the T&D losses included in the reported value. DNV GL did not adjust capital project savings further at any location, even if visited. Further adjustment would have resulted in unequal treatment of sampled sites. No adjustment was made to ensure equal treatment of capital project savings across all sample points and accurate accounting of all savings achieved by Energy Trust across all its programs. Future SEM impact evaluations should conduct measurement and verification research for all SEM and capital measure savings claimed at each sampled site.

Report observations and recommend ways to improve how SEM energy savings are estimated

DNV GL has the following recommendations to improve SEM energy savings estimation. These recommendations are based on the activities and analysis completed as part of this evaluation or our experience with energy modeling.

- DNV GL recommends changing the modeling methodology to be degree-day based linear models with linear non-weather variables. The majority of program savings reviewed in the evaluation were determined using models based on the average temperature during the program period. DNV GL recommends utilizing degree-day estimates as independent variables, either heating, cooling, or both. DNV GL also recommends avoiding the use of a polynomial term in any regressions. Use of linear degree-day estimates more closely aligns with the inherent energy consumption processes in commercial buildings and provides more intuitive results. Use of a polynomial regression term may provide a better fit to the data, but is risky when extreme conditions occur.
- DNV GL recommends calculating program savings claims at the conclusion of each resource acquisition year. The current program practice is to forecast the energy savings that will be achieved in the next calendar year, based on consumption during a proportion of the current year. This forecast typically has occurred in the fall, with the first calculation occurring at the end of the initial engagement year. DNV GL recommends that Energy Trust consider calculating savings for the program year at the end of each program year, based on consumption and weather that occurred within that year. This calculation would take about the same time as the current process, but be completed a year later. The benefit to this methodology is that it reduces and, in many cases, eliminates the risk of over- or under-estimating savings. The challenge to this methodology is that the program must operate longer without claiming a full calendar-year savings for a site. However, DNV GL also recommends that the program calculate and document achieved savings at the end of the initial engagement year for the entire engagement year or part of the year. DNV GL observed many cases in which savings began occurring during the initial engagement year that were not claimed to occur until the following year.

Determine how engaged participants are and how well they have implemented and maintained changes in their organizations and buildings over time

- Program participants are engaged in energy management. Participants continue to track their energy use and savings resulting from SEM activities using either MT&R models or some other means. Most organizations find energy consumption and savings tracking tools helpful and valuable.
- Energy champions are the primary leaders of energy teams. These champions are essential to providing proper direction and accountability toward SEM programs. Organizations with energy teams are more engaged in SEM and more motivated to complete SEM activities. Organizational changes are a risk to continued SEM engagement. Organizational changes negatively affected the implementation of the SEM programs at one-half of the organizations where changes occurred.

DNV GL has the following additional recommendations to improve the future evaluations of this program.

- DNV GL recommends enhancing program tracking records to connect sites with capital projects completed through other programs. The program data received did not include any identifiers – other than a participant name and site name – that could be used to connect SEM participants with participation records in Energy Trust's other programs. As a result, identifying capital projects that occurred at participant locations required substantial additional evaluation effort. Given the relative size of the SEM program, DNV GL recommends updating the program tracking records to include a site identifier that aligns with other program databases.
- DNV GL recommends enhancing program tracking data by ensuring that the energy consumption records for accounts and meters associated with participant locations are included in the Energy Trust database. Utility meter data supporting the top down estimation of savings for each site was stored in

unique Microsoft Excel files for each site. However, program tracking data did not include a comprehensive list of the utility accounts and meters associated with program sites. As a result, compiling utility meter data and associated account numbers required substantial additional evaluation effort. DNV GL recommends that Energy Trust consider updating the program tracking data to include the account and meter numbers associated with each participating site. The program also should validate these numbers against the utility data received by Energy Trust before claiming program savings. This change will reduce the evaluation burden and risk on the SEM program.

Appendix A. Sampling cohort results

This table shows the evaluation results for the sampling cohorts.

Sampling Cohort	Resource Fuel	Savings Year	Partici-pating Sites	Claimed Savings	Evaluated Sites	Evaluated Savings	Realization Rate	RP @ 90% CI of RR	Baseline Consumption	% of Baseline Consumption
PY2012 Start	Electric (kWh)	2013	45	5,299,318	15	7,350,568	139%	66%	175,140,706	4.2%
		2014		7,473,477		5,898,544	79%	61%	174,406,536	3.4%
		2015		8,154,081		10,922,778	134%	36%	190,828,116	5.7%
	Gas (therms)	2013	45	126,942	9	-18,452	-15%	700%	4,001,720	-0.5%
		2014		246,460		-38,477	-16%	162%	4,568,476	-0.8%
		2015		366,840		92,950	25%	131%	5,218,181	1.8%
PY2013 Start	Electric (kWh)	2014	38	6,090,749	14	7,854,524	129%	39%	92,387,383	8.5%
		2015		8,523,241		9,382,123	110%	40%	85,230,585	11.0%
	Gas (therms)	2014	38	186,524	5	38,438	21%	208%	1,621,336	2.4%
		2015		205,311		79,654	39%	28%	1,852,570	4.3%
PY2014 Start	Electric (kWh)	2015	53	9,591,894	21	7,152,088	75%	28%	98,710,210	7.2%
	Gas (therms)	2015	53	272,758	8	483,856	177%	30%	4,545,003	10.6%
Corporate	Electric (kWh)	2014	31	460,031	8	2,585,176	562%	56%	51,983,992	5.0%
		2015		690,273		395,218	57%	177%	19,502,371	2.0%
	Gas (therms)	2014	31	63,293	6	154,452	244%	20%	1,137,166	13.6%
		2015		82,057		219,473	267%	15%	1,474,293	14.9%

Appendix B. Detailed results

This appendix provides extrapolated results for specific sampling strata. The following information is provided for each stratum.

- 2013 – Savings acquisition in 2013 calendar year.
- 2014 – Savings acquisition in 2014 calendar year.
- 2015 – Savings acquisition in 2015 calendar year.
- First Eligible Year – Savings acquisition in first year after SEM engagement year.
- First Site Year – Savings acquisition in first year program claimed savings for site, irrespective of fuel.

Estimate	Group	Estimate						Significantly Different From…	
		Modeled Savings	Claimed Savings	Base Consumption	Realization Rate	% of Baseline Consumption	RP of RR @ 90% C.I.	0, Modeled Savings	100%, Modeled / Claimed
Claim Savings Weighted: kWh, Total	2013	7,350,568	5,299,318	175,140,706	138.7%	4.2%	66.3%		Yes
	2014	16,338,244	14,024,257	318,777,912	116.5%	5.1%	30.4%		Yes
	2015	27,852,207	26,959,489	394,271,281	103.3%	7.1%	21.0%		Yes
	First Eligible Year	26,566,297	22,016,853	448,719,815	120.7%	5.9%	20.7%		Yes
	First Site Year	25,988,688	23,317,174	435,134,071	111.5%	6.0%	18.3%		Yes
Claim Savings Weighted: Therm, Total	2013	-18,452	126,942	4,001,720	-14.5%	-0.5%	699.9%	Yes	
	2014	155,938	496,277	7,340,910	31.4%	2.1%	73.4%	Yes	Yes
	2015	846,577	926,966	13,051,759	91.3%	6.5%	25.6%		Yes
	First Eligible Year	645,624	696,487	12,479,337	92.7%	5.2%	35.2%		Yes
	First Site Year	785,931	719,536	12,970,841	109.2%	6.1%	24.1%		Yes
Claim Savings Weighted: kWh, 2012 Cohort	2013	7,350,568	5,299,318	175,140,706	138.7%	4.2%	66.3%		Yes
	2014	5,898,544	7,473,477	174,406,536	78.9%	3.4%	61.2%		Yes
	2015	10,922,778	8,154,081	190,828,116	134.0%	5.7%	35.8%		Yes
	First Eligible Year	8,974,509	5,874,179	205,638,229	152.8%	4.4%	42.4%		Yes
	First Site Year	7,268,178	6,902,184	196,129,882	105.3%	3.7%	38.0%		Yes
Claim Savings Weighted: kWh, 2013 Cohort	2013	0	0	0	0.0%	0.0%			
	2014	7,854,524	6,090,749	92,387,383	129.0%	8.5%	39.3%		Yes
	2015	9,382,123	8,523,241	85,230,585	110.1%	11.0%	40.3%		Yes
	First Eligible Year	7,854,524	6,090,749	92,387,383	129.0%	8.5%	39.3%		Yes
	First Site Year	8,953,853	6,355,994	89,045,390	140.9%	10.1%	33.2%		Yes

Estimate	Group	Estimate						Significantly Different From…	
		Modeled Savings	Claimed Savings	Base Consumption	Realization Rate	% of Baseline Consumption	RP of RR @ 90% C.I.	0, Modeled Savings	100%, Modeled / Claimed
Claim Savings Weighted: kWh, 2014 Cohort	2013	0	0	0	0.0%	0.0%			
	2014	0	0	0	0.0%	0.0%			
	2015	7,152,088	9,591,894	98,710,210	74.6%	7.2%	28.4%	Yes	Yes
	First Eligible Year	7,152,088	9,591,894	98,710,210	74.6%	7.2%	28.4%	Yes	Yes
	First Site Year	7,152,088	9,591,894	98,710,210	74.6%	7.2%	28.4%	Yes	Yes
Claim Savings Weighted: kWh, Corporate	2013	0	0	0	0.0%	0.0%			
	2014	2,585,176	460,031	51,983,992	562.0%	5.0%	56.0%	Yes	
	2015	395,218	690,273	19,502,371	57.3%	2.0%	177.5%		
	First Eligible Year	2,585,176	460,031	51,983,992	562.0%	5.0%	56.0%	Yes	
	First Site Year	2,614,570	467,102	51,248,590	559.7%	5.1%	54.8%	Yes	
Claim Savings Weighted: kWh, Size=Low	2013	2,304,320	2,751,403	83,644,649	83.8%	2.8%	0.0%		
	2014	4,983,834	6,792,796	139,746,395	73.4%	3.6%	8.9%	Yes	Yes
	2015	12,707,006	15,029,767	187,404,563	84.5%	6.8%	3.5%	Yes	Yes
	First Eligible Year	8,731,623	11,774,466	188,885,888	74.2%	4.6%	5.3%	Yes	Yes
	First Site Year	8,173,026	12,301,364	189,589,944	66.4%	4.3%	5.6%	Yes	Yes
Claim Savings Weighted: kWh, Size=Medium	2013	1,401,044	964,290	54,809,307	145.3%	2.6%	195.2%		
	2014	6,097,017	2,928,349	97,697,269	208.2%	6.2%	70.2%		Yes
	2015	10,242,884	4,323,662	127,090,003	236.9%	8.1%	49.2%	Yes	Yes
	First Eligible Year	8,956,396	3,644,209	120,701,838	245.8%	7.4%	39.0%	Yes	Yes
	First Site Year	9,087,902	3,698,252	123,181,741	245.7%	7.4%	39.1%	Yes	Yes
Claim Savings Weighted: kWh, Size=High and Certainty	2013	3,589,190	1,583,625	34,280,361	226.6%	10.5%	107.0%		Yes
	2014	5,067,520	4,303,113	79,666,501	117.8%	6.4%	36.5%		Yes
	2015	4,428,510	7,606,061	74,987,689	58.2%	5.9%	49.2%	Yes	Yes
	First Eligible Year	8,362,463	6,598,179	131,462,186	126.7%	6.4%	45.9%		Yes
	First Site Year	8,065,312	7,317,559	113,248,159	110.2%	7.1%	30.8%		Yes
Claim Savings Weighted: Therm, 2012 Cohort	2013	-18,452	126,942	4,001,720	-14.5%	-0.5%	699.9%	Yes	
	2014	-38,477	246,460	4,568,476	-15.6%	-0.8%	162.0%	Yes	
	2015	92,950	366,840	5,218,181	25.3%	1.8%	131.2%	Yes	
	First Eligible Year	-23,636	173,912	5,195,120	-13.6%	-0.5%	549.6%	Yes	
	First Site Year	78,315	187,546	5,568,297	41.8%	1.4%	82.0%	Yes	Yes

Estimate	Group	Estimate						Significantly Different From…	
		Modeled Savings	Claimed Savings	Base Consumption	Realization Rate	% of Baseline Consumption	RP of RR @ 90% C.I.	0, Modeled Savings	100%, Modeled / Claimed
Claim Savings Weighted: Therm, 2013 Cohort	2013	0	0	0	0.0%	0.0%			
	2014	38,438	186,524	1,621,336	20.6%	2.4%	208.0%	Yes	
	2015	79,654	205,311	1,852,570	38.8%	4.3%	28.5%	Yes	
	First Eligible Year	38,438	186,524	1,621,336	20.6%	2.4%	208.0%	Yes	
	First Site Year	62,456	186,733	1,602,544	33.4%	3.9%	62.1%	Yes	
Claim Savings Weighted: Therm, 2014 Cohort	2013	0	0	0	0.0%	0.0%			
	2014	0	0	0	0.0%	0.0%			
	2015	483,856	272,758	4,545,003	177.4%	10.6%	29.6%	Yes	Yes
	First Eligible Year	483,856	272,758	4,545,003	177.4%	10.6%	29.6%	Yes	Yes
	First Site Year	483,856	272,758	4,545,003	177.4%	10.6%	29.6%	Yes	Yes
Claim Savings Weighted: Therm, Corporate	2013	0	0	0	0.0%	0.0%			
	2014	154,452	63,293	1,137,166	244.0%	13.6%	20.3%	Yes	Yes
	2015	219,473	82,057	1,474,293	267.5%	14.9%	15.5%	Yes	Yes
	First Eligible Year	154,452	63,293	1,137,166	244.0%	13.6%	20.3%	Yes	Yes
	First Site Year	176,918	72,499	1,302,567	244.0%	13.6%	20.3%	Yes	Yes
Claim Savings Weighted: Therm, Size=Low	2013	0	1,744	0	0.0%	0.0%			
	2014	59,002	90,266	669,642	65.4%	8.8%	94.4%		
	2015	306,720	156,082	2,547,801	196.5%	12.0%	6.2%	Yes	Yes
	First Eligible Year	274,731	153,608	2,539,504	178.9%	10.8%	19.5%	Yes	Yes
	First Site Year	303,088	155,022	2,530,498	195.5%	12.0%	10.6%	Yes	Yes
Claim Savings Weighted: Therm, Size=Medium	2013	-14,787	58,012	1,677,375	-25.5%	-0.9%	36.3%	Yes	Yes
	2014	80,089	156,240	4,305,199	51.3%	1.9%	60.6%	Yes	Yes
	2015	145,295	261,489	6,093,745	55.6%	2.4%	87.3%		Yes
	First Eligible Year	181,903	213,029	5,542,097	85.4%	3.3%	59.3%		Yes
	First Site Year	183,481	228,957	5,974,750	80.1%	3.1%	60.1%		Yes
Claim Savings Weighted: Therm, Size=High and Certainty	2013	-3,221	67,186	2,273,918	-4.8%	-0.1%	4006.4%		
	2014	93,121	249,771	3,019,524	37.3%	3.1%	64.8%	Yes	Yes
	2015	420,981	509,395	4,777,351	82.6%	8.8%	47.3%		Yes
	First Eligible Year	199,332	329,850	4,772,087	60.4%	4.2%	98.1%		
	First Site Year	308,292	335,557	4,826,366	91.9%	6.4%	52.7%		Yes

Appendix C. Program staff interview guide

Interview Guide
Energy Trust Commercial SEM Impact Evaluation
Staff Interviews

October 22, 2015

This guide was developed to facilitate an approximately one hour interview with the program staff involved with the delivery of Energy Trust of Oregon's (Energy Trust) Commercial Strategic Energy Management (SEM) program, cohorts 1 to 4 and the corporate participants. The overarching purpose of these interviews is to help the evaluation team to understand the program and the participants so we can conduct a thorough and informed evaluation. The specific objectives for the interview are to:

- Understand the SEM implementation, how it has changed over time, and how the program influences or drives participants' modeled energy savings.
- Understand MT&R models and program savings estimation methods.
- Understand how specific projects were documented, what data is likely to exist, and how we can access data.
- Understand the collection of energy consumption data, who collected the data, and what documentation exists regarding which meters are used at a given facility. Also identify key staff job titles who would potentially have this data.
- Identify typical job titles of staff on the participant's team likely to be knowledgeable about how energy is used and how energy use may have been affected over the course of the program.
- (SEG only) Review other project documents to make sure we understand them correctly. What are they? What is their purpose?
- (SEG only) Identify participants or buildings with unique or unusual circumstances, such as those who needed a special approach to the MT&R or savings estimates or had MT&Rs that were especially difficult to develop.
- Understand the participant engagement and the level of organizational commitment for each cohort and corporate participant. Determine the participation timeline for each organization and the details of their participation in the program.
- Discuss what impact the turnover of the energy champion has had on savings and participation. Investigate how teams manage the transition. What do successful transitions look like? For instance do successful transitions occur with strong team versus teams that delegate most responsibility to the energy champion?
- Investigate how the program can make savings persist past the initial program activity. How can the program make the energy saving culture of participant organizations persist beyond a one-time initiative?
- Understand the SEM ecosystem and how the program's interaction with a subset of building operators may end up influencing other building operating practices at other organizations over time.
- Determine the best methods for participant outreach and recruitment, and whether they are aware of any changes to the customer contacts

Interview Guide

Interview Date: ______________________________

Interviewee Name(s): ______________________________

Interviewer Name(s): ______________________________

Interview Duration: ______________________________

Hello, I'm Jennifer Barnes with DNV GL. As you know, DNV GL has been selected to conduct an impact evaluation of Energy Trust's Commercial SEM program for cohorts 1 to 4 and the corporate participants. I have a series of questions that are intended to help us to understand how you implement the program and to learn about the participants. I anticipate that the interview will take just over an hour.

Throughout the interview, I'm going to use specific terminology to distinguish between the *SEM practices* we hope the participants adopt, like tracking their energy use, forming an energy team and holding regular meetings, and the *O&M or capital measures* that they adopt as an outcome of these SEM practices.

Do you have any questions before we begin?

Background

Can you each describe your role with the program and how long you've been involved?

SEM Implementation

- *Understand the SEM implementation, how it the program influences or drives the models' participants' modeled energy savings.*
- *(SEG only) Review other project documents to make sure we understand them correctly. What are they? What is their purpose?*

Can you please describe the major components of the program as you rolled it out for *cohort 1*? Be sure to note:

- What group activities you conducted
- What activities the participants conducted independently
- Which activities or milestones were required and which were optional

How did each of the steps or activities drive energy savings?

What changes have you made since then and what motivated those changes?

Can you please briefly describe each of the program documents and their purpose? We're aware of the:

- MT&R spreadsheet (this appears to have been named "Dashboard" for the earlier cohorts, correct?)
- MT&R Report
- Executive Sponsor Report (how many of these should there be ideally?)
- Opportunity Action List
- Energy Management Assessment Report
- SEM Action Plan
- Participant report out presentation

We want to understand the implementation timeframes. We have the key dates for cohorts 3 and 4. Is there a schedule for 1 and 2? Corporate?

SEM Continuation

Can you describe the SEM continuation offering?

- What support do you provide?
- Are there structured activities or meetings?
- What is required of the participants?
- How long can a participant remain in continuation?

What results or outcomes do you see from participation in SEM continuation?

- Do participants add more buildings or conduct additional activities?

Can you give us a list of participants in continuation, both current enrollees and past, and the years in which they participated?

MT&R Models, Data Availability and Estimation Methods

- *Understand MT&R models and program savings estimation methods.*
- *Understand how specific projects were documented, what data is likely to exist, and how we can access data.*

- *Understand the collection of energy consumption data, who collected the data, and what documentation exists regarding which meters are used at a given facility. Also identify key staff job titles who would potentially have this data.*
- *Identify typical job titles of staff on the participant's team likely to be knowledgeable about how energy is used and how energy use may have been affected over the course of the program.*
- *(SEG only) Identify participants or buildings with unique or unusual circumstances, such as those who needed a special approach to the MT&R or savings estimates or had MT&Rs that were especially difficult to develop.*

We want to make sure we understand the MT&R models thoroughly. Can you please describe your general approach to building the MT&R models? Be sure to address:

- What were the criteria you used to develop the models? How did you determine when a model was acceptable to use to compute energy savings for the program?
- What independent variables did you routinely test in the MT&R models?
- How did you choose which variables to keep in a given model? Were any statistical tests used (R2, P value, F test)?
- What independent variables were used most frequently?

Can you describe situations where a representative model could not be developed?

What changes have you made to the structure of or approach to the models over time?

Were there participants or buildings with unique or unusual circumstances that made the model difficult to develop?

- Which participants or buildings?
- What were these circumstances?
- How did you overcome them?

Will there be only one MT&R per meter/building or could one model aggregate several? If so, how did you group them?

What's the best way for us to figure out which MT&R to use or which is most current?

Is there anything that tells us about the O&M measures they did?

- Is this ever reflected in the MT&R?
- How did you determine how much savings to deduct for capital measures?
- Did they (or you) try to pencil out the potential savings from identified opportunities?

Can you describe how specific projects were documented, say for a retrofit or capital project? Or for a procedural or instrumentation change?

- What data are likely to exist?
- What are our options to access these data?

Please describe the collection of energy consumption data.

- Who typically collected the energy data, and what sources did they use (eg accounting reports, management reports, bills?)
- What documentation exists regarding which meters are used at a given facility?
- What is the best way to find this information?
- What are the typical job titles of key staff who would potentially have these data?

Can you identify typical job titles of staff on the participant's team likely to be knowledgeable about how energy is used in their facilities and how energy use may have been affected over the course of the program (energy manager, facility manager, maintenance manager, other?).

- Are the site contacts provided likely to include these people?

Participant Engagement with SEM

- *Understand the participant engagement and the level of organizational commitment for each cohort and corporate participant. Determine the participation timeline for each organization and the details of their participation in the program.*

We want to understand the level of organizational commitment of the participant organizations. Can you run through your thoughts on each cohort and the corporate participants? Note any standouts that were either highly motivated or not very motivated. Can you provide us with documentation for the participation timeline for each organization?

Energy Teams

- *Discuss what impact the turnover of the energy champion has had on savings and participation. Investigate how teams mange the transition. What do successful transitions look like? For instance do successful transitions occur with strong team versus teams that delegate most responsibility to the energy champion?*

Has there been much turnover of the energy team members in each cohort?

How does turnover affect the level of the team's engagement and energy savings?

Is the impact different if the energy champion leaves the organization?

Have you seen teams that handle this transition well or better than other teams?

- What are the differences between the two?
- Are there differences in successful transitions between teams that spread the responsibilities equally among members or those where the champion takes on most of the responsibilities?

Savings Persistence

- *Investigate how the program can make savings persist past the initial program activity. How can the program make the energy saving culture of participant organizations persist beyond a one-time initiative?*

What do you think is the best predictor of persistence of the SEM practices? Put another way, what characteristics do organizations need to possess to make them successful at maintaining the SEM practices and energy savings from the O&M measures they've implemented?

Are there participants you believe will be more successful at maintaining SEM practices over time? Why do you think this?

What does the program do to support the persistence of SEM practices? What more could the program do if resources were no object?

Diffusion of SEM Practices

- *Understand the SEM ecosystem and how the program's interaction with a subset of building operators may end up influencing other building operating practices at other organizations over time.*

Are you aware of participants that are leveraging their energy teams and SEM practices to save energy in buildings outside of Energy Trust territory?

- Which participants?
- How are they doing this?
- Are they using the same energy team or forming different teams?
- Are they applying the SEM practices the same way or adopting a subset of activities?

Participant Recruitment

- *Determine the best methods for participant outreach and recruitment, and whether they are aware of any changes to the customer contacts*

How were participants recruited into the program?

Did these methods change over time?

Which methods were most successful?

Are there other methods that you haven't tried that you think may be successful? Why?

Clarifications

What is a policy bonus? (this was noted in a cohort 3 summary document)

Appendix D. Participant interview guide

Interview Guide

Energy Trust Commercial SEM Impact Evaluation

Participant Interviews

1 BACKGROUND

This interview guide is designed to help you complete an approximately 45 minute interview with participants in Energy Trust of Oregon's Commercial SEM program. The objectives of this interview are to:

- Document participant O&M actions and capital measures, both implemented and planned
- Investigate the persistence of energy savings, O&M activities, and SEM practices (energy policies, energy team, tracking, etc.)
    - Investigate how the program can make savings persist past the initial program activity. How can the program make the energy saving culture of participant organizations persist beyond a one-time initiative?
- Determine whether the MT&R has been maintained
- Determine what data are available for the evaluation
- Determine if there are operating conditions or changes to the facility that may affect the energy savings or the validity of the MT&R model
- Investigate possible reasons for large variances in realization rates
- Understand the level of the participant's organizational engagement and commitment
- Discuss what impact the turnover of the energy champion has had on savings and participation. Investigate how teams mange the transition. What do successful transitions look like? For instance do successful transitions occur with strong teams versus teams that delegate most responsibility to the energy champion?
- Understand the SEM ecosystem and how the program's interaction with a subset of building operators may end up influencing other building operating practices at other organizations over time.

2 INTERVIEW RECRUITMENT

Send an advance email to the Energy Champion contact with a follow up phone call two to three days later if the contact doesn't respond via email.

2.1 Recruitment Advance Email

Dear <ENERGY CHAMPION>,

My firm, DNV GL, has been engaged by Energy Trust of Oregon to conduct an evaluation of their commercial Strategic Energy Management program. I understand you previously worked or are currently working with Energy Trust and their contractor, Strategic Energy Group, to use energy more efficiently at your organization. We'd like to schedule a 45 minute telephone interview to follow up on the progress you made during your SEM engagement and gather some information about your facilities and their energy usage. We're interested in speaking with you and the person who is most knowledgeable about these buildings, their operations and how energy is consumed at these buildings. The information collected from this research is critical to help Energy Trust guide future strategic energy management programs.Specifically, we are interested in the following buildings:

- Building 1
- Building 2
- Building 3

We are interested in the following types of activities as identified in your opportunity register <ADAPT FOR EACH PARTICIPANT>:

- HVAC equipment schedule, control and operational improvements, include economizers and fans
- Lighting operations and controls
- Boiler operations and controls.

Please note that all information obtained is confidential. I have attached a letter verifying the study for your review.

May we contact you to arrange a time for the interview?

Best regards,

2.2 Recruitment Telephone Script

Hello, my name is __________________________ and I work for DNV GL, an energy consulting firm. I'm calling on behalf of Energy Trust of Oregon; they are interested in getting your feedback on several aspects of the Commercial SEM program in which you participated in <COHORT YEAR>. This feedback will help Energy Trust improve the programs and services they offer to customers like you.

Are you the person who is most familiar with the operation of your buildings, specifically:

- Building 1
- Building 2
- Building 3

If no, ask: Can you refer me to that person or persons?

If yes, proceed with: I anticipate that the interview will take approximately 45 minutes. This interview is for research purposes only; your feedback will be reported to Energy Trust anonymously and will not affect the status of any Energy Trust projects you are involved with.

When is a good time to hold this interview?

3 INTERVIEW

Interview Date: ____________________

Interviewee Name(s): ____________________

Interviewer Name(s): ____________________

Interview Duration: ____________________

3.1.1.1 Interview Recording

If you record the interview, you must obtain explicit permission from the respondent(s).

3.1.1.2 Confidentiality

If respondents ask, tell them yes, their answers will remain confidential.

3.1.1.3 Legitimacy

If the respondent questions the legitimacy of the interview, give them the evaluation manager's contact information:

Andy Eiden
Evaluation Project Manager
Energy Trust of Oregon
Andy.Eiden@energytrust.org
503.445.2945

3.1.1.4 Introduction

Hello, my name is ________________________ and I work for DNV GL. Thank you for taking the time to speak with me. I am interested in getting your feedback on several aspects of the Commercial SEM program in which you participated in <COHORT YEAR> and the following buildings specifically:

- Building 1
- Building 2
- Building 3

I anticipate that the interview will take approximately 45 minutes. This interview is for research purposes only; your feedback will be reported to Energy Trust anonymously and will not affect the status of any Energy Trust projects you are involved with. Do you have any questions before we begin?

Great. Let's get started. Throughout the interview, I'm going to make a distinction between the SEM practices, like tracking your energy use, forming an energy team and holding regular meetings, and the operations and maintenance changes or capital investments that result in energy savings that you may have made as an outcome of the SEM practices you undertook.

3.1.1.5 MT&R Models and Estimation Methods

Are you currently using the Monitoring, Targeting and Reporting or MT&R model and workbook developed for you to track your energy use?

- If yes ask:
  - How frequently do you update it?
  - Who, specifically, maintains it? Is it a different person for each building?
- If no ask:
  - When was the last time you updated it?
  - Are you using another type of electronic system to track your energy usage over time?

Do you have the data available to update it?

- If yes ask: Can we get it?
- If no ask: What is missing?

Is this the same for all of the buildings we are discussing today?

3.1.1.6 Actions Taken as Part of SEM

For <Building 1>, which actions on your opportunity register have you completed? (Read actions as needed)

Now let's talk about the current operations around the actions completed on your opportunity register, and how they have evolved:

- Which actions do you continue to do?
- Which are no longer part of your operation? Why? (probe each action of the register for the buildings)

Have you completed more items on your opportunity register? If yes ask: Which ones?

Have you added any more items to your opportunity register?

Who has the primary responsibility for implementing the actions? Is it the energy team, building operators, or someone else?

Do you engage the building occupants in any way? If so, what have you done?

<REPEAT FOR EACH BUILDING>

3.1.1.7 Capital Projects

In <Building 1>, I see you completed the following capital upgrades during your program participation at the following buildings:

<POPULATE WITH CAPITAL PROJECTS FROM ENERGY TRUST PROJECT LIST>

Are these projects still operational? (ask for each one).

(Ask measure-specific relevant questions to determine if we could assess the savings from these projects):

What information about the project is available? For example: Size/capacity, load factor, efficiency, schedule, controls, operational system, design drawings, manufacturer model number, equipment cut sheets, etc?

Have you completed any capital upgrades since then? For which building? What were they?

<REPEAT THIS SECTION FOR EACH BUILDING>

3.1.1.8 Understanding Energy Consumption & Facilities Changes

The MT&R model approach assumes that the major changes happening during the program period are because of your participation in SEM. However, other unrelated events can also affect energy consumption, such as changing the hours of operation or the way spaces are used.

For <Building 1>, can you describe any significant changes in how building operations have changed since you began the SEM program, in particular we are interested in <COHORT YEAR> compared to <BASELINE YEAR> ?

Have you:

- Have you changed any of the space uses?
- Has the building use schedule changed?
- Has there been a change in occupancy?
- Have you added or removed energy using equipment?
- Have you retrofit any equipment?
- Have you built additions to the building or the conditioned space?

Are there other ways that you've changed the way you use energy (gas/electricity) since <BASELINE YEAR>?

Do you think the outside air temperature or weather is a major factor in the amount of energy consumed in a day?

Other than the temperature outside, what other factors do you think drive energy consumption at this building?

For these other factors, is there any information recorded that we could use to analyze how these factors impact energy consumption?

Is this building operated differently on weekends? Holidays?

Are there any periods during the year when the facility is operated differently, something like a school being operated differently during summer break?

Is there anything else about this building that you think we should understand in order to accurately model its energy consumption?

<REPEAT FOR EACH BUILDING>

3.1.1.9 O&M Practices/Participant Engagement

[NOTE TO INTERVIEWER: THE REMAINING QUESTIONS ARE NOT BUILDING-SPECIFIC]

Do you still have an executive sponsor for SEM?

- If yes ask: Can you describe their involvement? For instance, do they require regular updates from the team? How frequently?

Did you set an energy policy as a result of your participation in the program? If so, is it still in place?

Did you develop a numeric energy reduction goal? Is it still in place or have you updated it?

Do you still have an energy team?

- If yes ask:
  - How frequently do you meet?
  - What is the relative share of responsibility between the team and the energy champion?
  - Is energy management a documented part of the energy champion's or any energy team members' job description or job responsibilities?

Has the energy champion changed since <COHORT YEAR>?

- If yes ask: Who took over for them? Were they an energy team member or new to the team? How did the transition go? Do you think that this changeover had an impact on savings and your participation in the program?

3.1.1.10 Spillover

Aside from those buildings participating in the program, have any other facilities within your company adopted strategic energy management practices as a result of your participation in the SEM program?

- If yes, ask: How many facilities? What type of buildings are they? Where are they located?

3.1.1.11 Wrap up

Is there anything else you'd like us to know?

Is there anything you'd like Energy Trust to get back to you about? [If they raise an issue here, tell them that we will need to disclose their name/identity to Energy Trust on this particular issue only]

Appendix E. Sampling memo

DNV·GL

Memo to:
Dan Rubado, Energy Trust of Oregon

Copied to:
Jennifer Barnes, DNV GL
Santosh Lamichhane, DNV GL
Andrew W. Wood, DNV GL

Memo No.: 001
From: Julia Vetromile, DNV GL
Mike Witt, DNV GL
Date: October 9, 2015
Prep. by: Mike Witt, DNV GL

Commercial SEM Evaluation, Site Sampling Plan

This memorandum summarizes DNV GL's proposed sampling plan for the evaluation of the Energy Trust of Oregon's Commercial SEM program, as presented by DNV GL on the 25th of September, 2015 and subsequently altered to address concerns.

Sampling Unit, Unit of Analysis and Sample Frame

The Commercial SEM Program actions may target a site's electricity consumption, natural gas consumption, or both. The objective of this evaluation is to provide separate electricity and gas results so DNV GL is proposing to treat each fuel separately. Therefore, the sampling unit and unit of analysis in this study is not the site in general but the fuel-specific savings at a site.

Energy Trust provided DNV GL with the file "SEM Impact Evaluation Participant Sites.xlsx" which shows all sites that participated in some manner in this program, even if no savings were claimed for the location. There are 318 unique sites in the file and therefore 636 potential fuel-specific sites. This file is the initial sample frame for this study. DNV GL will randomly select and ultimately review 80 fuel-specific sites from this frame as part of the impact evaluation.

For this evaluation, the population of interest is the set of fuel-specific sites provided with non-zero savings. If a fuel-specific site had never reported non-zero savings, then it was inferred that the site had never truly participated in the SEM program (with respect to that fuel). Of the total 636 fuel-specific sites listed in the population file, 344 were eliminated for this reason. The remaining 292 fuel-specific sites, including three that reported negative savings, comprise the sample frame. These 292 sites are associated with 27 distinct customers.

Stratification

Stratification is an important and commonly used design feature in most data collection efforts. Stratification refers to the process of partitioning the sample frame into distinct groups (called strata) and sampling is done independently within groups. Stratification is often used to (1) improve precision of the final estimates and (2) control the sample size by subgroups of interest during the analysis. Precision is improved if strata are formed so that the population is relatively homogeneous within each stratum and relatively heterogeneous between strata.

Studies that involve analyzing data that could be highly variable between units often benefit by creating what is referred to as a **certainty stratum**. Generally, those units with the very largest and very smallest values of some size measure are placed in this stratum. This stratum is referred to as "certainty" because all frame units are selected for the data collection effort from this stratum. So the sampling variance associated with estimates created from this stratum is zero (since a census is being taken). A certainty stratum will be defined for this study. This is noted below.

For this study, the sample will be selected independently within stratum defined by the following:

- **Fuel Type**: Electric or Gas
- **Cohort**: 2012, 2013, 2014 and Corporate. For the purpose of this sample design, Cohort numbers 3 and 4 on the original sample frame (both 2014) have been pooled into one stratification variable.
- **Size**: This is defined for those sites not in the Certainty stratum (defined below). Size strata were created within each fuel type by cohort group by looking at the distribution of energy saving among all sites within the group. Those sites in the lower 25% were considered size="Low", those in the 25%-75% group were considered "Medium", and those in the upper 25% were considered "Large". We chose to stratify by size in order to ensure an adequate number of small, medium and large sites (in terms of energy savings) in the respondent sample.
- **Certainty**: Within each of the 27 customers, the sites with the largest energy saving on the sample frame was identified and placed in this certainty stratum. Additionally, a few electric and gas sites with the largest savings, regardless of customer, were placed in this stratum. Ultimately, this stratum contained 30 sites: 28 electric and 2 gas sites. A list of the certainty sites is presented in **Appendix A**.

Note that since the largest site associated with each customer was placed in the certainty stratum, this means all customers on the sample frame will be selected for this study.

Sample Allocation to Strata

After the strata are formed, the next step was to allocate the sample of respondents to each stratum. We required each stratum to have at least 1 respondent allocated to it. And it was assumed the data collection effort for this study will yield a response rate of 80%. This response rate assumption meant:

(1) We are assuming 80% of the 30 certainty sites will be respondents. This meant a sample size of 56 [80 – (.80*30)] remained. This 56 was the sample size allocated to the noncertainty strata.

(2) The desired respondent sample size within each noncertainty stratum was defined to not be larger than 80% of the frame total. This was done in order to ensure an adequate number of frame sites were available in order to achieve the target respondent sample (assuming the 80% response rate).

The noncertainty sample of 56 was allocated to each stratum roughly proportional to the total energy saved among sites in the stratum. A few minor deviations to this rule were applied. For example, as noted above we required at least 1 respondent to be allocated to each noncertainty stratum.

The final sample allocation, the number of sites and the number of customers by stratum are presented in **Table 1**. Note that customers can have sites in multiple rows in this table, so the customer column will not sum to the total row. And note the respondent sample size in the certainty strata are presented as fractional values. This represents the expected number of respondents in these strata.

The results in **Table 1** show that the sample is allocated and this study will proceed with the goal of obtaining 52 respondents for electricity and 28 respondents for gas.

Table 1. Number of Customers, Sites and the Allocated Sample to Each Stratum

Fuel Type	Certainty	Cohort	Size	Number of Customers on the Sample Frame	Number of Sites on the Sample Frame	Desired Number of Respondents
Electric	Noncertainty	2012	Low	3	10	1
Electric	Noncertainty	2012	Medium	6	18	3
Electric	Noncertainty	2012	High	3	9	4
Electric	Noncertainty	2013	Low	3	9	1
Electric	Noncertainty	2013	Medium	4	16	3
Electric	Noncertainty	2013	High	3	8	5
Electric	Noncertainty	2014	Low	4	10	1
Electric	Noncertainty	2014	Medium	9	20	2
Electric	Noncertainty	2014	High	5	10	5
Electric	Noncertainty	Corporate	Low	1	7	1
Electric	Noncertainty	Corporate	Medium	1	14	2
Electric	Noncertainty	Corporate	High	2	7	2
Electric	Certainty	2012	Certainty	7	7	5.6
Electric	Certainty	2013	Certainty	5	5	4
Electric	Certainty	2014	Certainty	12	13	10.4
Electric	Certainty	Corporate	Certainty	3	3	2.4
Electric	Total	Total	Total	27	166	52.4
Gas	Noncertainty	2012	Low	2	8	1
Gas	Noncertainty	2012	Medium	4	15	3
Gas	Noncertainty	2012	High	4	7	5
Gas	Noncertainty	2013	Low	3	6	1
Gas	Noncertainty	2013	Medium	4	12	1
Gas	Noncertainty	2013	High	4	6	2
Gas	Noncertainty	2014	Low	5	9	1
Gas	Noncertainty	2014	Medium	8	16	2
Gas	Noncertainty	2014	High	3	8	4
Gas	Noncertainty	Corporate	Low	2	10	1
Gas	Noncertainty	Corporate	Medium	1	18	2
Gas	Noncertainty	Corporate	High	2	9	3
Gas	Certainty	2013	Certainty	1	1	0.8
Gas	Certainty	2014	Certainty	1	1	0.8
Gas	Total	Total	Total	21	126	27.6

Sample Selection

As noted previously, all sites within the certainty strata will be selected for the study. And sites in the noncertainty strata will be selected using a simple random sample selection methodology. Our ultimate objective is to obtain the desired number of respondents in each noncertainty stratum. So the sample selection will proceed by putting the entire frame in a random order and working sites within each stratum in the pre-defined order until the specified number of respondents is attained within each stratum.

Expected Precision

Table 2 shows the expected sample size and precision for various groups of interest. Rows in this table represent domains of interest defined by fuel type, cohort group and design strata. The columns in this table represent the expected sample size and precision for estimates of savings by calendar year. Note that:

(1) The 2014 cohort will only be contributing sample to the 2014 calendar year estimates since by definition, this cohort began in 2014.

(2) The 2013 cohort began in 2013 and can contribute to the estimates of savings in both calendar years 2013 and 2014. We assumed the energy savings estimate would be cumulative, so that energy saved in calendar year 2013 for this cohort is equal to the 2013 estimated value on the frame and energy saved in calendar year 2014 would equal the frame value associated with 2013 plus the frame value associated with 2014.

(3) Similarly, the 2012 cohort began in 2012 and can contribute to the estimates of savings in 2012, 2013 and 2014. So similar to the 2013 cohort, we assumed the energy savings estimate would be cumulative. So that energy saved in calendar year 2012 for this cohort is equal to the 2012 estimated value on the frame, the energy saved in calendar year 2013 is the sum of the frame 2012 and 2013 values and the energy saved in 2014 is the sum of the frame 2012, 2013 and 2014 values.

As noted above, sites are only included in the sample frame for a given year if there are incremental savings (positive or negative) claimed in that year. Annual savings estimates for a given program year will be based on the incremental savings achieved. Total savings across several years will be cumulative given an expected measure life of 3 years.

The precision estimates suggest that estimates of savings for total electricity would have a relative precision of 9.6% and the precision would range from 13% to 23% for the cohort specific estimates. And the precision for total gas estimates would be 13.5%, with a range of 19% to 35% for the cohort estimates.

Precision estimates are reported in **Table 2** for groups defined by design stratum, but note that estimates will not be computed for each stratum during the analysis. These design strata are defined for sample selection purposes only and are not meant to be analytic domains of interest in their own right.

The precision estimates presented in **Table 2** assume the variability in the final adjusted saving estimates from this study will be approximately equal to the variability observed with the current estimates of savings as indicated on the sample frame.

DNV GL also plans to look at the time trend of savings across cohorts for sites involved in continuation and those not involved. The goal of this assessment is to better understand persistence of savings and how they decay or increase over time. We expect there to be a difference between those involved in continuation and those not. The variability of the growth or decay is savings with or without continuation is not known, and thus the precision of the analysis cannot be determined. However, the lower bound (best case) would be the persistence of the expected precision for each cohort (for example, 14% for the 2012 cohort).

Table 2. Frame Number of Sites, Sample Size and Expected Precision

Group	All Years			Program Year 2012			Program Year 2013			Program Year 2014		
	Frame Number of Sites	Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence
Total, Electricity												
Total	166	52	9.6%	27	10	14.1%	97	31	10.9%	166	52	9.6%
Cohort, Electricity												
Cohort 2012	44	14	13.0%	27	10	14.1%	42	13	14.3%	44	14	13.2%
Cohort 2013	38	13	15.0%	0	0		33	12	17.8%	38	13	15.0%
Cohort 2014	53	18	20.7%	0	0		0	0		53	18	20.7%
Corporate	31	7	23.3%	0	0		22	6	41.2%	31	7	23.3%
Design Strata, Electricity												
Noncertainty, Cohort 2012, Savings=Low	10	1	90.9%	2	0	252.1%	9	1	151.1%	10	1	123.0%
Noncertainty, Cohort 2012, Savings=Medium	18	3	30.9%	12	2	37.6%	17	3	32.5%	18	3	30.9%
Noncertainty, Cohort 2012, Savings=High	9	4	28.2%	7	3	26.2%	9	4	31.7%	9	4	28.2%
Noncertainty, Cohort 2013, Savings=Low	9	1	144.7%	0	0		6	1	197.0%	9	1	144.7%
Noncertainty, Cohort 2013, Savings=Medium	16	3	36.6%	0	0		15	3	48.3%	16	3	36.6%
Noncertainty, Cohort 2013, Savings=High	8	5	22.4%	0	0		7	4	31.5%	8	5	22.4%
Noncertainty, Cohort 2014, Savings=Low	10	1	81.4%	0	0		0	0		10	1	81.4%
Noncertainty, Cohort 2014, Savings=Medium	20	2	68.3%	0	0		0	0		20	2	68.3%
Noncertainty, Cohort 2014, Savings=High	10	5	34.7%	0	0		0	0		10	5	34.7%
Noncertainty, Corporate, Savings=Low	7	1	59.5%	0	0		4	1	37.3%	7	1	59.5%
Noncertainty, Corporate, Savings=Medium	14	2	35.3%	0	0		9	1	88.1%	14	2	35.3%
Noncertainty, Corporate, Savings=High	7	2	24.9%	0	0		6	2	82.9%	7	2	24.9%
Certainty, Cohort 2012	7	6	16.3%	6	5	16.8%	7	6	17.0%	7	6	16.3%
Certainty, Cohort 2013	5	4	23.0%	0	0		5	4	17.9%	5	4	23.0%
Certainty, Cohort 2014	13	10	27.4%	0	0		0	0		13	10	27.4%
Certainty, Corporate	3	2	46.3%	0	0		3	2	59.4%	3	2	46.3%

Group	All Years			Program Year 2012			Program Year 2013			Program Year 2014		
	Frame Number of Sites	Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence	Frame Number of Sites	Expected Sample	Expected Precision at 90% Confidence
Total, Gas												
Total	126	28	13.5%	13	5	33.0%	78	17	21.4%	126	28	13.5%
Cohort, Gas												
Cohort 2012	30	9	25.7%	13	5	33.0%	26	7	33.3%	30	9	25.7%
Cohort 2013	25	5	27.5%	0	0		23	4	33.9%	25	5	27.5%
Cohort 2014	34	8	19.1%	0	0		0	0		34	8	19.1%
Corporate	37	6	34.8%	0	0		29	5	35.9%	37	6	34.8%
Design Strata, Gas												
Noncertainty, Cohort 2012, Savings=Low	8	1	125.0%	1	0	0.0%	8	1	137.1%	8	1	125.0%
Noncertainty, Cohort 2012, Savings=Medium	15	3	64.0%	7	1	68.1%	13	3	71.4%	15	3	64.0%
Noncertainty, Cohort 2012, Savings=High	7	5	13.1%	5	4	20.8%	5	4	17.6%	7	5	13.1%
Noncertainty, Cohort 2013, Savings=Low	6	1	61.7%	0	0		4	1	77.3%	6	1	61.7%
Noncertainty, Cohort 2013, Savings=Medium	12	1	92.4%	0	0		12	1	98.6%	12	1	92.4%
Noncertainty, Cohort 2013, Savings=High	6	2	53.6%	0	0		6	2	75.4%	6	2	53.6%
Noncertainty, Cohort 2014, Savings=Low	9	1	63.0%	0	0		0	0		9	1	63.0%
Noncertainty, Cohort 2014, Savings=Medium	16	2	59.3%	0	0		0	0		16	2	59.3%
Noncertainty, Cohort 2014, Savings=High	8	4	24.6%	0	0		0	0		8	4	24.6%
Noncertainty, Corporate, Savings=Low	10	1	86.5%	0	0		6	1	136.6%	10	1	86.5%
Noncertainty, Corporate, Savings=Medium	18	2	50.3%	0	0		15	2	65.6%	18	2	50.3%
Noncertainty, Corporate, Savings=High	9	3	48.8%	0	0		8	3	44.1%	9	3	48.8%
Certainty, Cohort 2013	1	1	0.0%	0	0		1	1	0.0%	1	1	0.0%
Certainty, Cohort 2014	1	1	0.0%	0	0		0	0		1	1	0.0%

Appendix F. Model diagnostics summary memo

Memo to:		**Memo No.:**	002
Andy Eiden & Dan Rubado,		**From:**	Andrew Wood, DNV GL
Energy Trust of Oregon			Julia Vetromile, DNV GL
		Date:	March 9, 2016
Copied to:		**Prep. by:**	Andrew Wood, DNV GL
Jennifer Barnes, DNV GL			

COMMERCIAL SEM EVALUATION, MODEL DIAGNOSTICS SUMMARY

This memorandum summarizes DNV GL's review of MT&R models sampled for impact evaluation. The final sample design was provided in a site sampling plan memo, dated October 9, 2015. Since that time, DNV GL has completed a review of the models associated with the sampled savings claims.

1.1 Review of Sample Selected

This impact evaluation sampled a portion Commercial SEM program sites for review. The program years being evaluated include 27 unique participants which manage a total of 292 unique program sites (gas and electric savings are considered unique models, so one location may have two sites). From this population, the evaluation sampled 30 program sites in a certainty stratum and 56 program sites through additional stratified random sampling. These 86 program sites consist of the sampled reviewed. The sampled program sites are spread across all cohorts and both fuels. This table shows the distribution of sample points by fuel, cohort, and stratum type.

Table 1: Evaluation sample summary

Sampled Fuel & Cohort	Certainty	Random	Grand Total
Electric	**28**	**30**	**58**
1	7	8	15
2	5	9	14
3	7	4	11
4	6	4	10
Corporate	3	5	8
Gas	**2**	**26**	**28**
1		9	9
2	1	4	5
3	1	3	4
4		4	4
Corporate		6	6
Grand Total	**30**	**56**	**86**

1.2 Program Staff Interview Results

DNV GL interviewed SEG staff on October 28th, 2015. Our interviewed covered multiple topics. The following information learned during the interview is important context for this memo:

- The analysis template has evolved over time as SEG has worked to improve their process.
- In some cases, adjustments to the baseline were made.
- Cohort 1 used weather. SEG started testing other independent variables in Cohort 2. The current intention is to find the independent variables that drive consumption.
- Acquiring and updating models with data is laborious for the end user.
- Some participants use Energy Expert which models facilities differently. There are benefits and drawbacks to the system.

1.3 Model Types

The program has used three different modelling methods in its lifetime. The method is based on the implementer, data available, and the participant's preference. The following describes the three model types reviewed:

- MT&R – The SEG MT&R based models were used most often. Microsoft Excel was used as the modelling platform. Typically, these models used monthly consumption and other metrics to develop a regression model from a baseline period. This model then predicts consumption in future periods. Actual consumption is compared to predicted consumption to determine if savings is occurring. DNV GL was able to fully review these models.
- Energy Expert – Energy expert is a platform offered to Portland General Electric (PGE) customers. The platform is online and integrates with PGE's hourly consumption data for the facility. The software uses the baseline period to estimate consumption for each hour of the week across different temperature bins. For example, the model will estimate unique consumptions for Monday 18:00, 50-55°F and Monday 18:00, 55-60°F. DNV GL has had the opportunity to view the online platform, but could not fully review the models.
- Corporate Cohort – The corporate cohort was implemented by Ecova. According to their methodology document, Ecova used the Georgia Tech EnPI tool developed for Superior Energy Performance to generate consumption models. They then summarized the modeled consumption in a single spreadsheet at all locations for one participant. The models typically used CDD to model electric consumption and HDD to model gas consumption. However, some models used other parameters such as average temperature or a non-weather variable such as customer count at a restaurant. Actual consumption is compared to modeled consumption to estimate program savings. DNV GL received summary workbooks associated with the sampled sites to review, however not all modeling parameters could be reviewed.

The table on the following page summarizes the use of these different models across the sample.

Table 2: Model types in sample by sampled site

Sampled Fuel & Cohort	SEG MT&R	Energy Expert	Ecova Model	Grand Total
Electric	**41**	**9**	**8**	**58**
1	12	3		15
2	8	6		14
3	11			11
4	10			10
Corporate			8	8
Gas	**22**		**6**	**28**
1	9			9
2	5			5
3	4			4
4	4			4
Corporate			6	6
Grand Total	**63**	**9**	**14**	**86**

1.4 MT&R Model Review

This section documents our review of the MT&R models developed by SEG for the program. DNV GL reviewed all MT&R models received in our sample. In this section we document our conclusions based on the review. We use the tables and charts to communicate the results of our review. After each table or chart we document the evaluations conclusions that are supported by the chart or table above. Additional charts are provided at the end of this memo.

1.4.1 Basic MT&R Structure

Overall, the basic MT&R structure is the same across the program and years. The evaluation found that models for later cohorts were more automated and were often easier to follow than the models for the original cohort. The basic pieces of the models are:

- Basic project information
- Input data: consumption, weather, other independent variables
- Event log
- Regression model
- Savings estimation

1.4.2 MT&R Review Conclusions

The table on the following page summarizes the independent variables used in the models reviewed, including the non-MT&R models. Some sampled sites contained more than one model, so the total number of models reviewed is higher than the number of sites originally sampled.

Table 3: Independent variables used in reviewed models

Model Type	Cohort	Fuel Type	Total Models	Models using Avg. Mean Temp.	Models using Temp^2	Models using Non-Weather Indicator	Models using CDD	Models using HDD	Unknown Variables
SEG MT&R	1	Electric	20	20	18	2			
		Gas	16	16	15				
	2	Electric	11	11	10	3			
		Gas	6	6	4	2			
	3	Electric	11	10	7	3			
		Gas	4	4	2				
	4	Electric	10	8	6	5			
		Gas	4	4	3				
	1-4	**Electric Subtotal**	**52**	**49**	**41**	**13**			
	1-4	**Gas Subtotal**	**30**	**30**	**24**	**2**			
SEG Energy Expert	1	Electric	2						2
	2	Electric	6						6
Ecova Corporate	5	Electric	8			1	5	2	3
		Gas	6					6	
Grand Total			**104**	**79**	**65**	**16**	**5**	**8**	**11**

Evaluators Conclusions:

1. SEG's MT&R's relied heavily on average temperature, as either a linear or squared term. In our opinion, this does not represent the underlying process. A temperature term (linear or squared) cannot account for both the energy required heat and cool buildings. We prefer cooling degree day (CDD) and heating degree day (HDD) specifications as these specifications result in easy to interpret model parameters. Use of average temperature can result in intercepts that are difficult to interpret and in some cases represent the maximum load. The lack of intuitive parameters makes the models difficult to easily assess and compare.

   a. For example, a heating dependent gas model will have an intercept that is the consumption at 0°F. The coefficients for the temperature terms then adjust the consumption up or down from this value.

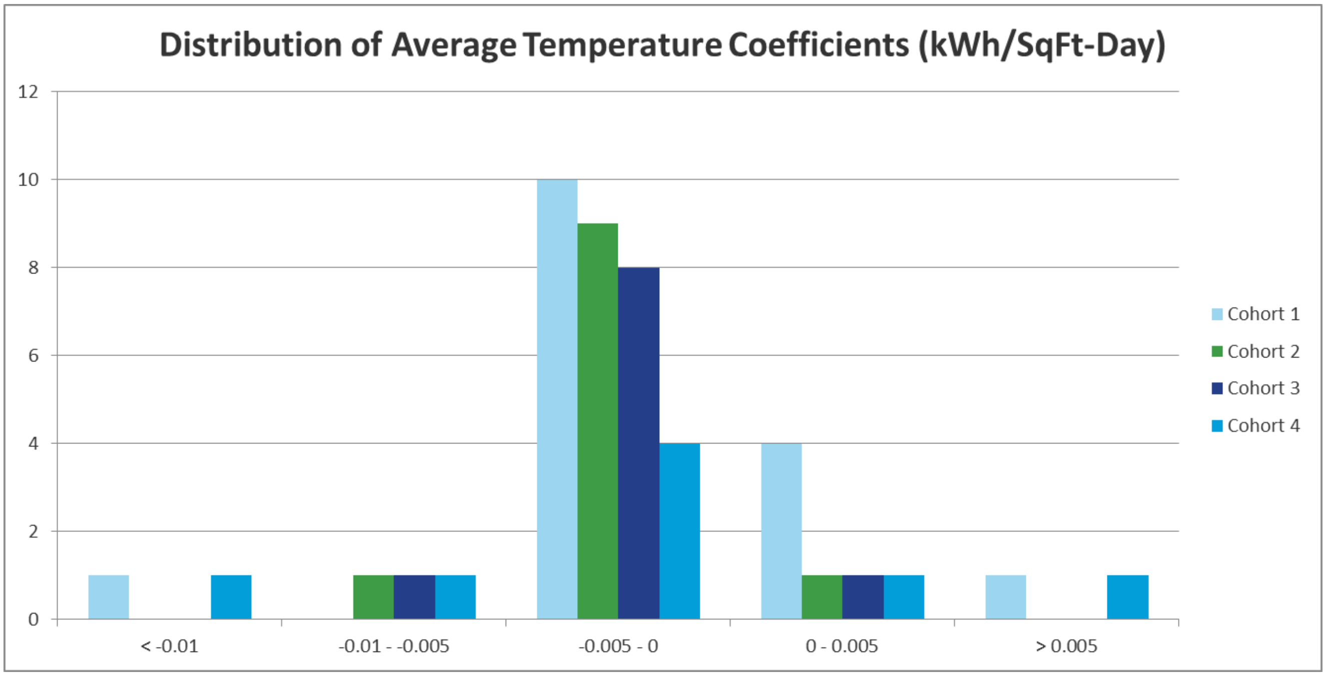
   b. Additionally, use of average temperature makes it more difficult to model consumption that increases both during hot and cold periods of the year.

2. Only 15 of the 82 MT&Rs reviewed use a non-weather indicator such as holidays. If non-weather parameters do drive consumption at any facility, then it is likely its model isn't accurately estimating consumption.

3. Ecova's models utilized CDD and/or HDD for many models. Average temperature was used for some models. For the three corporate sites, the models were apparently completed using a spreadsheet

model developed by Georgia Tech (EnPI tool) or a proprietary tool. The EnPI tool is not provided, but the model equations and data are available in separate tabs of a large summary spreadsheet for one customer. It is currently unclear how three of the models were completed, but they are expected to have used CDD and possibly HDD as input variables in most cases. Gas models typically used HDD.

This chart is a frequency distribution of the linear coefficient used in MT&R models. Chart is only for kWh/SqFt-Day-°F. A similar chart for gas is included at the end of the memo.

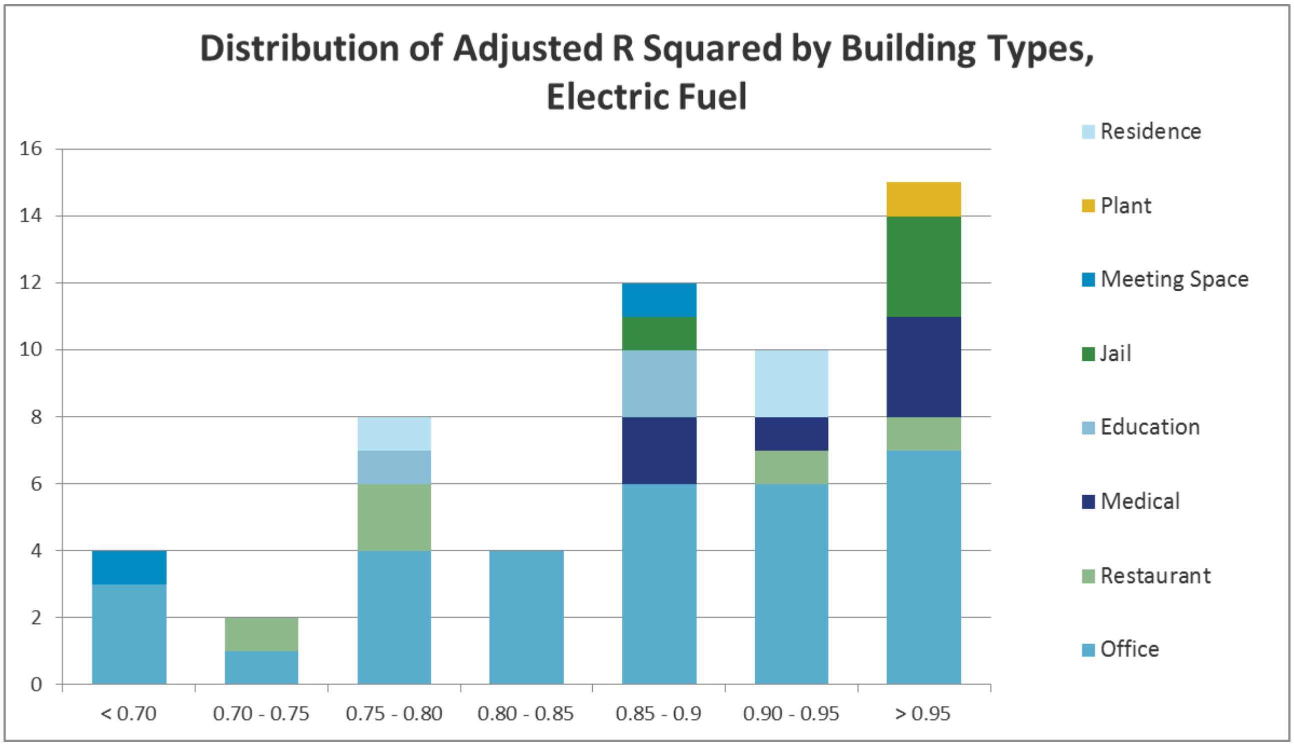


Evaluators Conclusions

1. The negative intercept coefficients continue to show how the use of average temperature creates non-intuitive coefficients. Intuitively, we expect consumption to go up when it gets warmer (more cooling) and expect some consumption to go up when it gets colder (more heating). The coefficient could therefore be negative (driven by heating) or positive (driven by cooling), but it would be impossible to accurately model this expected profile with one linear coefficient. To account for this, many of these models also use a Temp^2 term. This squared term becomes more and more dominant as the temperature gets higher, therefore making the linear term more influential at lower temperatures. The result is often that the linear coefficient is negative as shown in the chart. If a HDD and CDD based model were used, then unique coefficients would exist to adjust consumption during warm and cold periods. A squared term would then only need to be used if consumption changes non-linearly with changes in temperature.

2. Evaluators were hoping that these coefficients could be compared against one another, but comparison is challenging since the model specifications change from one building to the next.

This chart shows the frequency distribution of the R^2 terms associated with each electric model reviewed. A similar chart for gas models is included at the end of the memo.



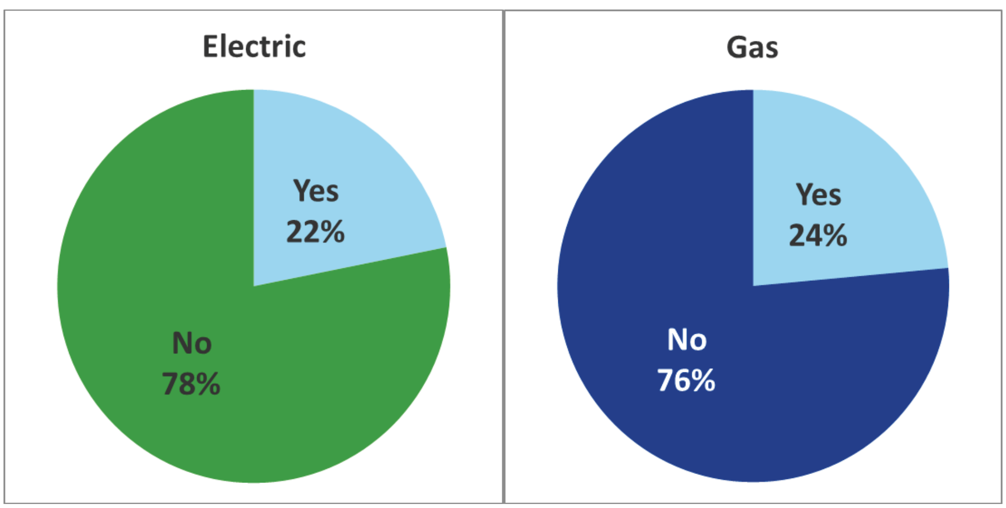
Evaluator Conclusions

1. Overall, DNV GL finds the R-Squared terms for these models to be higher than expected and too closely skewed to the high end of the range. This is strictly a judgment based only on our experience modeling consumption outside of this project. The same conclusion exists for the gas models. DNV GL does expect gas models to fit consumption better than electric due to the reduce number of gas end-uses, but the values are still higher than expected.
2. The R-squared values in this range implies at least one of the following exists;
   a. Another part of the model specification captures activity variation (e.g., site activity or school days).
   b. Activity is very constant over the evaluation period so that temperature is the only factor that drives variation in energy use.
   c. The model is over fit. An over fit model captures relationships between the data but does not capture the underlying process. Over fit models do not produce reliable estimates for evaluation (or forecasting) purposes.
   d. Cherry picking of input data. A few of the MT&Rs reviewed included instructions to adjust the baseline period to maximize the R-Squared term.

The concern about the selection of input data for the baseline model is supported by the first two of four instructions found in one of the Cohort 2 models. The instructions suggest adjusting the dates to find the best R2 value and then state that an R2 value below 0.75 was considered invalid. These instructions only exist in some files.

Single Regression Steps	
1	In the "Baseline Graphing Table" change the "Reference Number" in cell T101 to choose a period of at least 12 months that gives the best R2 number in the "KWH use vs Temperature" graph located just above the Baseline Graphing Table. To extend the range past 12 months simply perform the copy/drag function to extend and add more dates. Try different ranges of dates to find the best R2
2	Utilizing Excel's built in regression function run a regression of the "kWh/day" data and each of the variables to determine if the variable (typically Temp., Temp2 and Holidays) have a significant correlation. If the P values are over 0.05 re-run the regression removing one of the variables until the you have at least one variable remaining in use and have a P value less than 0.05 and an R2 value greater than .75

These two charts show the percentage of MT&R models for each fuel that used a non-standard baseline period. For both fuels, slightly less than one-quarter of models included non-standard baseline periods.



Evaluation Conclusions:

1. Modification of a baseline period is necessary if consumption during one metering period included unusual and anomalous consumption. The program must document why a non-standard baseline is used in the model
2. However, potential adjustment the baseline period to maximize the R2 term is a concern. The resulting model is a good fit, but not necessarily a predictor of annual consumption. Aligning with the table shown above, Cohort 2 had the highest rate of modified baselines.
3. MT&R files inconsistently document the reason substantiating any baseline adjustments. It has therefore been difficult for the evaluation to assess if the baseline adjustment was necessary.

1.5 Energy Expect Review

DNV GL attempted to review the input parameters associated with Energy Expect models. SEG provided DNV GL with an Energy Expert login for one site so the evaluation could learn more about the system. SEG staff also spent 30 minutes via Goto Meeting showing DNV GL how they use Energy Expert during the program.

Unfortunately, little information can be extracted from the platform other than charts either showing consumption or comparing consumption to the model's estimate. DNV GL is unable to observe the baseline period used to develop the estimates shown or any of the model's assumptions. The interval data used by Energy Expert cannot be extracted either.

Overall, the Energy Expert model appears robust and user friendly, but has little functionality for this evaluation. The evaluation will not be able to directly assess the model's validity without receiving the interval data for each model. The baseline periods used cannot be independently verified without support from the software providers. Assuming no further information is received, the evaluation will use the baseline periods documented in the final annual reports.

1.6 Corporate Cohort Model Review

DNV GL reviewed three unique excel files supporting the savings claimed for all sampled corporate cohort sites. In general, these files included: instructions on how to operate the file, input tabs for weather and consumption, and output tabs showing and in some cases calculating savings. None of the files reviewed showed the steps used to create the baseline model. A separate document, *Ecova Modeling and Savings Calculation Process.pdf* describes the process.

Overall, the modelling approach is similar to the MT&R files reviewed. The approach balances simplicity in approach with a desire for accurate savings estimates. The baseline models appear to be developed using the Georgia Tech EnPI tool or another proprietary tool, which were used to generate the regressions equation. In one case, the model equations and data are clearly identified for each site. In others, a proprietary tool provides results.

Information on actions completed is limited. For one participant, a detailed list of capital projects incented through Existing Buildings programs is provided. Final reports for each participant list energy conservation actions taken overall, but not by facility. No activity register exists to document any SEM actions taken by the facility that support the savings claimed. For one participant, site specific activity registers were initiated by site, but completion of activities was not documented.

The evaluation did find items of concern in these models. One example is the lack of non-weather parameters when there is a clear need for one. A college campus' dorms were aggregated together for electricity modelling. In this case, modelling was based strictly on the average temperature during the billing period. Highest consumption occurred during the colder winter months while the lowest consumption occurs during the summer. The temperature based model used a high intercept coefficient and a negative temperature coefficient. The result is fairly smooth wave like forecast of consumption. However, actual consumption shows a dramatic reduction in consumption during the summer. Without any more information, the quick conclusion is that this period aligns with a reduction in occupancy for the summer session. A review of all the dorm related consumption data in the file shows that this pattern existed every year. The appropriate variable in this case may have been "# of weeks in session" during the billing period. Use of weeks will account for the variation in length of breaks throughout the year. While this is just one example of our concern, it shows that concerns with the modelling approach are not exclusive to the MT&R files.

1.7 Summary of Evaluation Diagnostics

DNV GL has come to the following conclusions based on our review of the MT&R models.

- The MT&R models were created to balance accuracy of energy modeling with user simplicity. Since the models are expected to be operated by building operator or energy champion, further sophistication will not guarantee an increase in accuracy especially since this will place greater burden on the building operator. However, the evaluation is concerned by the simplicity of the models and expects that not all models are accurately modeling the drivers of facility consumption.
- No significant difference in modeling methodology exists across the four SEG served cohorts or the building types involved in the program. While the MT&R spreadsheets were updated with improvements for each cohort, the modeling methodology is consistent.
- The MT&R models almost always use a linear average temperature term and often use a squared average temperature term. This is the simplest approach possible and expected to result in the least user created modeling error. However, use of these terms does lead to non-intuitive regression coefficients that prevent simple review or comparison.
- There was limited use of non-temperature independent variables. In order for these variables to be useful for the program, the data associated with them must be collected and reported regularly. Perhaps the program assumed that the increase in time burden to record these parameters would have reduced operator commitment to the program. Non-weather variables were used in cases where the data was easily available, such as annual holidays, or deemed necessary for modeling, such as a waste water treatment plant.
- In some cases, the baseline period appears to have been chosen to maximize the R2 term associated with the baseline regression model. There is no evidence that this occurred in any specific model or systematically throughout the program, but the evaluation believes the R2 terms are higher than expected for simple commercial buildings. DNV GL is therefore concerned that the models may artificially represent a good fit to the baseline data and do not accurately forecast energy consumption.
- Energy Expert creates a more robust model tied to the hour of each weekday and the variation in temperature during each hour. This modeling method is similar to many demand response models and is expected by the evaluation to provide a more accurate forecast of consumption than a monthly consumption based temperature regression model. Unfortunately, little information can be extracted from the software platform without support of the utility company or software provider.

1.8 Evaluation Next Steps

The evaluation team recommends deviating from the original project plan to only adjust some models, and instead independently calculate energy savings to test the validity of the program models and overall savings estimates. This recommendation is primarily based on our concerns regarding overall program model simplicity and methodology. By developing an independent estimate for each model, we can evaluate the site-level models, their methodology, and the overall program savings estimated.

1.8.1 Proposed Model Specification

DNV GL recommends using the following specification for the development of an evaluation based estimate of savings. This outline describes the general approach that we are taking toward completing these objectives. Overall, there two analytical objectives for this analysis:

1. Create an independent site-level model of energy consumption to test if an independent model provides similar savings estimates.
2. Calculate program savings based on the sampling approach implemented. This will test if the program's estimate of savings is significantly different than our savings estimate.

Site-level energy consumption model specification

The following is our proposed specification for independently testing the program's savings claims.

- All models will be developed with monthly billing data. This will be a change in interval frequency for Energy Expert sites and one other site estimated using weekly consumption. However, use of only monthly data allows for a consistent and efficient approach to test program modelling.
- The baseline period will be defined as either the twenty four months or the twelve months prior to the cohort kick-off meeting, depending on data availability.
- Cooling and/or heating degree days. We prefer model specification that describes energy consumption as a response to temperature above and below a basis temperature. We also prefer using models with a variable degree basis as we recognize that not all building operators use the same set points and that not all buildings have similar building envelopes.
  - Hourly weather will be used to determine the degree days in the monthly billing period. Hourly weather allows for an estimation of the cooling load and heating load that can exist within the same day. Since commercial facilities often operate with set-back schedules (often implemented by this program) and have substantially different internal loads during the day than at night, hourly weather based degree days are expected to provide a much better independent variable for predicting energy consumption.
- When possible, models will include a non-weather "site activity" independent variable. Our interviews with facility personnel and visits to sites will attempt to identify non-weather parameters that drive facility consumption. If possible, we collect the data associated with the identified parameters associated with the analysis period. Example parameters include the number of working days, observed holidays, building occupancy, and customer or guest counts. Similar to the program, acquisition or accurate data covering the entire analysis period will be a challenge.
- Estimate models following a PRSIM-like approach. We estimate the model on the pre-period data. We estimate 5 unique models with the following specifications:

1. Site activity plus cooling and heating degree days with a separate basis temperatures for cooling and heating
2. Site activity plus cooling and heating degree days with a common basis temperature for cooling and heating
3. Site activity plus cooling degree days
4. Site activity plus heating degree days
5. Only site activity

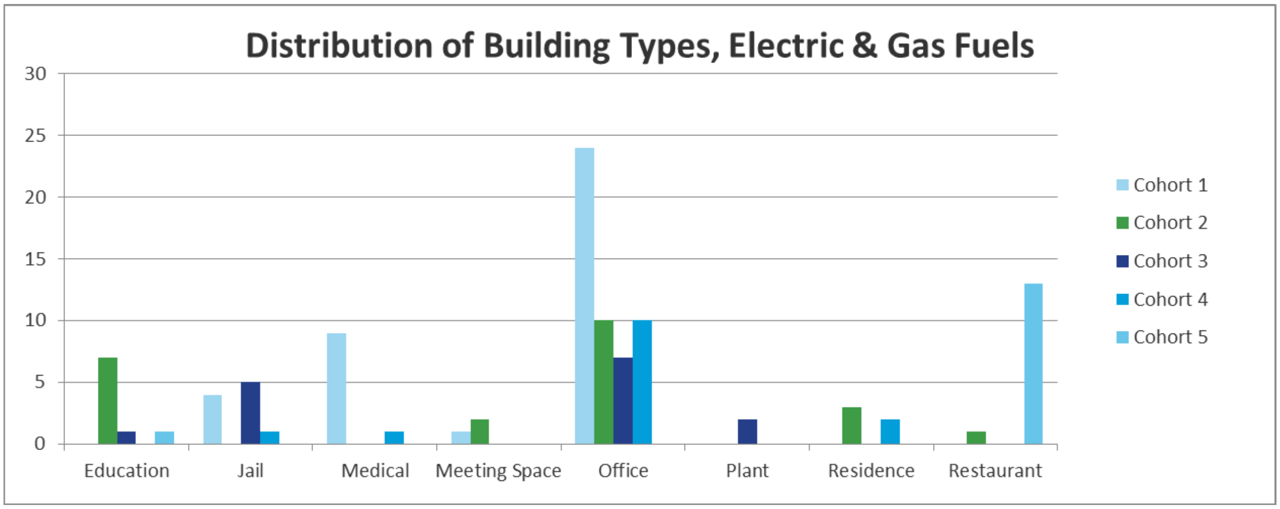
- Select the model with the best fit among models with plausible coefficients. Some estimation results will likely show results such as a basis temperature over 90 °F or have a negative coefficient on a degree day term. We select the model with the highest fit (as measured by adjusted R^2) after eliminating implausible models.
- Our last step is to calculate a weather-adjusted savings estimate in the post-period. This step includes the following:
  1. Apply the baseline site-level model for each site to the actual weather during each subsequent year. This results in estimated values for each month in the post-period. We define the difference between the estimated values and observed values as the residuals.
  2. Calculate annual site-level savings as the cumulative sum of residuals during a calendar year.

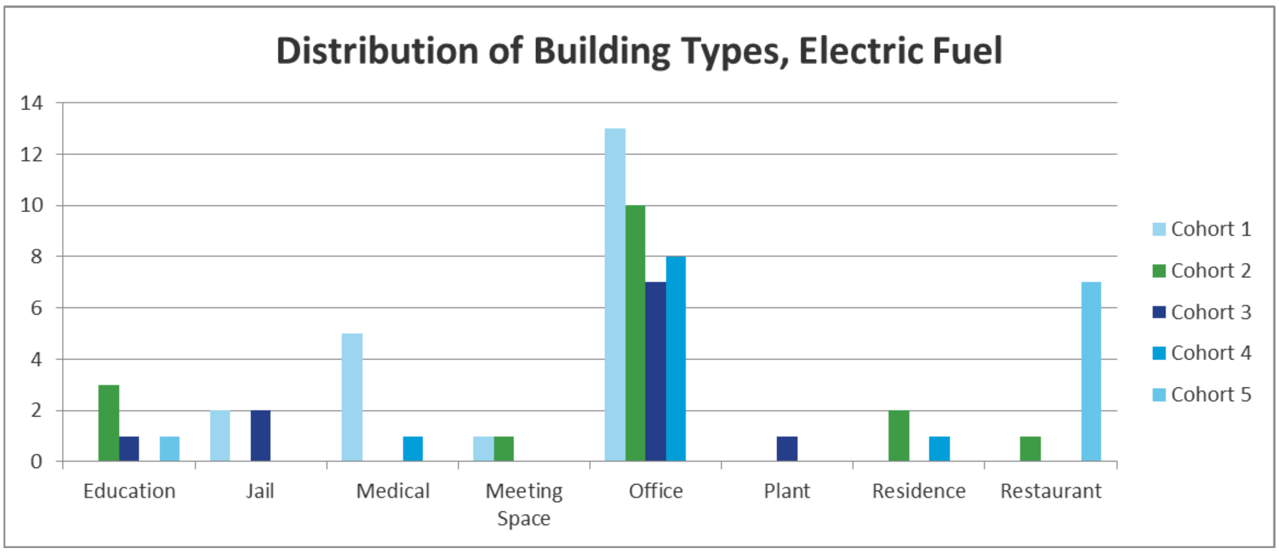
The savings calculated using the above evaluation model will be compared to the incremental and total site savings estimated by the program for the three years following an incremental savings claim or through 2015, depending on which comes first.

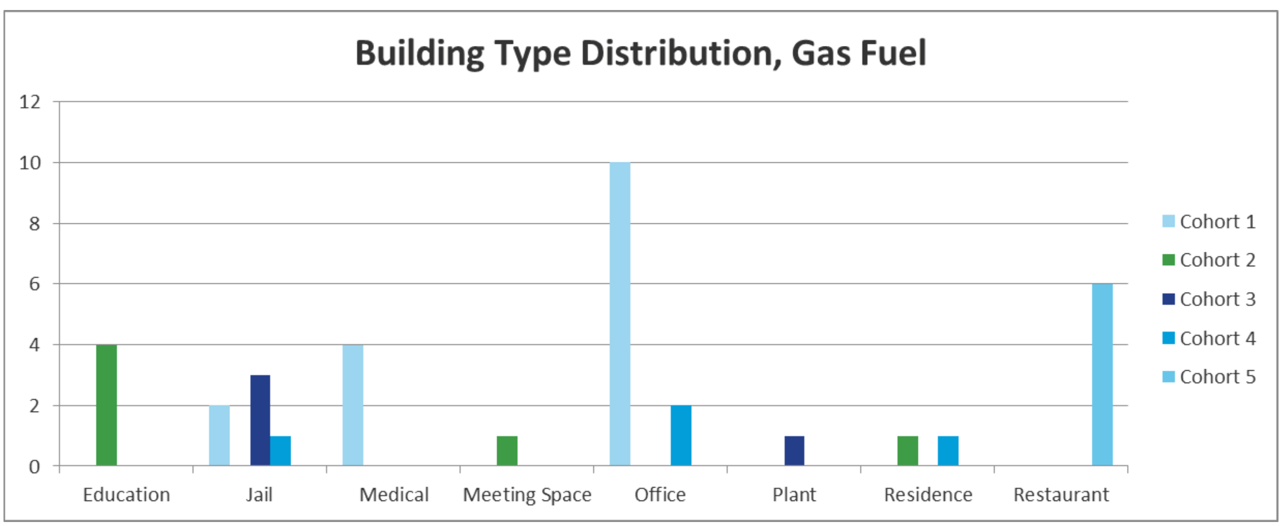
APPENDIX A: SUPPLEMENTAL CHARTS

The following charts were analyzed as part of the evaluation's diagnosis of the energy models sampled.

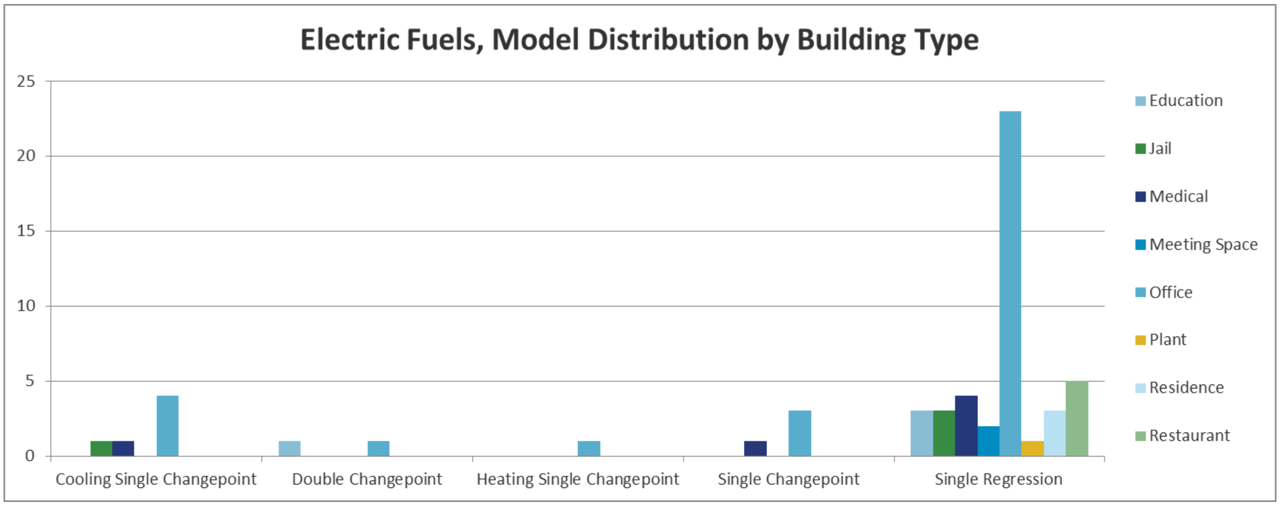
Distribution by Cohort & Building Type

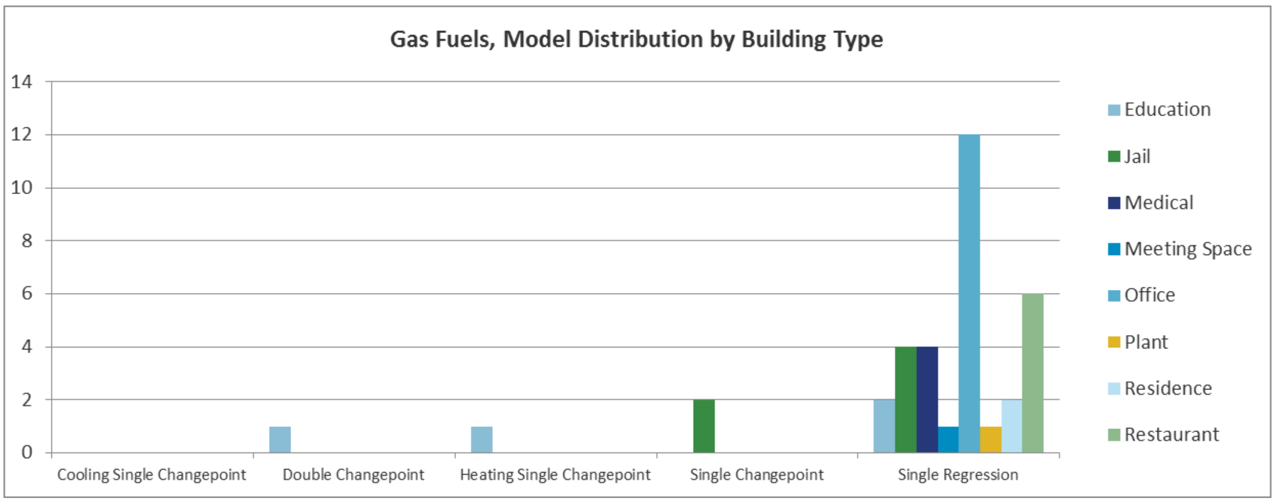




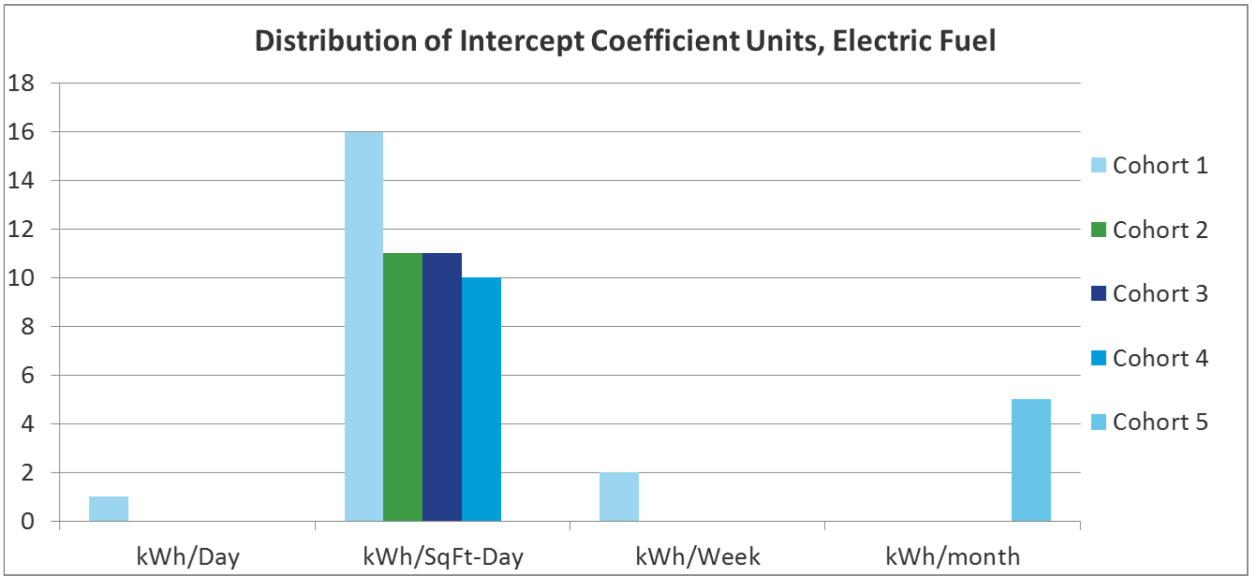


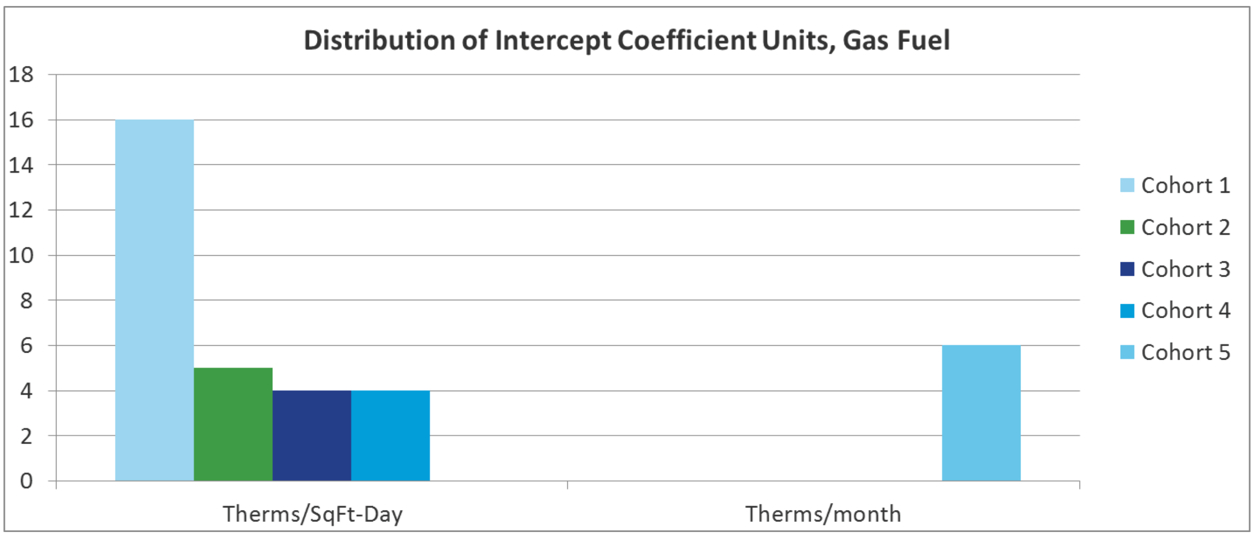
Model Type Distribution, MT&R



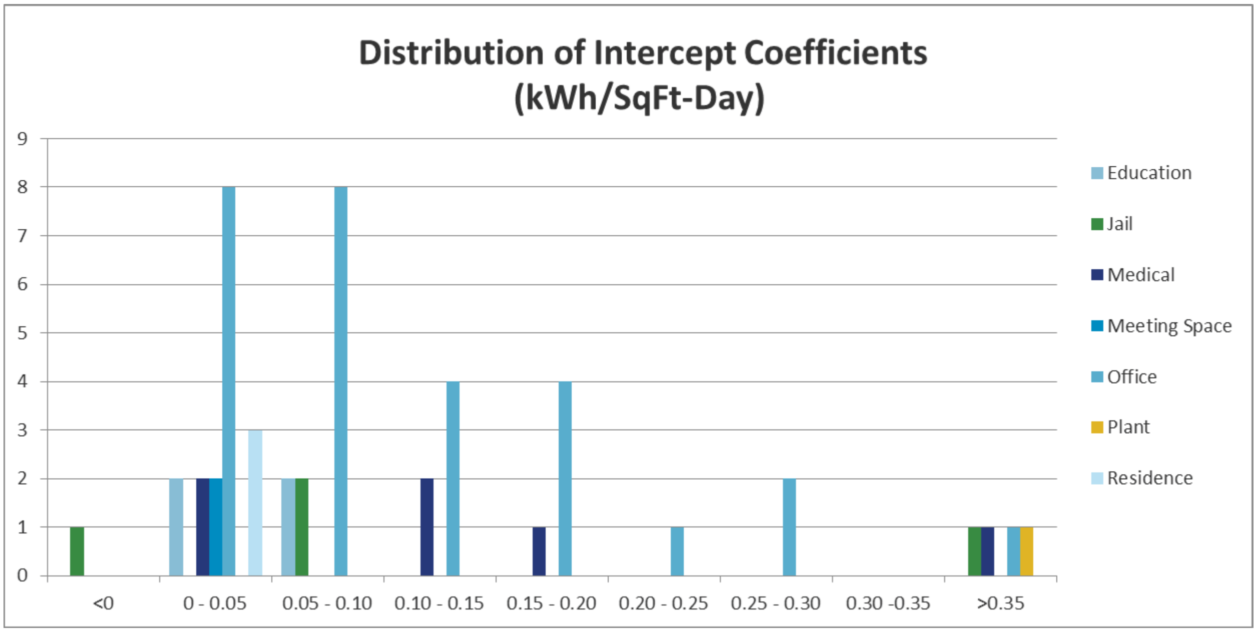


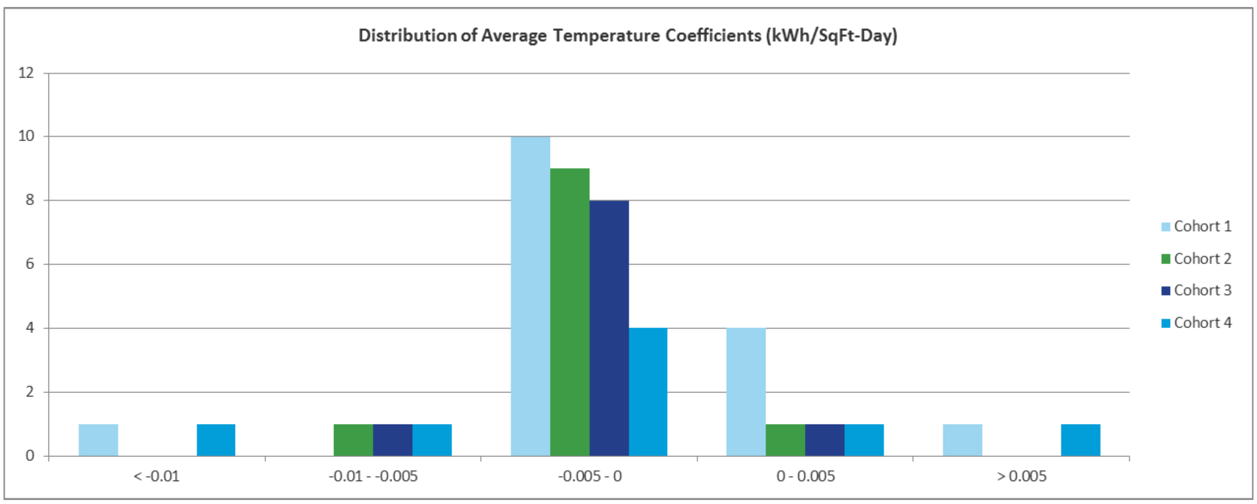
Intercept Coefficients Units

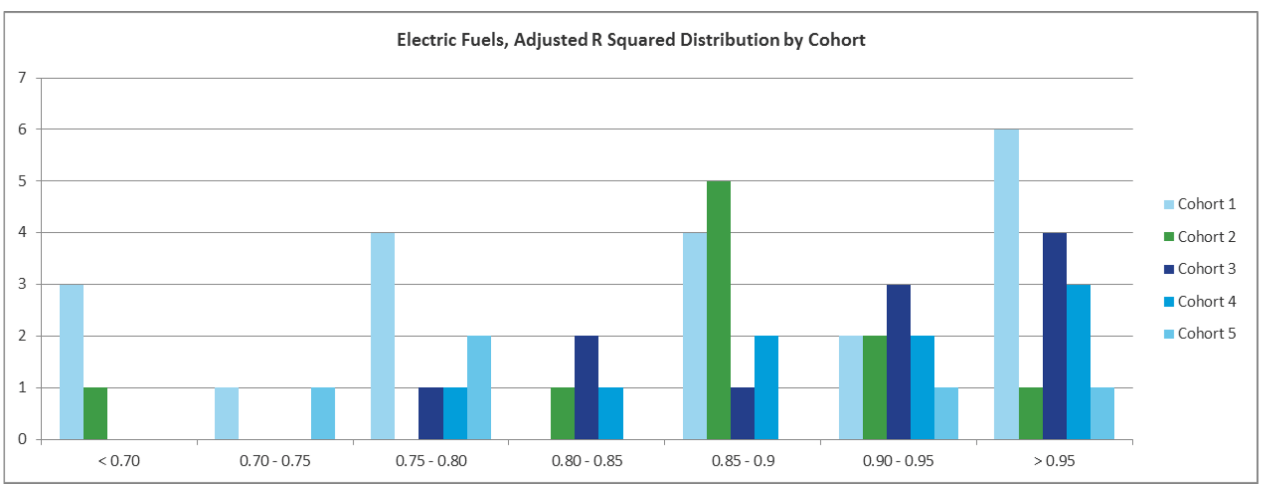




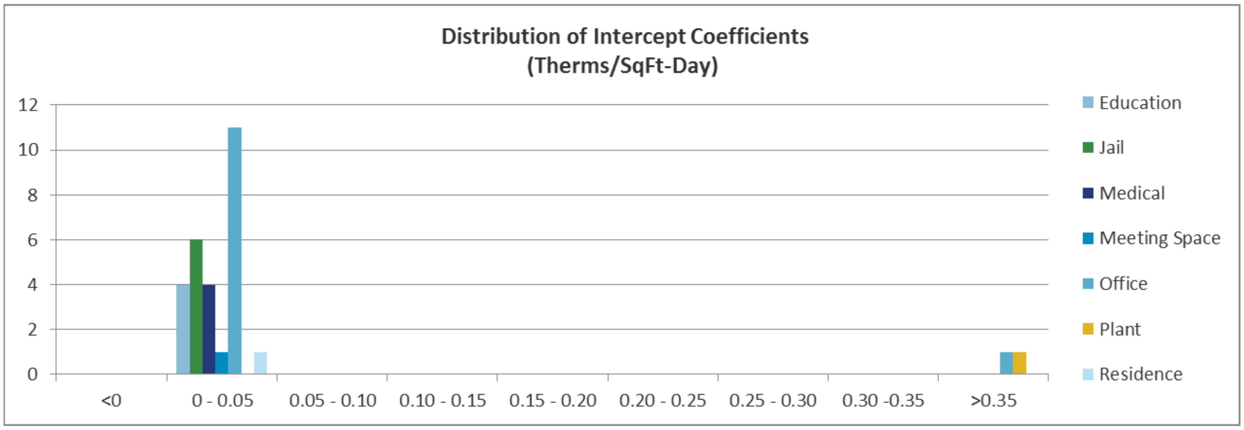
Electric Model Coefficients, MT&R

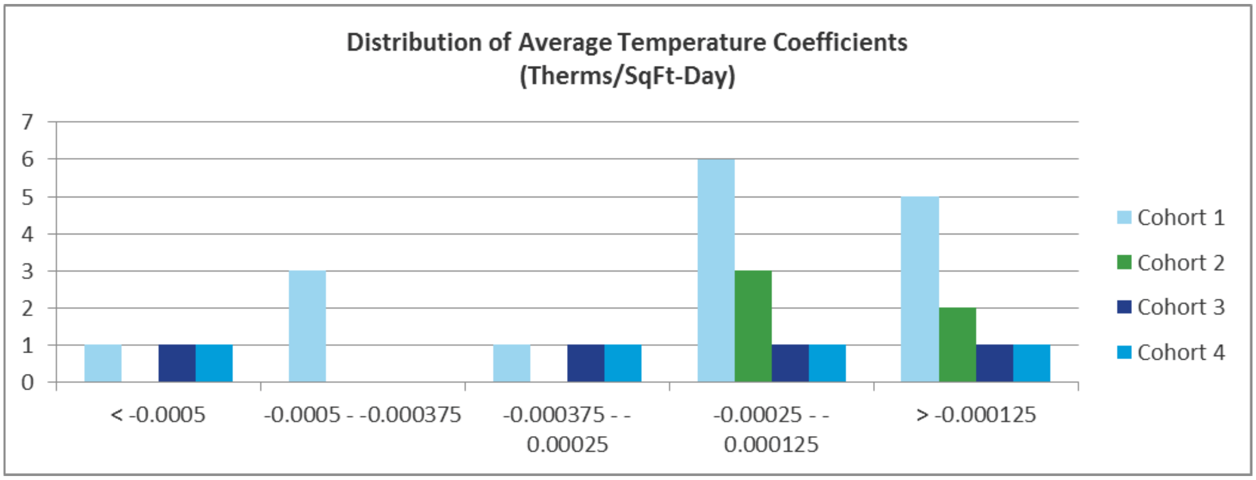


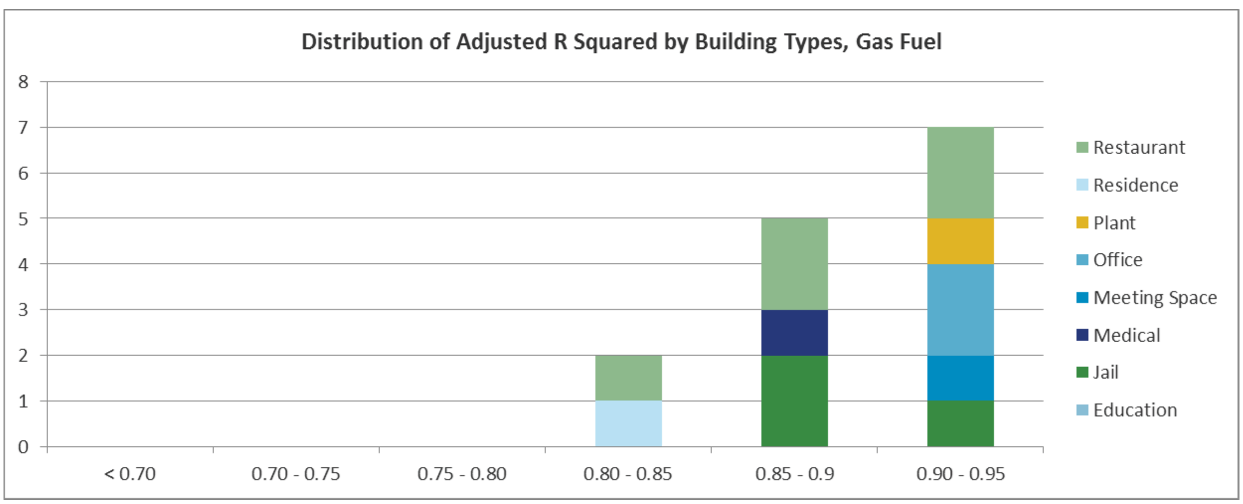




Gas Model Coefficients, MT&R







About DNV GL