MEMO

Date: February 7, 2020
To: Board of Directors
From: Dan Rubado, Evaluation Project Manager
Subject: Staff Response to Recurve Residential Ceiling Insulation Impact Analysis

Energy Trust used an impact analysis tool built by Recurve Analytics to evaluate the gas and electricity savings of ceiling insulation retrofit projects in single-family homes in Oregon completed from 2013 to 2017. The tool uses monthly utility billing data to conduct pre/post analyses of whole home energy usage. Energy usage data are weather-normalized using typical meteorological year data. Normalized annual energy usage in the year immediately preceding the project is compared with that of the year immediately following the project. The change in normalized annual energy usage is then evaluated against changes in energy usage during the same time period in two comparison groups—a site-level matched non-participant comparison group and a group of homes that completed ceiling insulation projects in later years (future participants). These calculations provide two estimates of the average annual energy savings resulting from the measure, given typical weather conditions.

We restricted the analysis to ceiling insulation projects where no other efficiency measures were installed in the home during the analysis period. This was done to isolate the energy impact of ceiling insulation in the analysis, although these projects may not be representative of typical ceiling insulation projects, in which many different measures are installed together—principally wall and floor insulation. Several standard data screens were applied to remove atypical projects and zero in on the impact of ceiling insulation. As shown in the Recurve snapshot reports that follow, energy savings were large and significant across the board, although they were more robust for gas-heated homes than electrically-heated homes. These results indicate that ceiling insulation continues to be one of the highest impact residential efficiency measures that Energy Trust supports.

Heating zones are geographic areas defined by the Regional Technical Forum, based on the number of heating degree-days during a typical winter. Heating zone 1 represents areas of the state with relatively mild winters, such as Western Oregon. Heating zones 2 and 3 represent areas of the state with cold winters, like the mountains and Central and Eastern Oregon. We focused our analysis on projects located in heating zone 1 because Energy Trust's ceiling insulation measures are stratified by heating zone and there were a sufficient number of zone 1 projects available for analysis. We were unable to quantify ceiling insulation savings in heating zone 2, due to a small number of projects.

For heating zone 1 from 2013 to 2017, overall average gas savings in gas-heated homes ranged from 104 to 106 therms per year (+/- 10 therms) or 16% of baseline gas usage. There were 477 treatment homes analyzed, which had an average annual baseline gas usage of 655 therms. They were concentrated in the Portland metro area, although they were distributed all along the I-5 corridor in Western Oregon. Both the matched and future comparison groups provided a good representation of the baseline gas usage in the treatment group and a reasonable point of comparison as similar homes that did not install ceiling insulation. The large sample size, good precision and close match between groups give us high confidence in this result.

We analyzed gas savings in the three most recent years of ceiling insulation projects in gas-heated homes separately (2015, 2016, and 2017) to see if there were any changes in savings that might be occurring over time. To preserve large enough sample sizes to analyze individual years, we combined projects in heating zones 1 and 2, although most projects were in heating zone 1. We did not see a coherent time trend and the year-to-year differences could easily be explained by variability in gas usage, lower sample sizes and lower precision. Each of the annual savings estimates was within 20% of the overall gas savings estimate. Results for 2015 were relatively robust and closely aligned with the overall gas savings estimate. However, 2016 and 2017 projects had lower sample sizes and precision and the savings estimates were substantially higher and lower than the overall estimate, respectively.

We also analyzed the electric savings for gas-heated homes, which result from reduced furnace fan runtime and summer cooling savings. In heating zone 1 from 2013 to 2017, average electric savings ranged from 155 to 180 kWh per year (+/- 170 kWh) or 2% of baseline electricity usage. There were only 238 homes available for this analysis with average annual baseline electricity usage of 8,680 kWh. The magnitude of these savings is relatively small compared to the variability in electricity usage, so the savings value is uncertain. However, the electric savings are borderline statistically significant and show that ceiling insulation in gas-heated homes most likely does have an impact on electricity usage. The comparison groups provided only rough matches to the treatment homes and may provide a somewhat skewed point of comparison. Thus, we have some confidence in the direction of the result but note that the point estimate has low precision.

Overall average electric savings in electrically-heated homes in heating zone 1 ranged from 1,560 to 1,910 kWh per year (+/- 580), or 9-11% of baseline electricity usage—not as high or precise as the gas savings estimate but still relatively large and statistically significant. There was higher variability in electricity usage and smaller sample sizes available for electrically-heated homes. In addition, neither comparison group provided a good match to the treatment homes, so these points of comparison may be somewhat skewed. However, the precision of the savings estimate was moderately good compared with the magnitude of savings, so we have moderate confidence in the result.

In the table below, we summarize results of the various ceiling insulation analysis scenarios we looked at. Results are provided for kWh and therm savings for gas- and electrically-heated homes for projects completed from 2013 to 2017. For individual year savings estimates, we combined the two heating zones. We present the midpoint savings estimate of the two comparison group methodologies (matched non-participants and future participants).

Table 1: Ceiling insulation energy savings analysis summary of results

Fuel Analyzed	**Heating Fuel**	**Heating Zone**	**Years**	**N***	**Baseline Energy Usage**	**Average Savings†**	**Absolute Precision†**	**Percent Savings†**	**Conf. Level**
Therms	Gas	1	2013-2017	477	655	105	+/- 10	16%	High
Therms	Gas	All	2015	113	659	109	+/- 20	17%	High
Therms	Gas	All	2016	77	698	126	+/- 28	18%	Moderate
Therms	Gas	All	2017	89	682	92	+/- 30	14%	Moderate

kWh	Gas	1	2013-2017	238	8,680	170	+/- 170	2%	Low
kWh	Electricity	1	2013-2017	107	16,990	1,730	+/- 580	10%	Moderate

* N is the final treatment group sample size in the analysis.

† The average savings, absolute precision and percent savings values represent the midpoint estimates between the two comparison group methodologies used.

These results reaffirm ceiling insulation is an important measure that continues to deliver large gas and electricity savings in Oregon. Energy Trust will use results from this and other Recurve analyses to update savings assumptions used in our standard residential measures when updates are made.

Impact Evaluation Report

Gas Impact of Ceiling Insulation in Program Year 2013, 2014, 2015, 2016, 2017

Result Summary

Measure: Ceiling Insulation	Program Year: 2013, 2014, 2015, 2016, 2017	Fuel: Gas	

Meter Data Filters:	DNAC: <75%	DNAC Percentile: None	Annual Consumption Percentile: Remove Top and Bottom 0.5%	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Model Filters:	Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1	
Metadata Filters:	Cooling Zone(s): All		Heating Fuel: Gas	

Thermostat Name: All	Heat Pump Baseline: All	Heating Zone(s): 1 - Hdd <= 6000	Multi Measure Filter: Single Measure Only
Heat Pump Manufacturer: All	Heat Pump Adv. Controls or Commissioning: All		

477 Treatment Meters	107 +/- 9 Therms Average Normal Year Pre-Post Difference in Consumption per Participant	16 +/- 1 % Percent Normal Year Pre-Post Difference in Consumption per Participant	655 Mean Baseline Consumption (Gas)	98% Realization Rate
2,375 Site-level Matched Meters	106 +/- 10 Therms Average Savings Relative to Site-level Matched Comparison Group	16 +/- 1% Percent Savings Relative to Site-level Matched Comparison Group	645 Mean Baseline Consumption (Gas)	97% Realization Rate
4,838 Future Participant Meters	104 +/- 10 Therms Average Savings Relative to Future Participant Group	16 +/- 1% Savings Relative to Future Participant Group	660 Mean Baseline Consumption (Gas)	95% Realization Rate

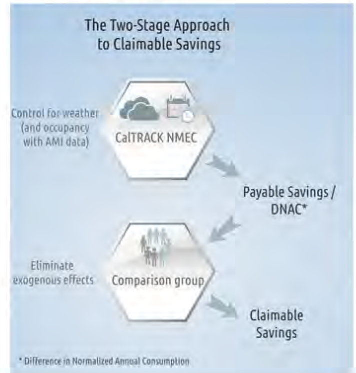
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

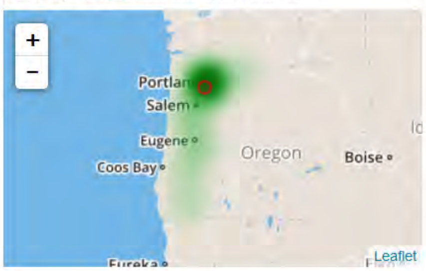
Two-Stage Approach



Treatment Group
The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality for modeling.

Treatment Site Locations

Projects Treatment Group Centroid



34.9 miles

80% of projects lie within this distance from treatment group centroid

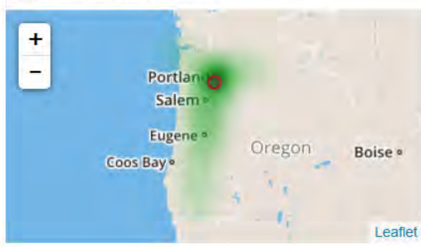
477
Meters

655
Mean Baseline Consumption (Gas)

Site-level Matched Comparison Group
This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the group. Matching was performed using monthly consumption in the baseline period as detailed in the Methodology section.

Site-level Matched Site Locations

Comparison Group Centroid Projects
Treatment Group Centroid



0.3 miles

Distance between treatment and comparison group centroids

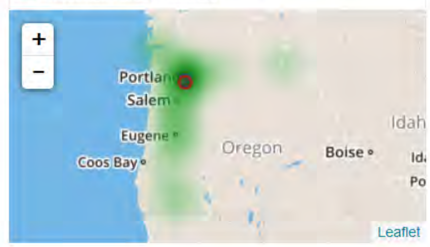
2,375
Meters

645
Mean Baseline Consumption (Gas)

Future Participant Group
The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations

Comparison Group Centroid Projects
Treatment Group Centroid



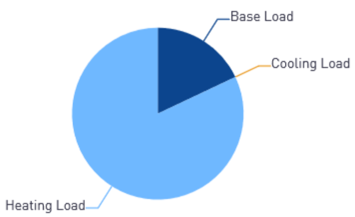
1.1 miles

Distance between treatment and future participant group centroids

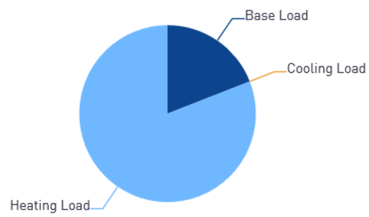
4,838
Meters

660
Mean Baseline Consumption (Gas)

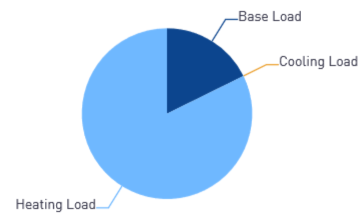
Load Disaggregation



Load Disaggregation



Load Disaggregation



2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

5,374	477	8.9%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

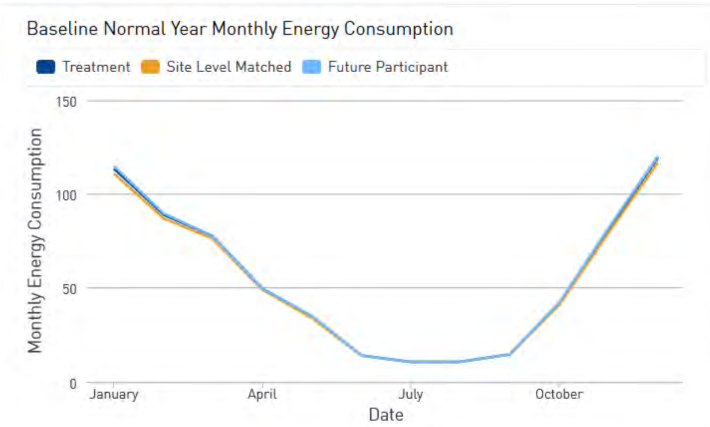
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2014, 2013, 2015, 2016, 2017 -- Fuel: Gas	--	5,374
Meters with valid consumption data in baseline and/or reporting periods.	--	354	5,020
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	4,345	675
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Gas	62	613
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 1 - Hdd <= 6000 -- Cooling Zone: All	20	593
Other measure-specific filters.	--	0	593
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	79	514
Meters with at least 5 site-level matched meters from the comparison group pool.	--	19	495
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	10	485

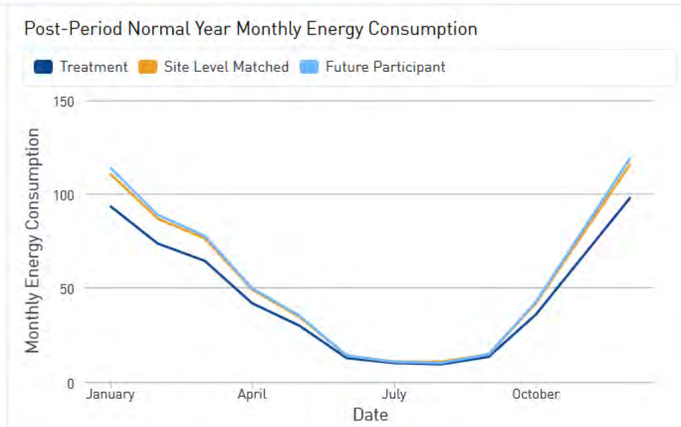
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: None	0	485
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	2	483
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	6	477
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	477

3. Modeling Results

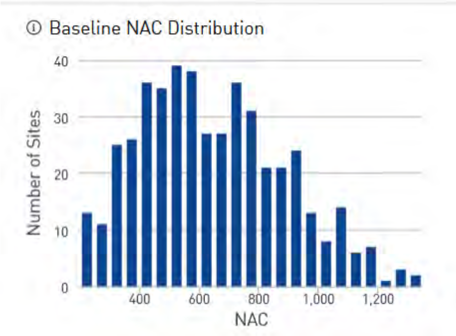
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

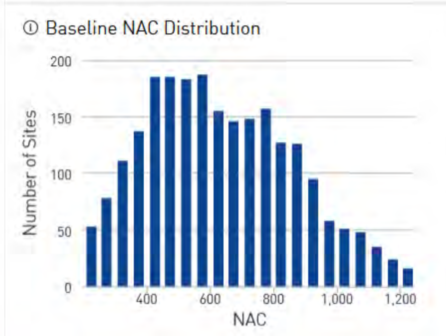




Treatment Group



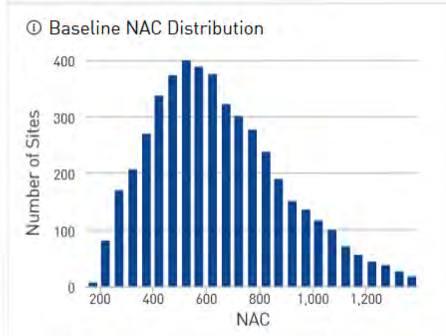
Site-level Matched Comparison Group



0.103

Annual Consumption p-value

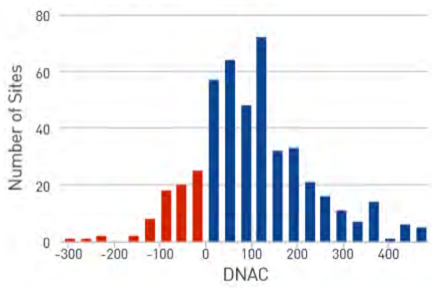
Future Participant Group



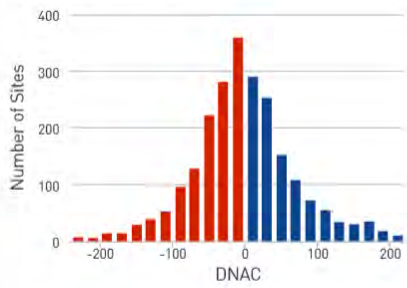
0.192

Annual Consumption p-value

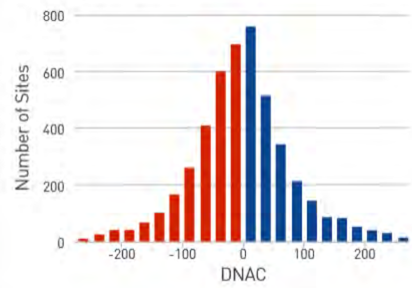
DNAC Distribution



DNAC Distribution

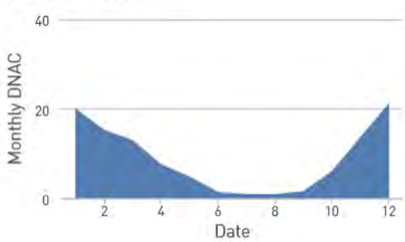


DNAC Distribution

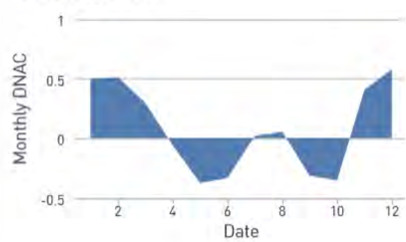


107 +/- 9 Therms	16 +/- 1 %	1 +/- 3 Therms	0 +/- 1 %	3 +/- 3 Therms	0 +/- 0 %
Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline

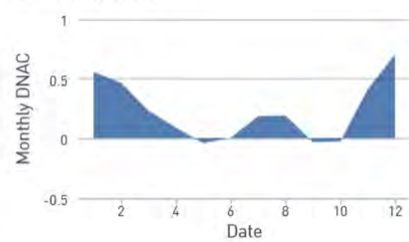
Monthly DNAC



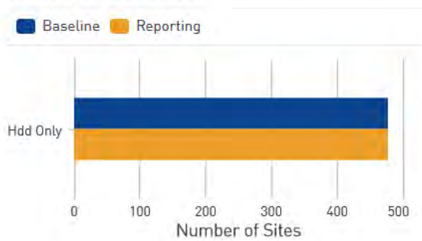
Monthly DNAC



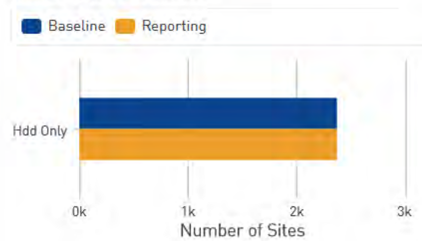
Monthly DNAC



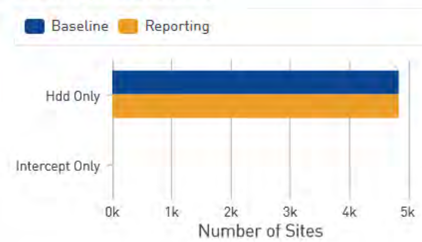
Model Type Distribution



Model Type Distribution



Model Type Distribution

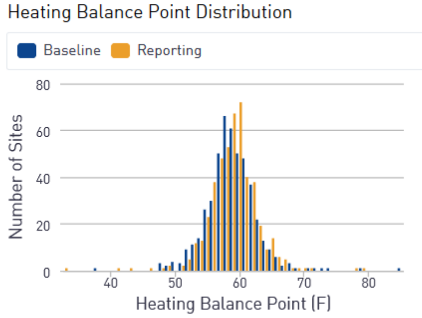


Treatment Group

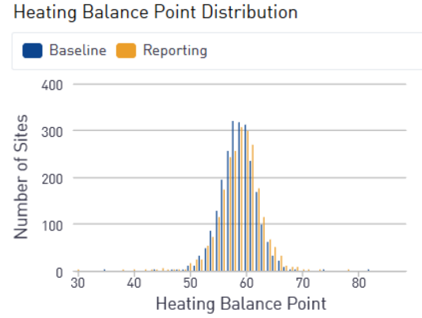
Site-level Matched Comparison Group

Future Participant Group

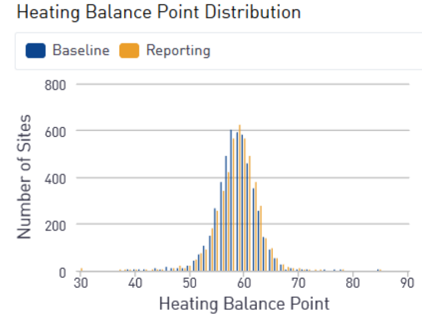
Heating Balance Point Distribution

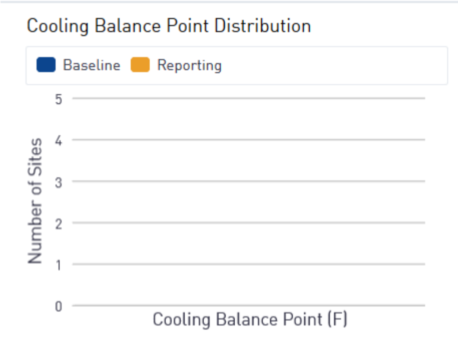


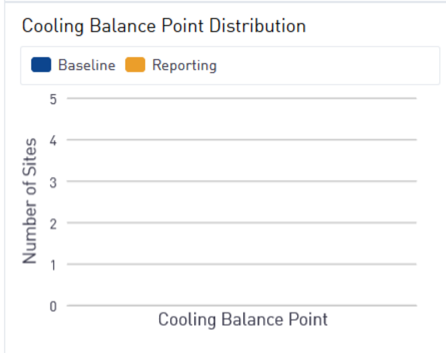
Heating Balance Point Distribution

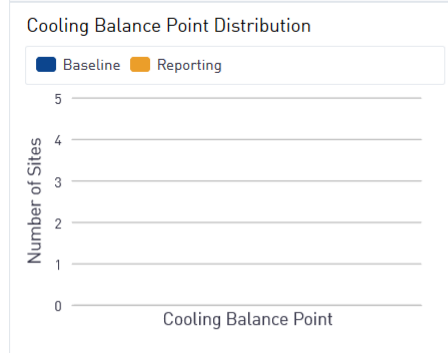


Heating Balance Point Distribution









4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

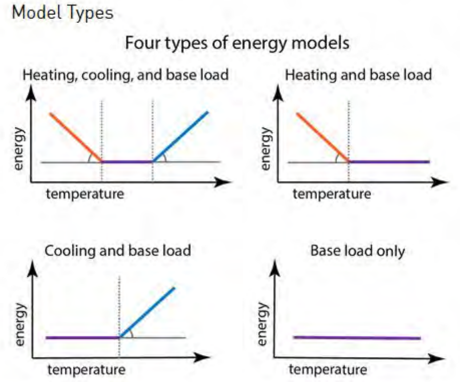
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

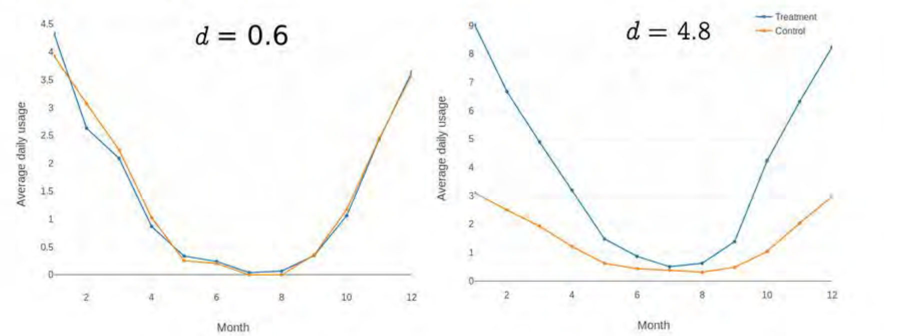
Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.

Model Types



Euclidean Distance Matching



RECURVE

Impact Evaluation Report

Gas Impact of Ceiling Insulation in Program Year 2015

Result Summary

Measure: Ceiling Insulation	Program Year: 2015	Fuel: Gas	

Meter Data Filters:	DNAC: <75%	DNAC Percentile: None	Annual Consumption Percentile: Remove Top and Bottom 0.5%
Model Filters:	Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1
Metadata Filters:	Cooling Zone(s): All		Heating Fuel: Gas

Thermostat Name: All	Heat Pump Baseline: All	Heating Zone(s): 1 - Hdd <= 6000, 2 - 6000 < Hdd < 7500, 3 - Hdd >= 7500	Multi Measure Filter: Single Measure Only
Heat Pump Manufacturer: All	Heat Pump Adv. Controls or Commissioning: All		

Last Consumption Data Update: October 1, 2019

Last Participation Data Update: October 1, 2019

CalTRACK Version: 2.0

113 Treatment Meters	99 +/- 19 Therms Average Normal Year Pre-Post Difference in Consumption per Participant	15 +/- 3 % Percent Normal Year Pre-Post Difference in Consumption per Participant	659 Mean Baseline Consumption (Gas)	102% Realization Rate
560 Site-level Matched Meters	123 +/- 20 Therms Average Savings Relative to Site-level Matched Comparison Group	19 +/- 3% Percent Savings Relative to Site-level Matched Comparison Group	636 Mean Baseline Consumption (Gas)	127% Realization Rate
974 Future Participant Meters	95 +/- 20 Therms Average Savings Relative to Future Participant Group	14 +/- 3% Savings Relative to Future Participant Group	672 Mean Baseline Consumption (Gas)	98% Realization Rate

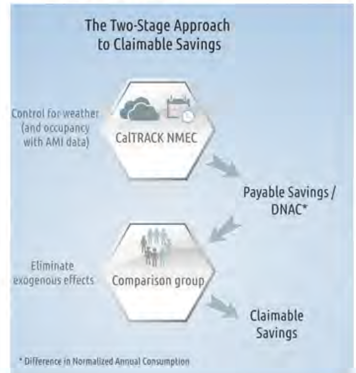
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

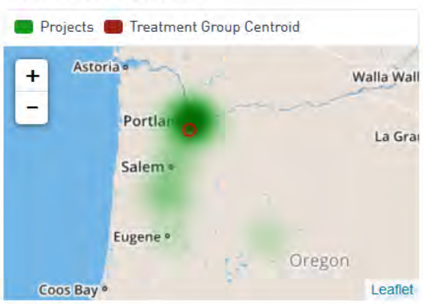
The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

Two-Stage Approach



Treatment Group
The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality for modeling.

Treatment Site Locations



38.6 miles

80% of projects lie within this distance from treatment group centroid

113

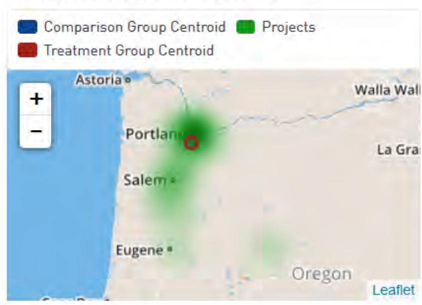
Meters

659

Mean Baseline Consumption (Gas)

Site-level Matched Comparison Group
This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the group. Matching was performed using monthly consumption in the baseline period as detailed in the Methodology section.

Site-level Matched Site Locations



0.2 miles

Distance between treatment and comparison group centroids

560

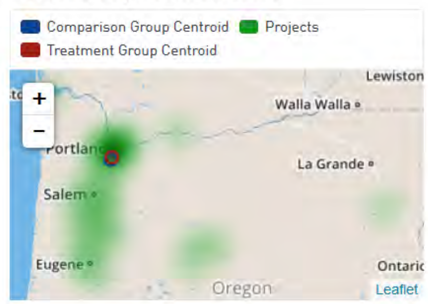
Meters

636

Mean Baseline Consumption (Gas)

Future Participant Group
The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations



2.4 miles

Distance between treatment and future participant group centroids

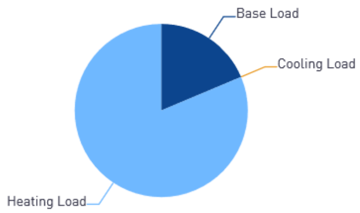
974

Meters

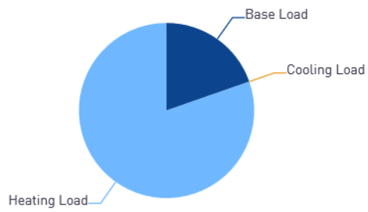
672

Mean Baseline Consumption (Gas)

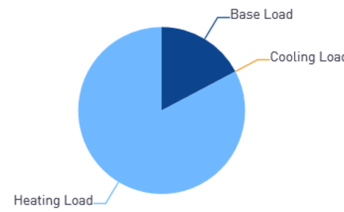
Load Disaggregation



Load Disaggregation



Load Disaggregation



2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

1,069	113	11%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2015 -- Fuel: Gas	--	1,069
Meters with valid consumption data in baseline and/or reporting periods.	--	60	1,009
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	868	141
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Gas	7	134
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 2 - 6000 < Hdd < 7500, 1 - Hdd <= 6000, 3 - Hdd >= 7500 -- Cooling Zone: All	2	132
Other measure-specific filters.	--	0	132
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	14	118
Meters with at least 5 site-level matched meters from the comparison group pool.	--	3	115
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	0	115

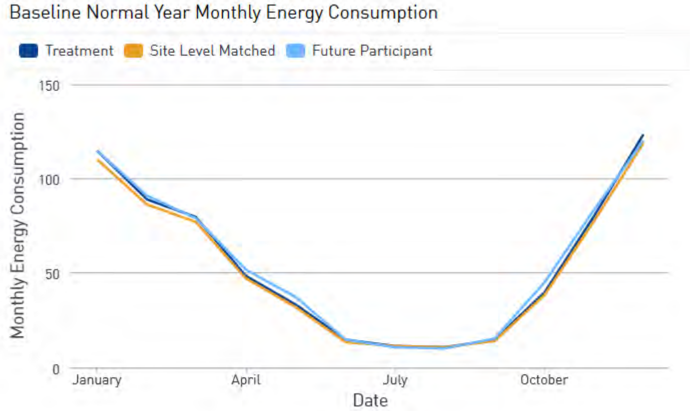
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	115
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	115
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	2	113
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	113

3. Modeling Results

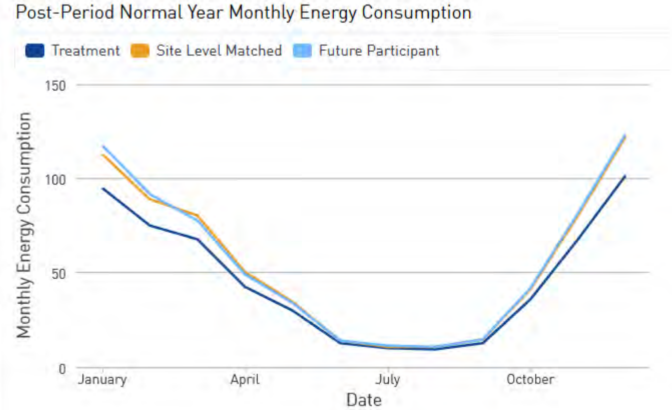
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

Baseline Normal Year Monthly Energy Consumption

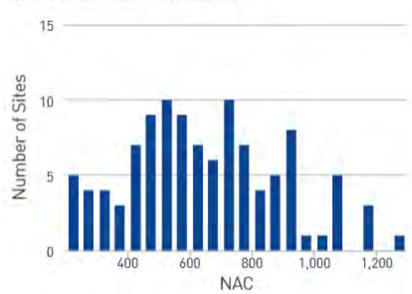


Post-Period Normal Year Monthly Energy Consumption



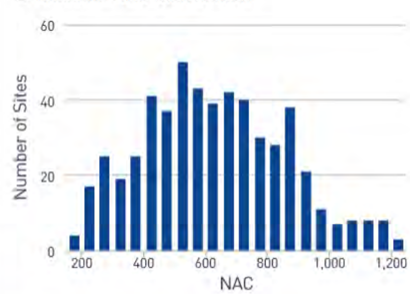
Treatment Group

ⓘ Baseline NAC Distribution



Site-level Matched Comparison Group

ⓘ Baseline NAC Distribution



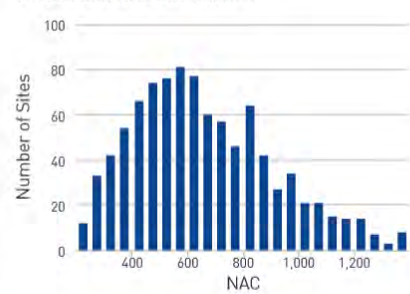
ⓘ

0.168

Annual Consumption p-value

Future Participant Group

ⓘ Baseline NAC Distribution



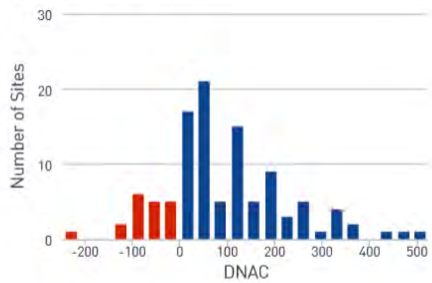
ⓘ

0.184

Annual Consumption p-value

Treatment Group

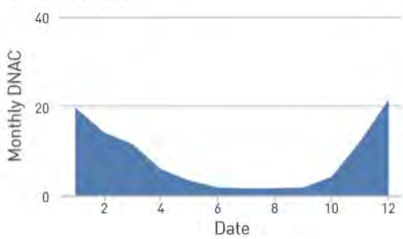
ⓘ DNAC Distribution



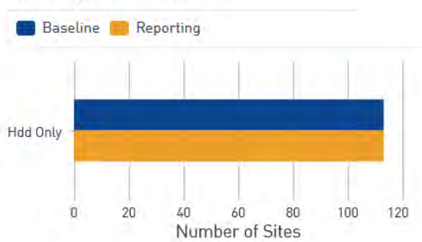
99 +/- 19 Therms
Average Difference in Normalized Annual Consumption per Participant

15 +/- 3 %
Difference in Normalized Annual Consumption as a Percent of Baseline

ⓘ Monthly DNAC

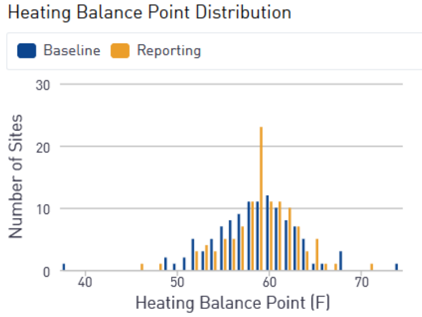


Model Type Distribution



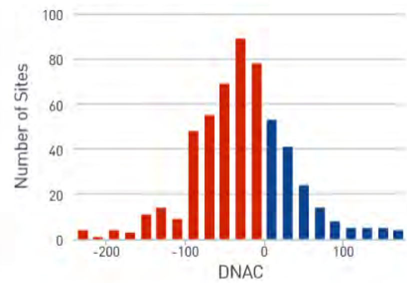
Treatment Group

Heating Balance Point Distribution



Site-level Matched Comparison Group

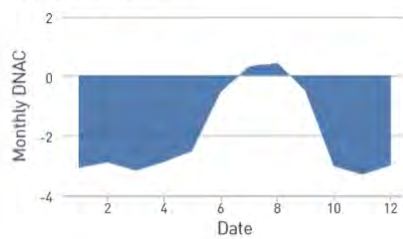
ⓘ DNAC Distribution



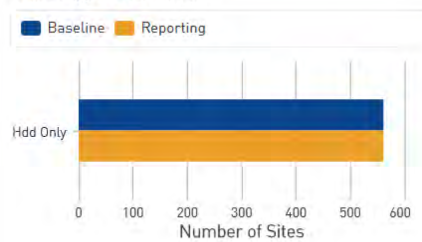
-24 +/- 7 Therms
Average Difference in Normalized Annual Consumption per Participant

-4 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline

ⓘ Monthly DNAC

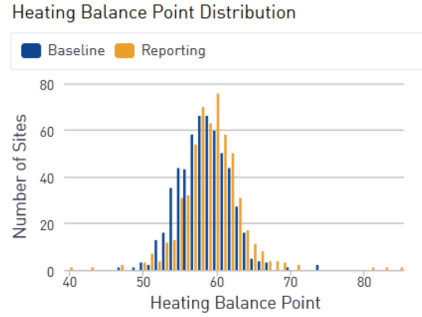


Model Type Distribution



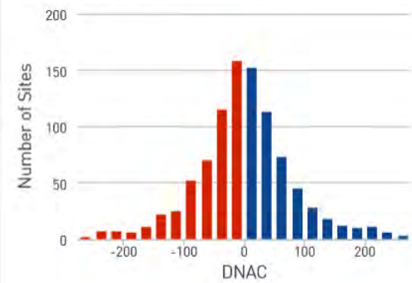
Site-level Matched Comparison Group

Heating Balance Point Distribution



Future Participant Group

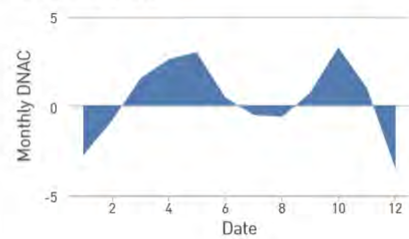
ⓘ DNAC Distribution



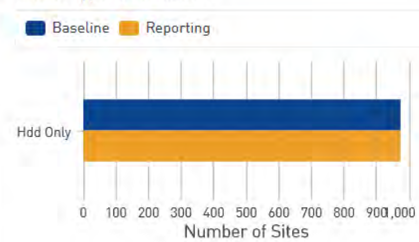
4 +/- 7 Therms
Average Difference in Normalized Annual Consumption per Participant

1 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline

ⓘ Monthly DNAC

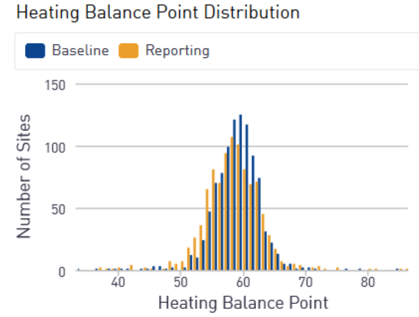


Model Type Distribution

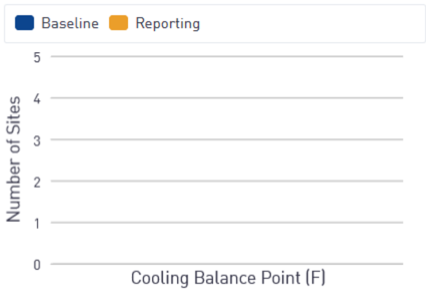


Future Participant Group

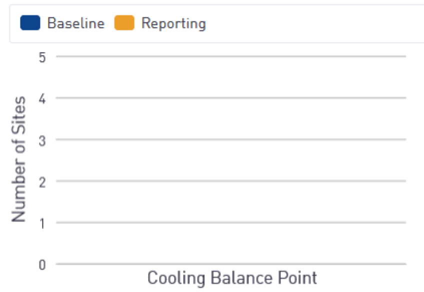
Heating Balance Point Distribution



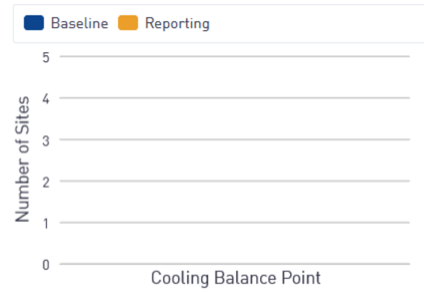
Cooling Balance Point Distribution



Cooling Balance Point Distribution



Cooling Balance Point Distribution



4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

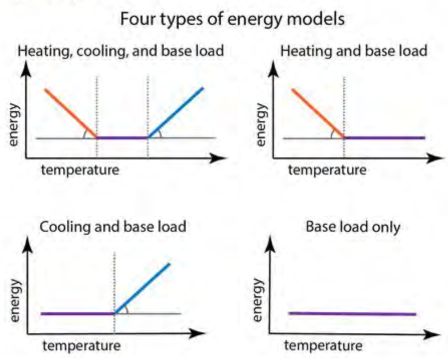
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

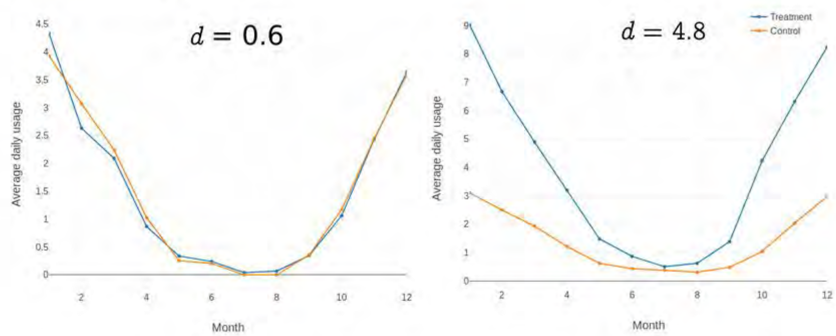
Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.

Model Types



Euclidean Distance Matching



Impact Evaluation Report

Gas Impact of Ceiling Insulation in Program Year 2016

Result Summary

Measure: Ceiling Insulation Measure: Ceiling Insulation		Program Year: 2016		Fuel: Gas	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:		DNAC: <75%	DNAC Percentile: None	Annual Consumption Percentile: Remove Top and Bottom 0.5%	
Model Filters:		Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1	
Metadata Filters:		Cooling Zone(s): All		Heating Fuel: Gas	
Thermostat Name: All	Heat Pump Baseline: All	Heating Zone(s): 1 - Hdd <= 6000, 2 - 6000 < Hdd < 7500, 3 - Hdd >= 7500		Multi Measure Filter: Single Measure Only	
Heat Pump Manufacturer: All	Heat Pump Adv. Controls or Commissioning: All				

77 Treatment Meters	122 +/- 26 Therms Average Normal Year Pre-Post Difference in Consumption per Participant	17 +/- 4 % Percent Normal Year Pre-Post Difference in Consumption per Participant	698 Mean Baseline Consumption (Gas)	135% Realization Rate
385 Site-level Matched Meters	111 +/- 28 Therms Average Savings Relative to Site-level Matched Comparison Group	16 +/- 4% Percent Savings Relative to Site-level Matched Comparison Group	691 Mean Baseline Consumption (Gas)	122% Realization Rate
902 Future Participant Meters	140 +/- 27 Therms Average Savings Relative to Future Participant Group	20 +/- 4% Savings Relative to Future Participant Group	661 Mean Baseline Consumption (Gas)	155% Realization Rate

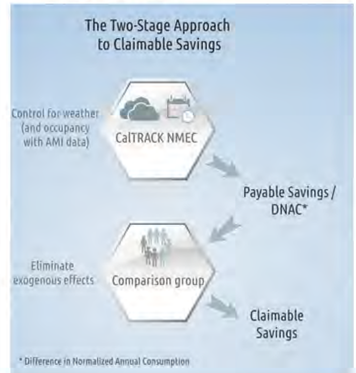
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

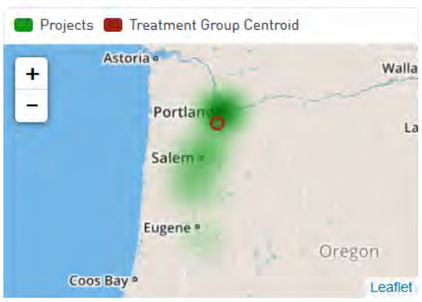
The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

Two-Stage Approach



Treatment Group
The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality for modeling.

Treatment Site Locations



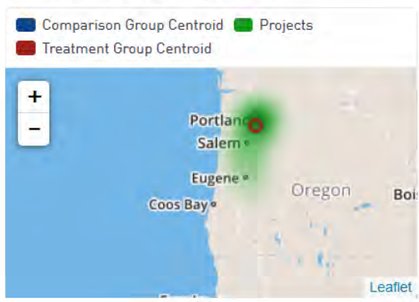
33.7 miles

80% of projects lie within this distance from treatment group centroid

77	698
Meters	Mean Baseline Consumption (Gas)

Site-level Matched Comparison Group
This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the group. Matching was performed using monthly consumption in the baseline period as detailed in the Methodology section.

Site-level Matched Site Locations



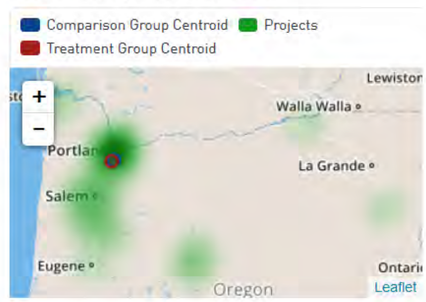
0.8 miles

Distance between treatment and comparison group centroids

385	691
Meters	Mean Baseline Consumption (Gas)

Future Participant Group
The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations

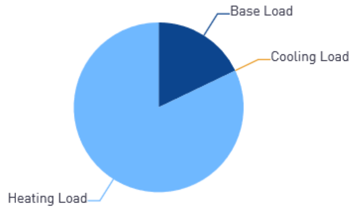


2.4 miles

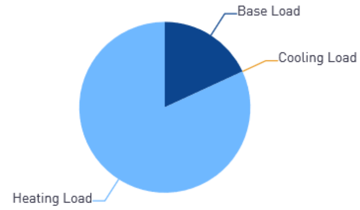
Distance between treatment and future participant group centroids

902	661
Meters	Mean Baseline Consumption (Gas)

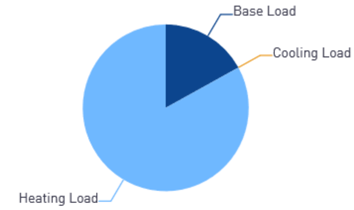
Load Disaggregation



Load Disaggregation



Load Disaggregation



2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

720	77	11%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2016 -- Fuel: Gas	--	720
Meters with valid consumption data in baseline and/or reporting periods.	--	31	689
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	573	116
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Gas	14	102
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 2 - 6000 < Hdd < 7500, 1 - Hdd <= 6000, 3 - Hdd >= 7500 -- Cooling Zone: All	0	102
Other measure-specific filters.	--	0	102
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	17	85
Meters with at least 5 site-level matched meters from the comparison group pool.	--	7	78
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	5	78

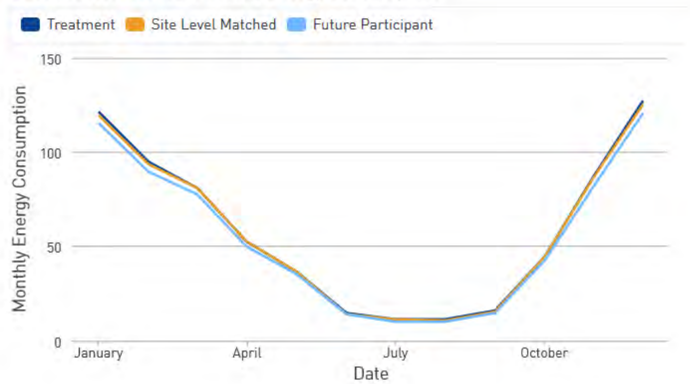
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	78
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	78
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	1	77
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	77

3. Modeling Results

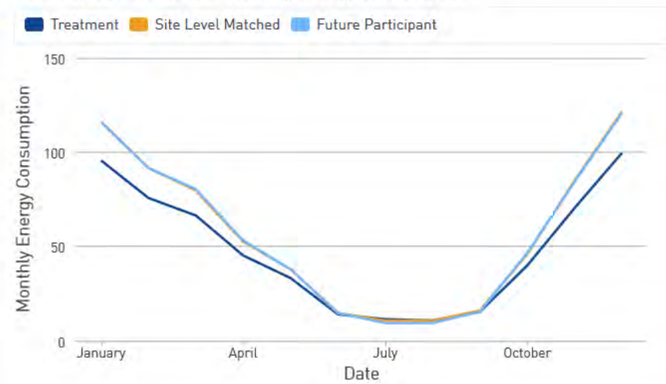
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

Baseline Normal Year Monthly Energy Consumption

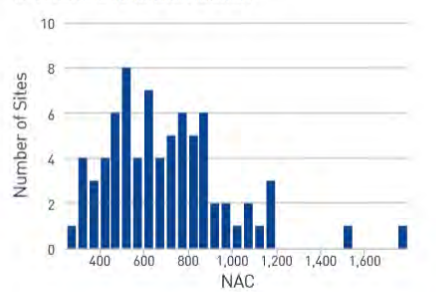


Post-Period Normal Year Monthly Energy Consumption



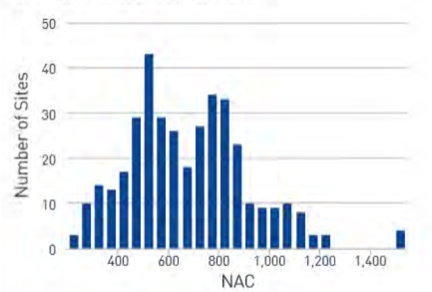
Treatment Group

Baseline NAC Distribution



Site-level Matched Comparison Group

Baseline NAC Distribution

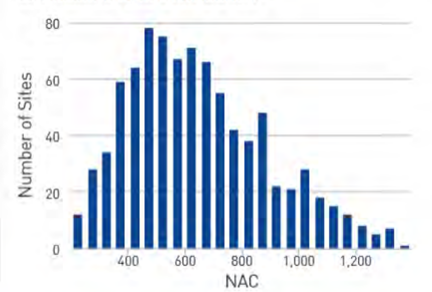


0.177

Annual Consumption p-value

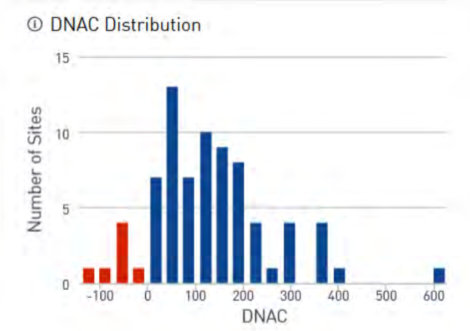
Future Participant Group

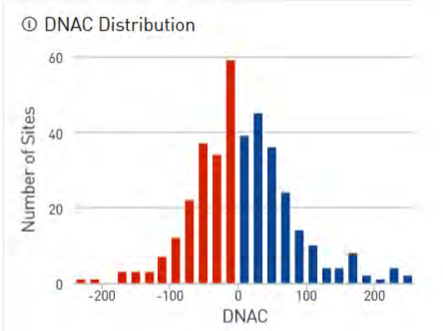
Baseline NAC Distribution

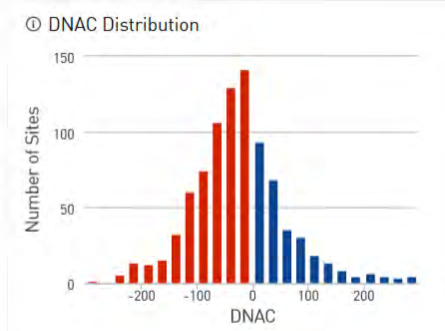


0.238

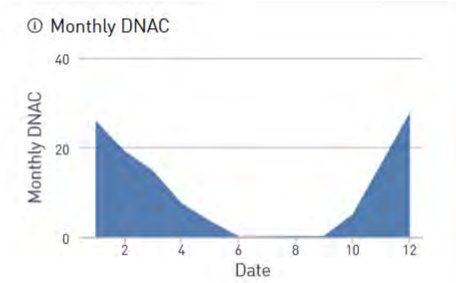
Annual Consumption p-value

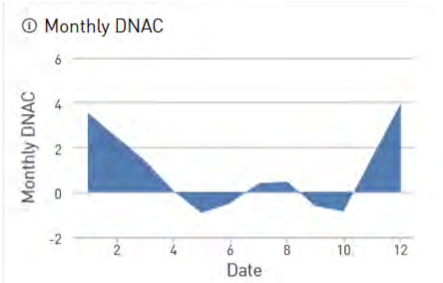


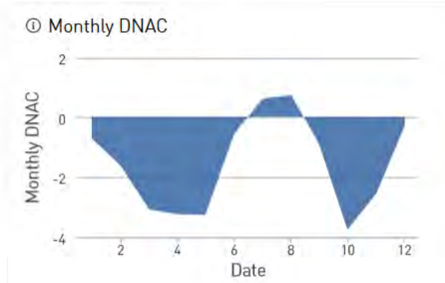


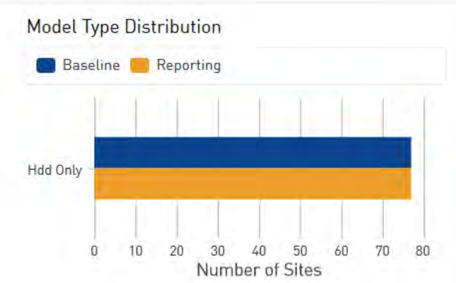


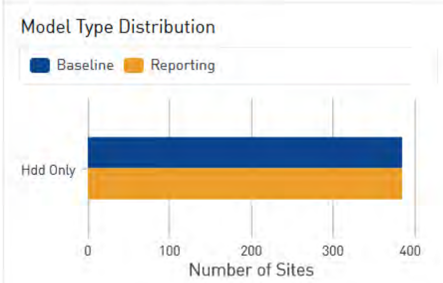
122 +/- 26 Therms	17 +/- 4 %	11 +/- 8 Therms	2 +/- 1 %	-18 +/- 7 Therms	-3 +/- 1 %
Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline

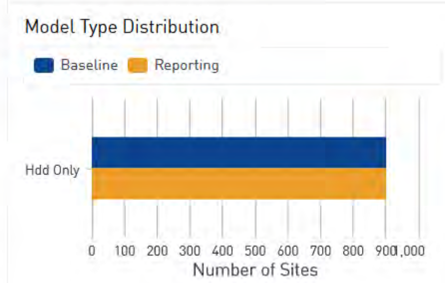




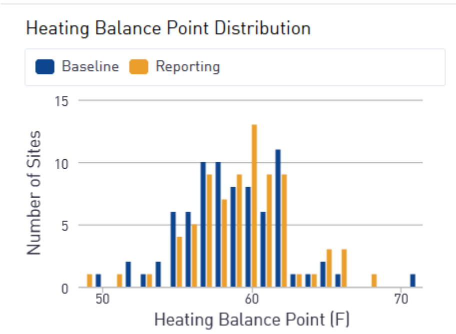


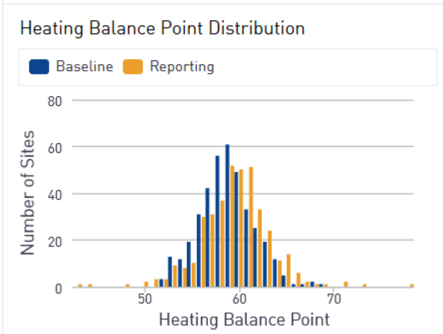


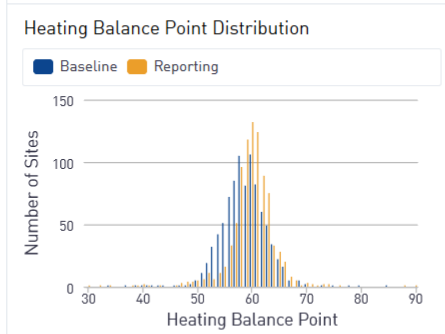


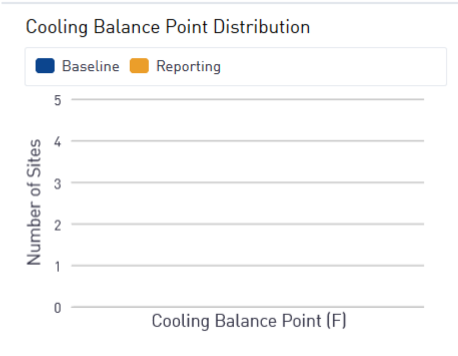


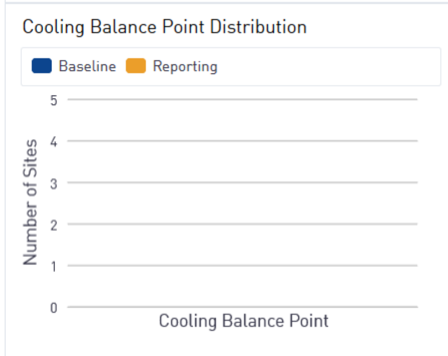
Treatment Group	Site-level Matched Comparison Group	Future Participant Group

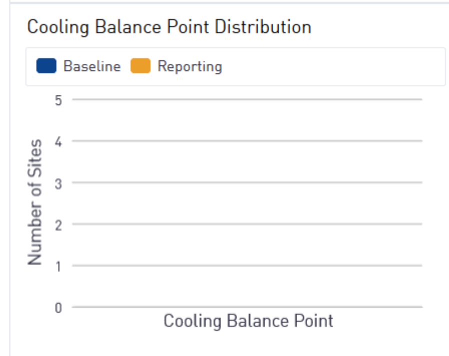












4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

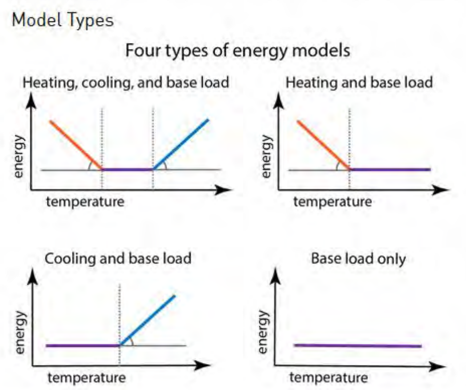
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

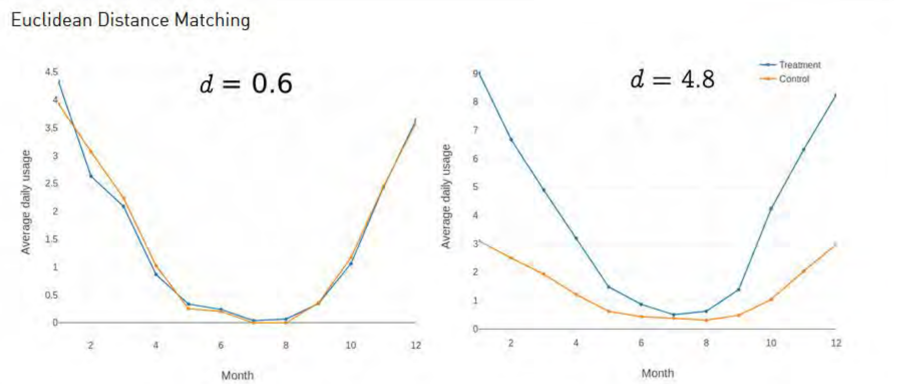
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.





Impact Evaluation Report

Gas Impact of Ceiling Insulation in Program Year 2017

Result Summary

Measure: Ceiling Insulation	Program Year: 2017		Fuel: Gas	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <75%	DNAC Percentile: All	Annual Consumption Percentile: Remove Top and Bottom 0.5%	
Model Filters:	Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1	
Metadata Filters:	Cooling Zone(s): All		Heating Fuel: Gas	
Thermostat Name: All	Heat Pump Baseline: All	Heating Zone(s): 2 - 6000 < Hdd < 7500, 1 - Hdd <= 6000, 3 - Hdd >= 7500	Multi Measure Filter: Single Measure Only	
Heat Pump Manufacturer: All	Heat Pump Adv. Controls or Commissioning: All			

89 Treatment Meters	102 +/- 28 Therms Average Normal Year Pre-Post Difference in Consumption per Participant	15 +/- 4 % Percent Normal Year Pre-Post Difference in Consumption per Participant	682 Mean Baseline Consumption (Gas)	108% Realization Rate
444 Site-level Matched Meters	101 +/- 29 Therms Average Savings Relative to Site-level Matched Comparison Group	15 +/- 4% Percent Savings Relative to Site-level Matched Comparison Group	680 Mean Baseline Consumption (Gas)	105% Realization Rate
409 Future Participant Meters	84 +/- 30 Therms Average Savings Relative to Future Participant Group	12 +/- 4% Savings Relative to Future Participant Group	700 Mean Baseline Consumption (Gas)	86% Realization Rate

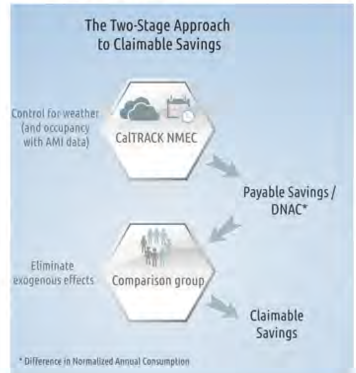
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

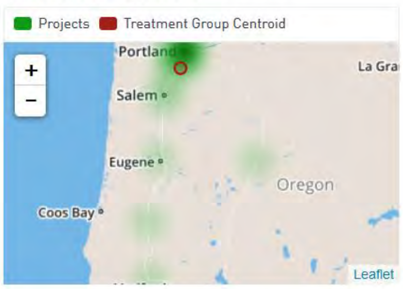
The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

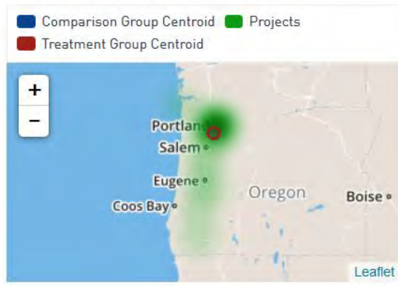
Two-Stage Approach



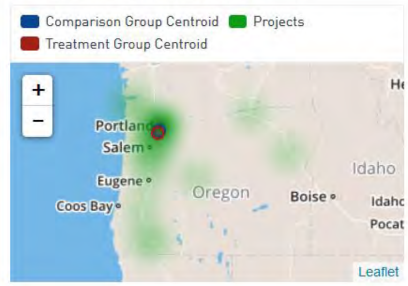
Treatment Site Locations



Site-level Matched Site Locations

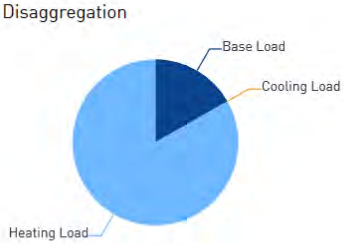


Future Participant Site Locations

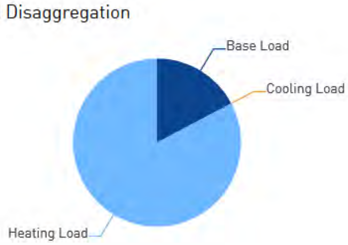


Treatment		Site-level Matched		Future Participant	
31.7 miles 80% of projects lie within this distance from treatment group centroid		2.3 miles Distance between treatment and comparison group centroids		4.4 miles Distance between treatment and future participant group centroids	
89 Meters	682 Mean Baseline Consumption (Gas)	444 Meters	680 Mean Baseline Consumption (Gas)	409 Meters	700 Mean Baseline Consumption (Gas)

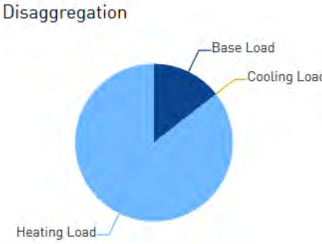
Load Disaggregation



Load Disaggregation



Load Disaggregation



2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

810	89	11%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

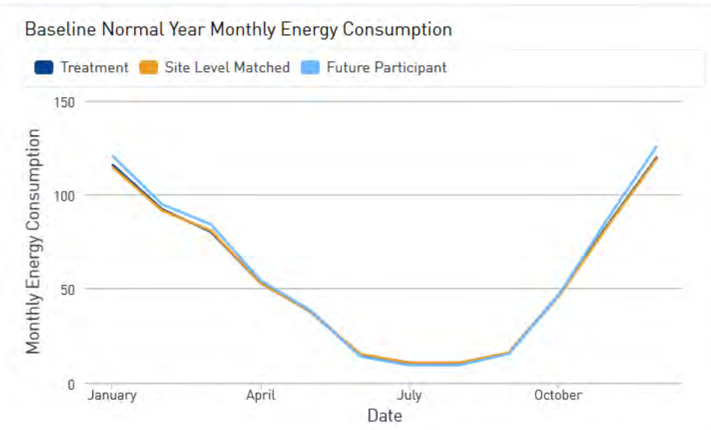
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2017 -- Fuel: Gas	--	810
Meters with valid consumption data in baseline and/or reporting periods.	--	50	760
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	629	131
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Gas	11	120
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 1 - Hdd <= 6000, 2 - 6000 < Hdd < 7500, 3 - Hdd >= 7500 -- Cooling Zone: All	5	115
Other measure-specific filters.	--	0	115
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	19	96
Meters with at least 5 site-level matched meters from the comparison group pool.	--	4	92
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	2	90

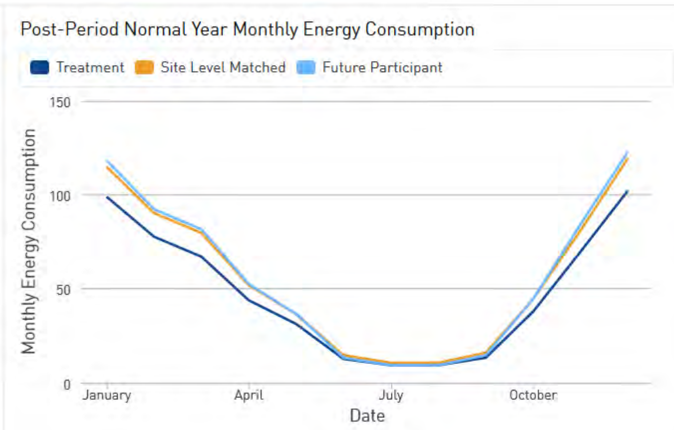
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	89
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	90
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	2	88
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	88

3. Modeling Results

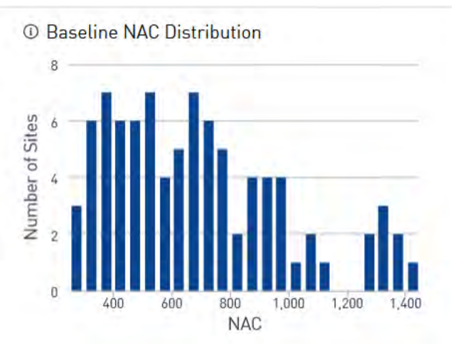
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

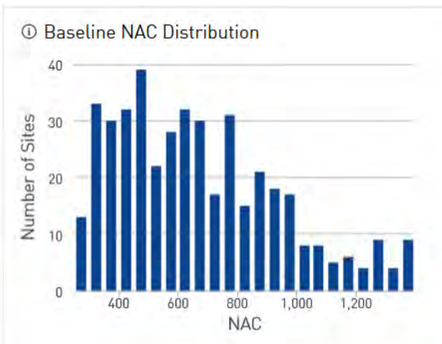




Treatment Group



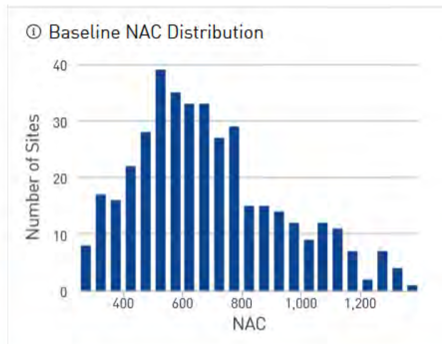
Site-level Matched Comparison Group



0.327

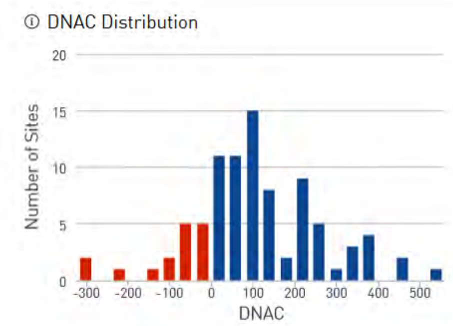
Annual Consumption p-value

Future Participant Group



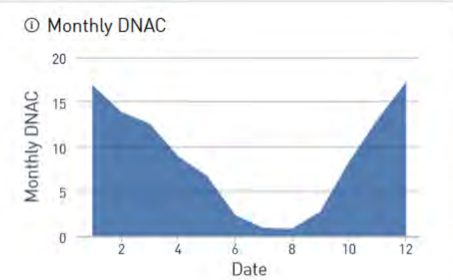
0.302

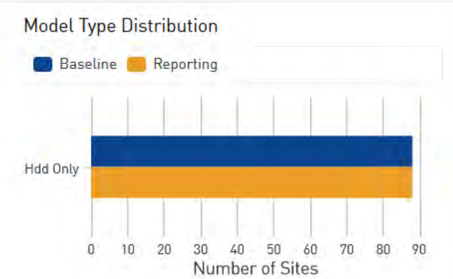
Annual Consumption p-value



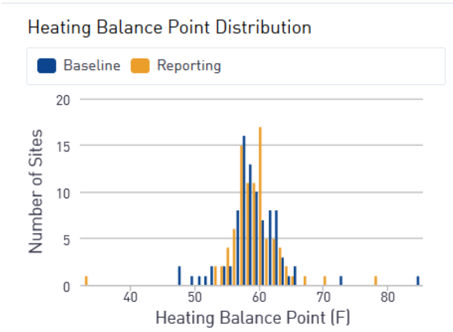
102 +/- 28 Therms
Average Difference in Normalized Annual Consumption per Participant

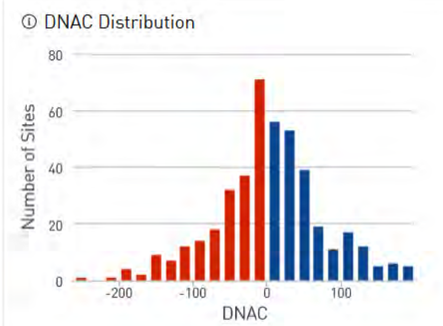
15 +/- 4 %
Difference in Normalized Annual Consumption as a Percent of Baseline





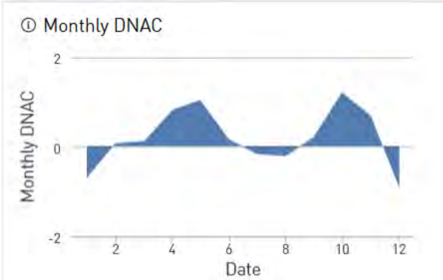
Treatment Group

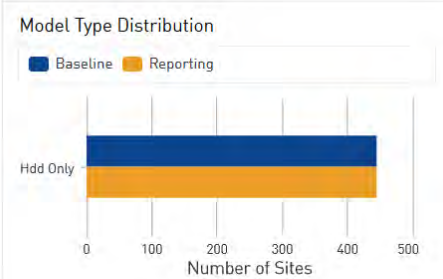




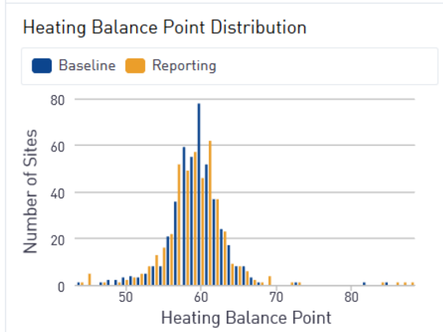
5 +/- 8 Therms
Average Difference in Normalized Annual Consumption per Participant

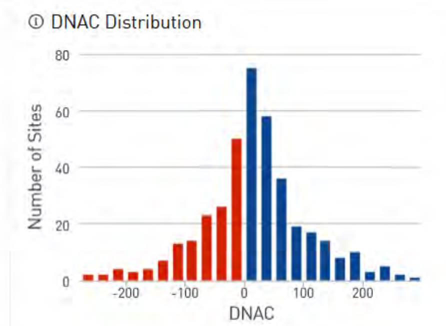
0 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline





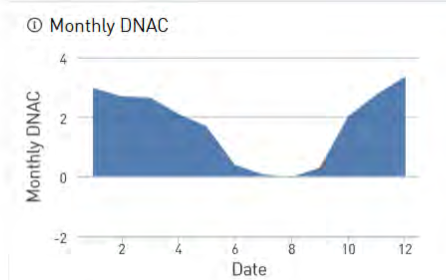
Site-level Matched Comparison Group

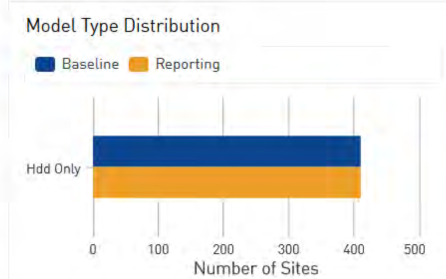




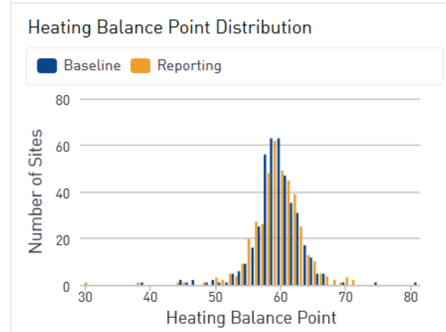
21 +/- 12 Therms
Average Difference in Normalized Annual Consumption per Participant

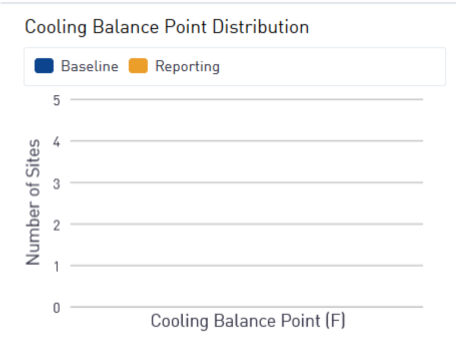
3 +/- 2 %
Difference in Normalized Annual Consumption as a Percent of Baseline

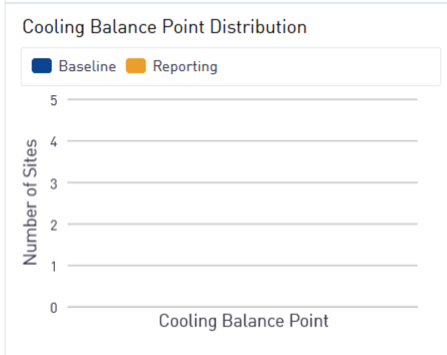


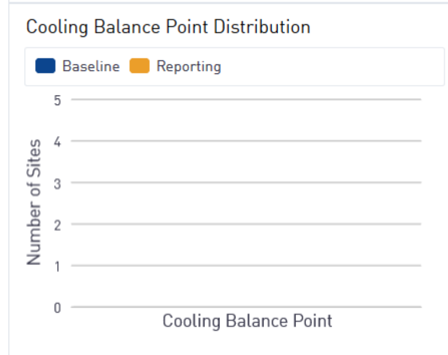


Future Participant Group









4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

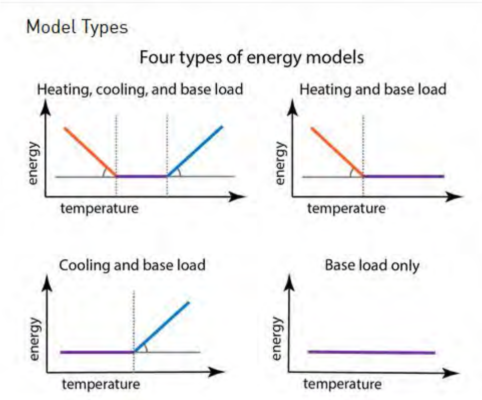
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

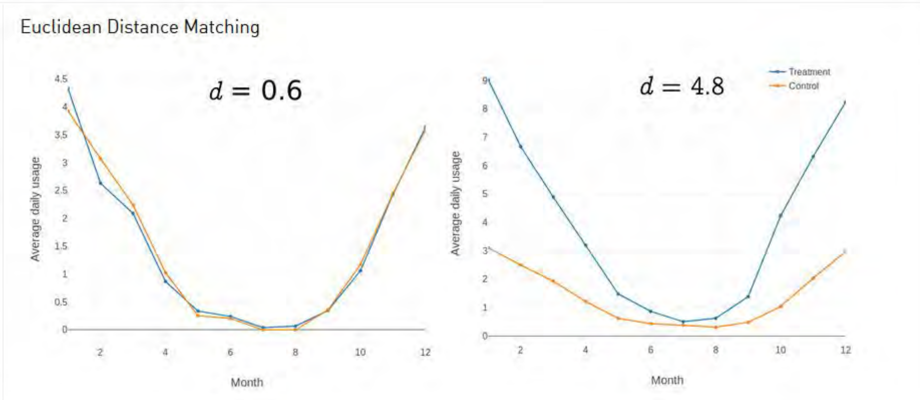
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.





Impact Evaluation Report

Electricity Impact of Ceiling Insulation in Program Year 2013, 2014, 2015, 2016, 2017

Result Summary

Measure: Ceiling Insulation	Program Year: 2013, 2014, 2015, 2016, 2017		Fuel: Electricity	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <75%	DNAC Percentile: None	Annual Consumption Percentile: Remove Top and Bottom 0.5%	
Model Filters:	Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1	
Metadata Filters:	Cooling Zone(s): All		Heating Fuel: Gas	
Thermostat Name: All / Heat Pump Baseline: All	Heating Zone(s): 1 - Hdd <= 6000		Multi Measure Filter: Single Measure Only	
Heat Pump Manufacturer: All	Heat Pump Adv. Controls or Commissioning: All			

238 Treatment Meters	251 +/- 162 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	3 +/- 2 % Percent Normal Year Pre-Post Difference in Consumption per Participant	8,680 Mean Baseline Consumption (Electricity)	531% Realization Rate
1,172 Site-level Matched Meters	180 +/- 171 kWh Average Savings Relative to Site-level Matched Comparison Group	2 +/- 2% Percent Savings Relative to Site-level Matched Comparison Group	8,523 Mean Baseline Consumption (Electricity)	382% Realization Rate
3,005 Future Participant Meters	155 +/- 169 kWh Average Savings Relative to Future Participant Group	2 +/- 2% Savings Relative to Future Participant Group	8,996 Mean Baseline Consumption (Electricity)	330% Realization Rate

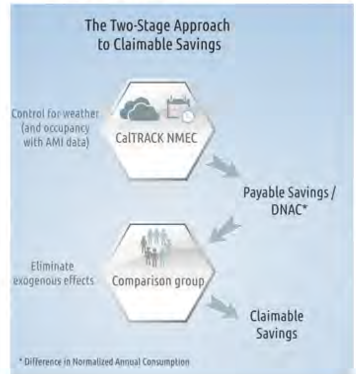
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

Two-Stage Approach

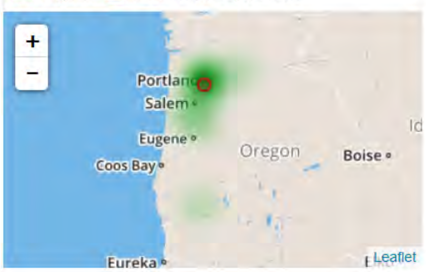


Treatment Group

The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality for modeling.

Treatment Site Locations

Projects | Treatment Group Centroid



34.1 miles

80% of projects lie within this distance from treatment group centroid

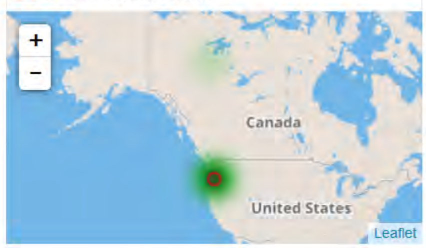
238	8,680
Meters	Mean Baseline Consumption (Electricity)

Site-level Matched Comparison Group

This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the group. Matching was performed using monthly consumption in the baseline period as detailed in the Methodology section.

Site-level Matched Site Locations

Comparison Group Centroid | Projects | Treatment Group Centroid



0.2 miles

Distance between treatment and comparison group centroids

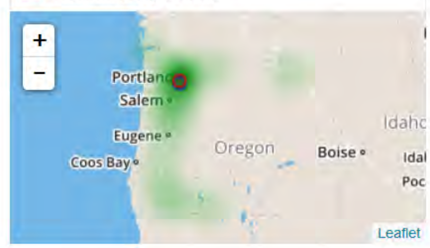
1,172	8,523
Meters	Mean Baseline Consumption (Electricity)

Future Participant Group

The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations

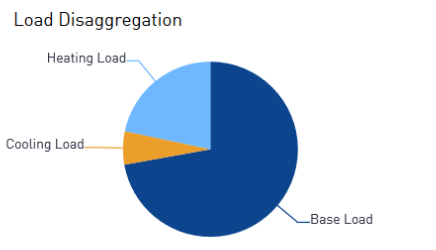
Comparison Group Centroid | Projects | Treatment Group Centroid

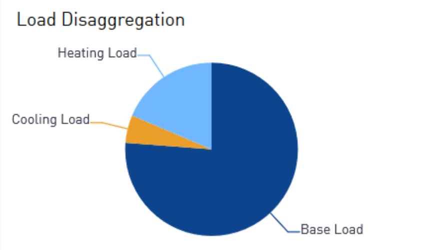


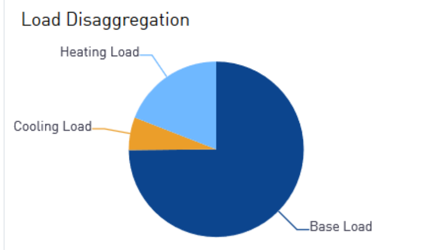
5.1 miles

Distance between treatment and future participant group centroids

3,005	8,996
Meters	Mean Baseline Consumption (Electricity)







2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

5,412	238	4.4%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

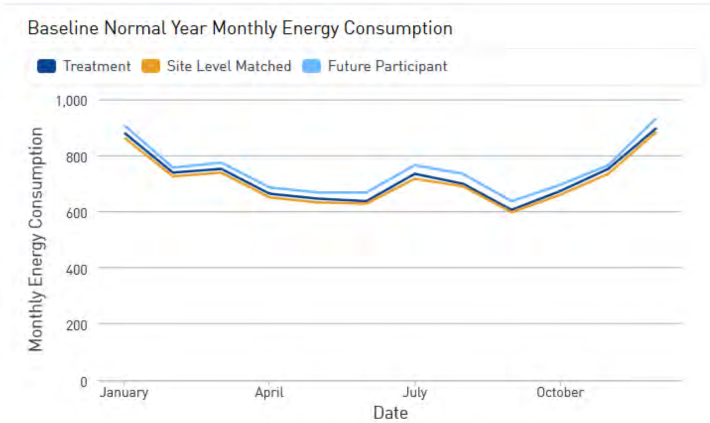
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2014, 2013, 2015, 2016, 2017 -- Fuel: Electricity	--	5,412
Meters with valid consumption data in baseline and/or reporting periods.	--	104	5,308
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	4,611	697
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Gas	216	481
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 1 - Hdd <= 6000 -- Cooling Zone: All	13	468
Other measure-specific filters.	--	0	468
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	96	372
Meters with at least 5 site-level matched meters from the comparison group pool.	--	24	348
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	13	335

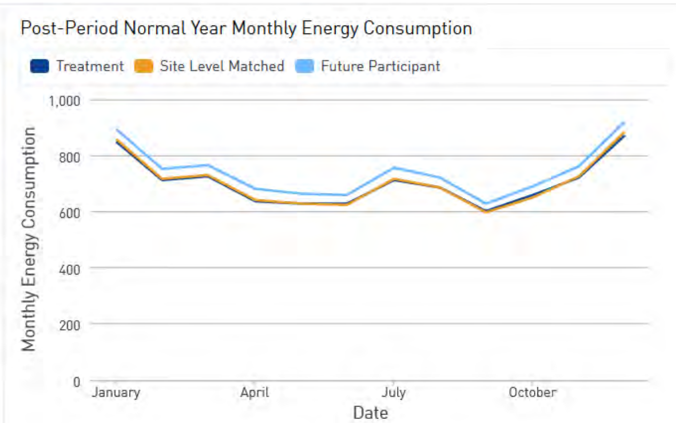
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: None	0	335
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	1	334
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	96	238
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	238

3. Modeling Results

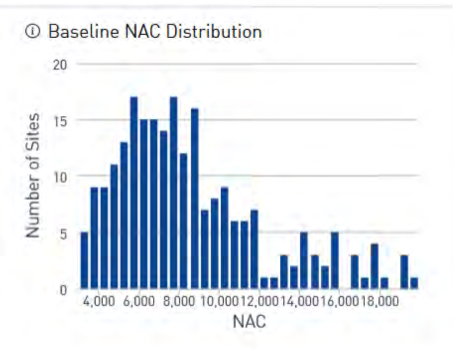
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

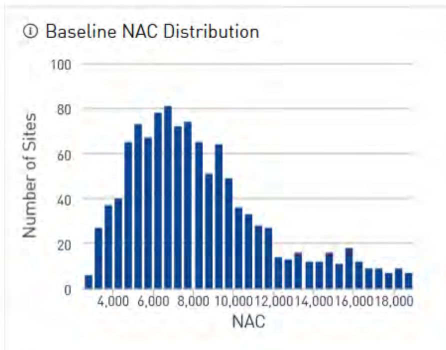




Treatment Group



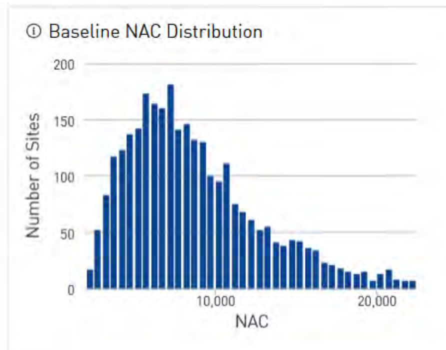
Site-level Matched Comparison Group



0.178

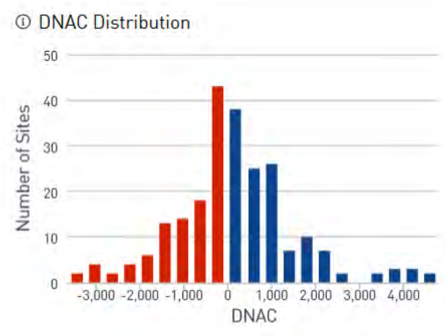
Annual Consumption p-value

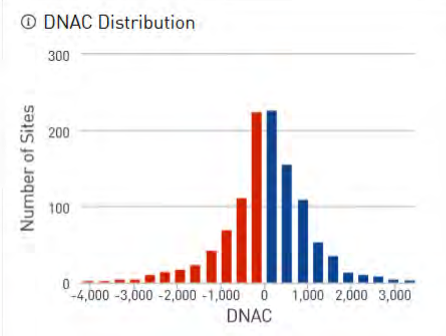
Future Participant Group

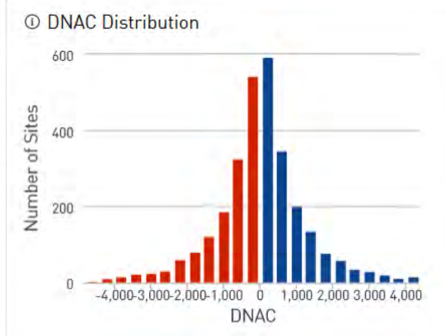


0.189

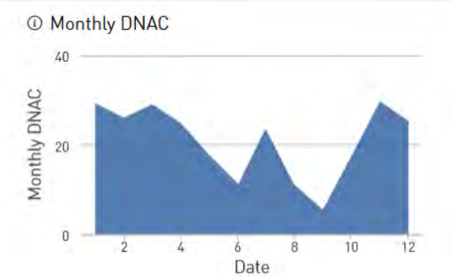
Annual Consumption p-value

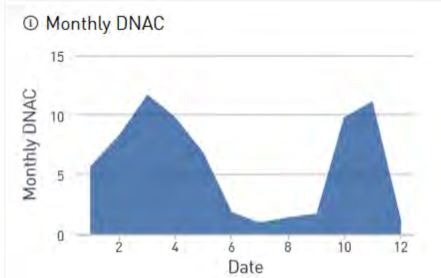


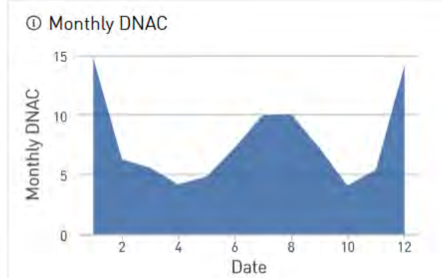


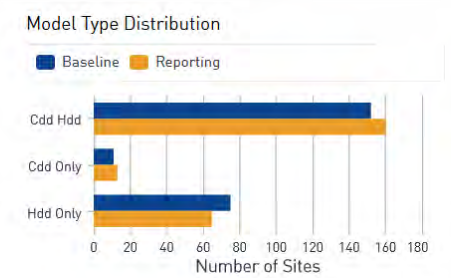


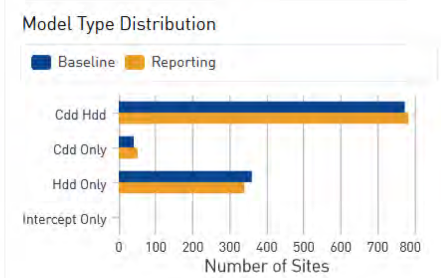
251 +/- 162 kWh	3 +/- 2 %	70 +/- 53 kWh	1 +/- 1 %	95 +/- 49 kWh	1 +/- 1 %
Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline	Average Difference in Normalized Annual Consumption per Participant	Difference in Normalized Annual Consumption as a Percent of Baseline

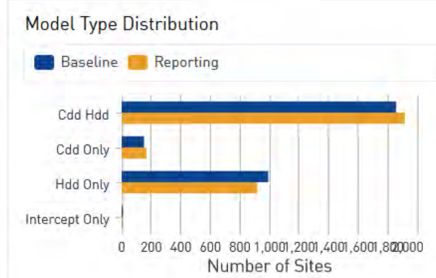




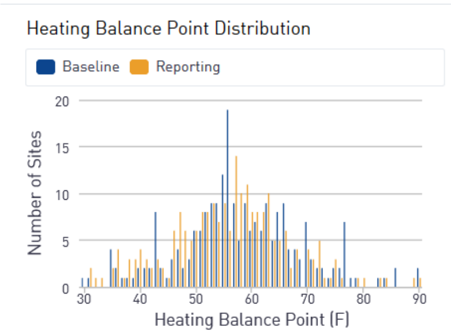


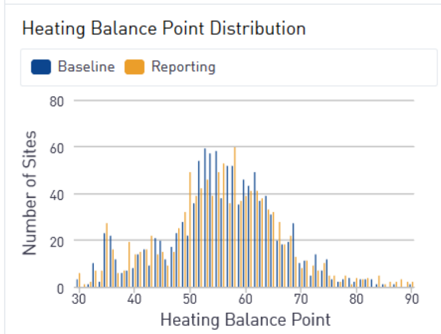


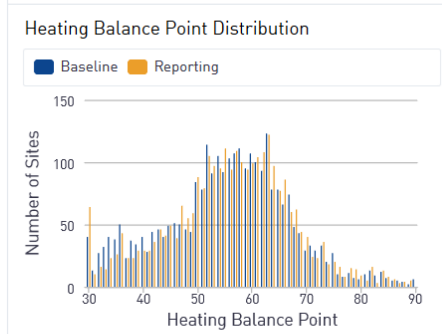


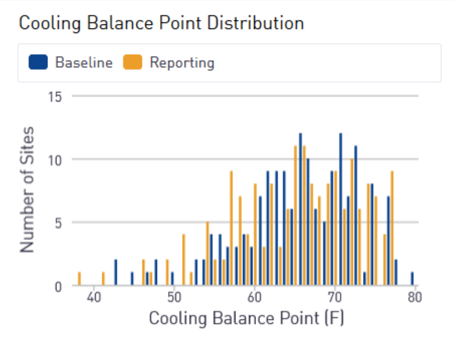


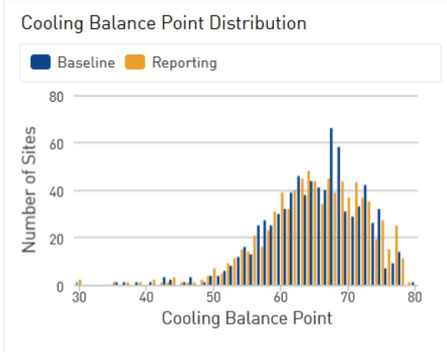
Treatment Group	**Site-level Matched Comparison Group**	**Future Participant Group**

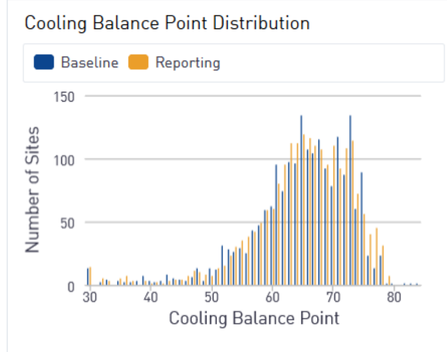












4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

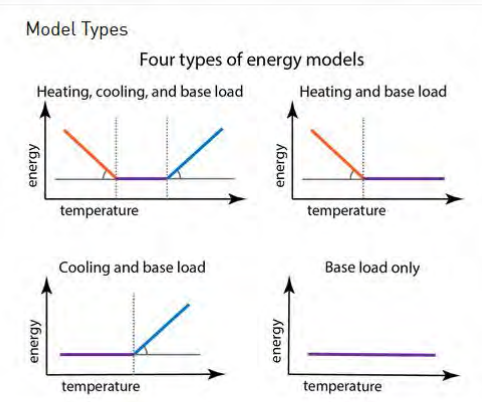
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

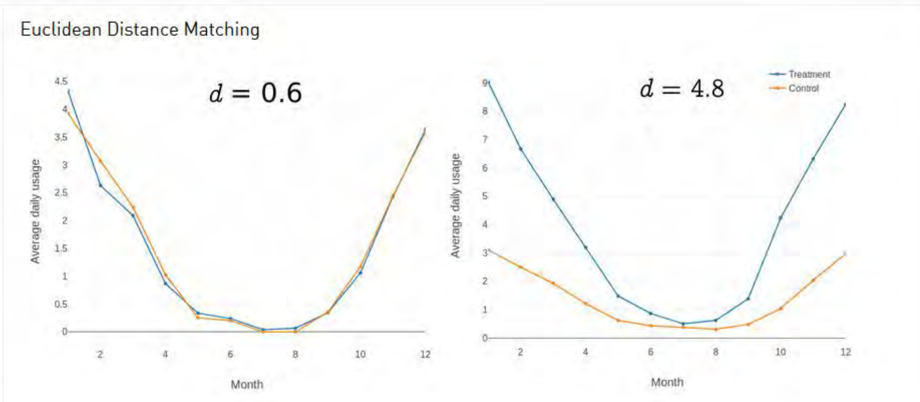
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.





Impact Evaluation Report

Electricity Impact of Ceiling Insulation in Program Year 2013, 2014, 2015, 2016, 2017

Result Summary

Measure: Ceiling Insulation	Program Year: 2013, 2014, 2015, 2016, 2017		Fuel: Electricity	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <75%	DNAC Percentile: None	Annual Consumption Percentile: Remove Top and Bottom 0.5%	
Model Filters:	Period Length: 11 Months or Longer	R-Squared: >0.5	CV(RMSE): < 1	
Metadata Filters:	Cooling Zone(s): All		Heating Fuel: Electricity	
Thermostat Name: All / Heat Pump Baseline: All	Heating Zone(s): 1 - Hdd <= 6000		Multi Measure Filter: Single Measure Only	
Heat Pump Manufacturer: All / Heat Pump Adv. Controls or Commissioning: All				

107 Treatment Meters	1897 +/- 550 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	11 +/- 3 % Percent Normal Year Pre-Post Difference in Consumption per Participant	16,991 Mean Baseline Consumption (Electricity)	75% Realization Rate
517 Site-level Matched Meters	1559 +/- 583 kWh Average Savings Relative to Site-level Matched Comparison Group	9 +/- 3% Percent Savings Relative to Site-level Matched Comparison Group	16,353 Mean Baseline Consumption (Electricity)	62% Realization Rate
1,208 Future Participant Meters	1910 +/- 569 kWh Average Savings Relative to Future Participant Group	11 +/- 3% Savings Relative to Future Participant Group	15,296 Mean Baseline Consumption (Electricity)	76% Realization Rate

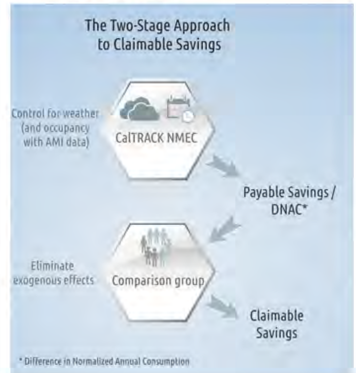
1. Introduction

This report contains the results of applying the two-stage approach (informed by the DOE's uniform methods chapter on whole building analysis) for calculating claimable savings to the selected portfolio of energy efficiency projects (see Figure). This approach begins with identification of two comparison groups for the treatment sample: (a) a site-level matched comparison group and (b) a future participant group. These groups are described below along with summary statistics (site locations, sample size, baseline consumption and baseline load disaggregation).

The CalTRACK methods are then applied to arrive at site-level savings, normalized for weather, and reflective of energy consumption changes for customers at the meter. Using a difference of differences for the treatment group with each comparison group accounts for population-level consumption changes (e.g. economic changes, rate changes, natural energy efficiency adoption etc.). The methods contained within this report are the outcome of a recent peer-reviewed study completed by Energy Trust of Oregon and Open Energy Efficiency (see "Methodology" section for more details).

The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model outputs and Difference in Normalized Annual Consumption (DNAC) results
Section 4. Methodology - Description of methods used in this report

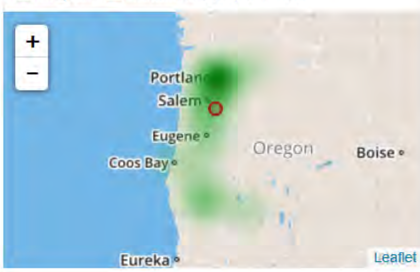
Two-Stage Approach



Treatment Group
The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality for modeling.

Treatment Site Locations

Projects Treatment Group Centroid



164.1 miles

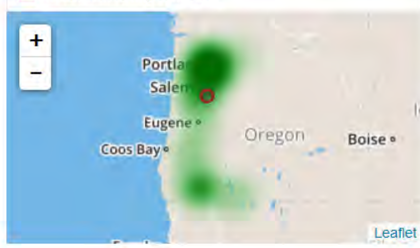
80% of projects lie within this distance from treatment group centroid

107	16,991
Meters	Mean Baseline Consumption (Electricity)

Site-level Matched Comparison Group
This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the group. Matching was performed using monthly consumption in the baseline period as detailed in the Methodology section.

Site-level Matched Site Locations

Comparison Group Centroid Projects
Treatment Group Centroid



1.0 miles

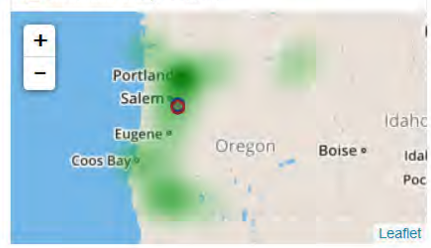
Distance between treatment and comparison group centroids

517	16,353
Meters	Mean Baseline Consumption (Electricity)

Future Participant Group
The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations

Comparison Group Centroid Projects
Treatment Group Centroid

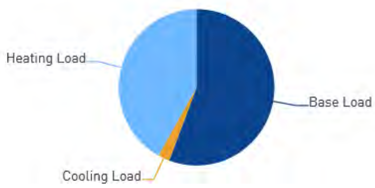


3.5 miles

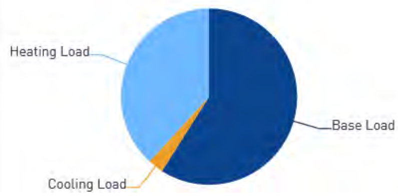
Distance between treatment and future participant group centroids

1,208	15,296
Meters	Mean Baseline Consumption (Electricity)

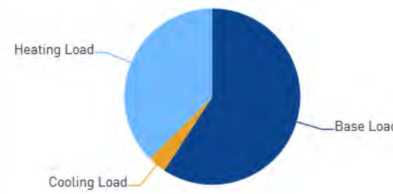
Load Disaggregation



Load Disaggregation



Load Disaggregation



2. Data Preparation

Consumption data preparation and cleaning followed best practices defined in the CalTRACK 2.0 billing methods. Some key aspects of the data cleaning process are highlighted here; please see the resources section for links to more detailed documentation. The initial and final sample sizes are shown below along with the percent of the treatment population that is represented by the sample. The sample attrition table shows the impact of each filtering criterion on sample size.

5,412	107	2%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Ceiling Insulation -- Year: 2014, 2013, 2015, 2016, 2017 -- Fuel: Electricity	--	5,412
Meters with valid consumption data in baseline and/or reporting periods.	--	104	5,308
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: Single Measure Only	4,611	697
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	487	210
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 1 - Hdd <= 6000 -- Cooling Zone: All	19	191
Other measure-specific filters.	--	0	191
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	64	127
Meters with at least 5 site-level matched meters from the comparison group pool.	--	5	122
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <75%	4	118

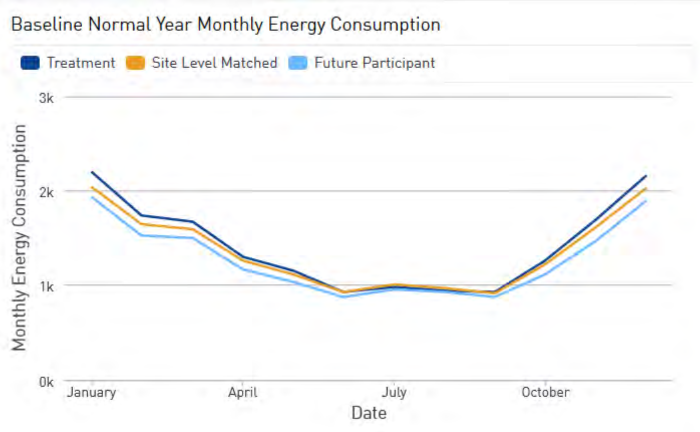
RECURVE

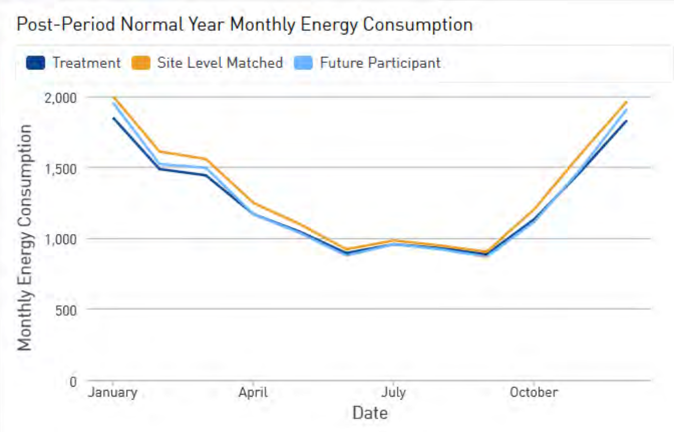
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: None	0	118
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	118
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	11	107
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	107

3. Modeling Results

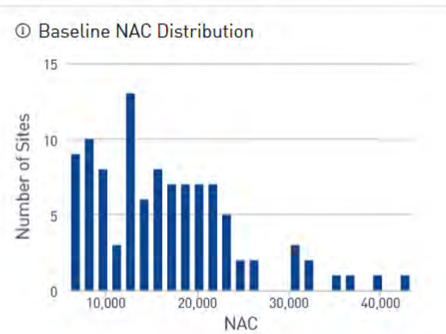
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting periods.

Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional layer of analysis.

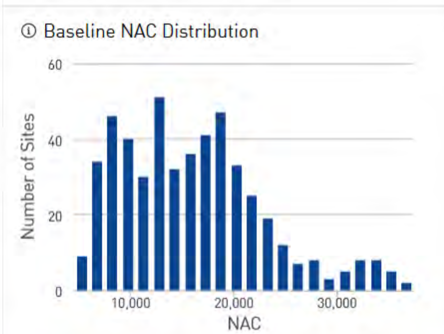




Treatment Group



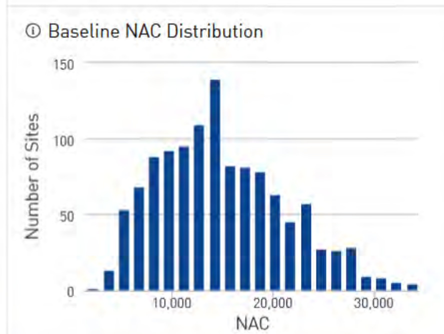
Site-level Matched Comparison Group



0.206

Annual Consumption p-value

Future Participant Group

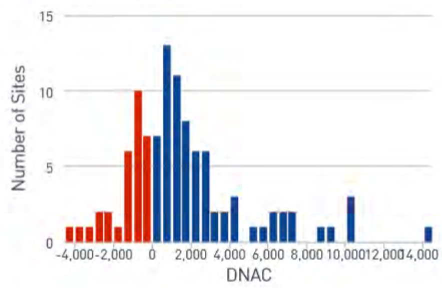


0.151

Annual Consumption p-value

Treatment Group

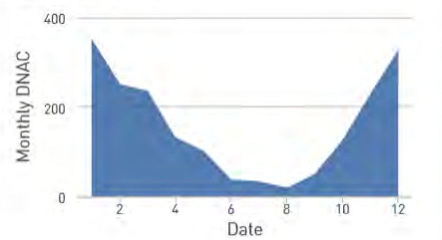
DNAC Distribution



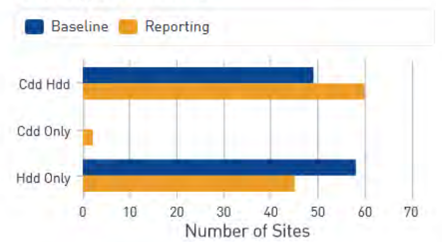
1897 +/- 550 kWh
Average Difference in Normalized Annual Consumption per Participant

11 +/- 3 %
Difference in Normalized Annual Consumption as a Percent of Baseline

Monthly DNAC

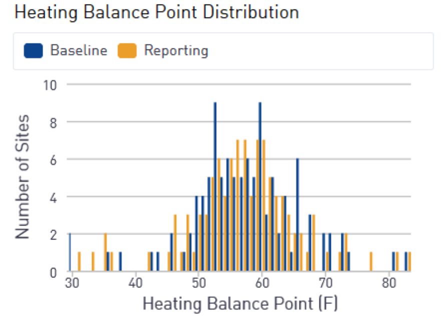


Model Type Distribution



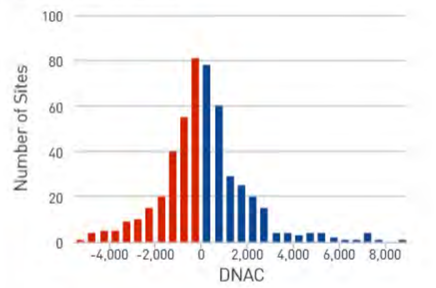
Treatment Group

Heating Balance Point Distribution



Site-level Matched Comparison Group

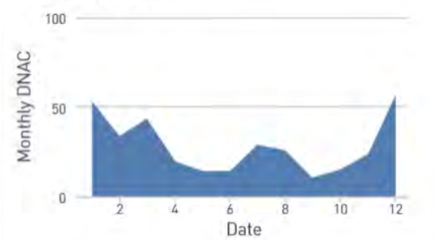
DNAC Distribution



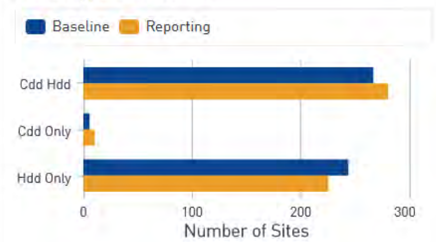
338 +/- 194 kWh
Average Difference in Normalized Annual Consumption per Participant

2 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline

Monthly DNAC

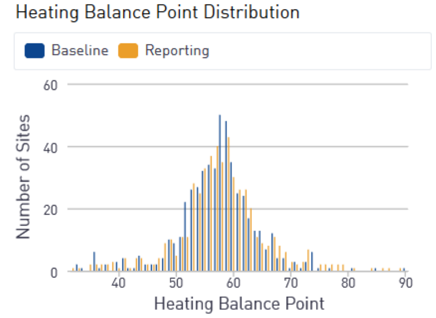


Model Type Distribution



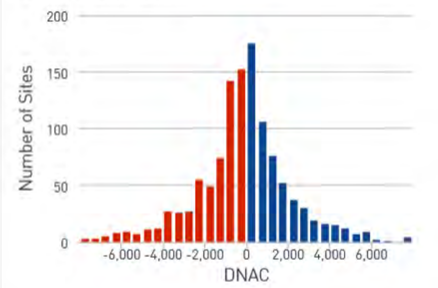
Site-level Matched Comparison Group

Heating Balance Point Distribution



Future Participant Group

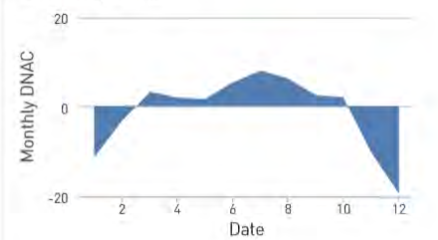
DNAC Distribution



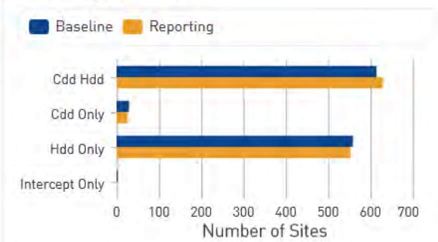
-13 +/- 147 kWh
Average Difference in Normalized Annual Consumption per Participant

-0 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline

Monthly DNAC

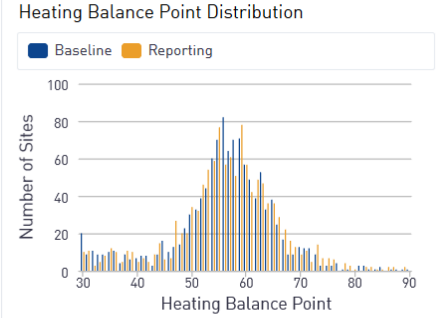


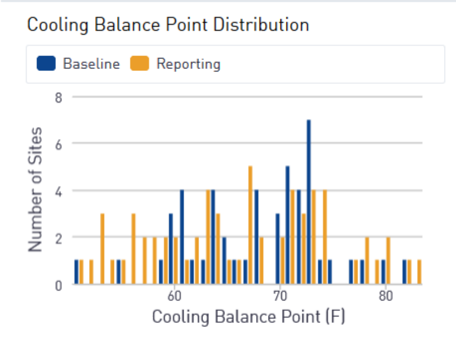
Model Type Distribution

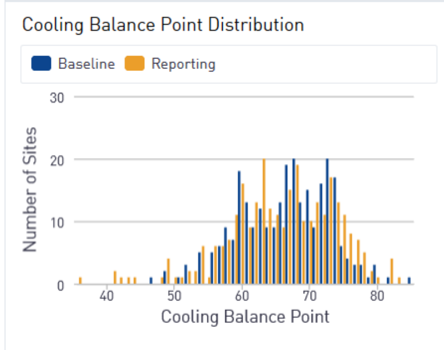


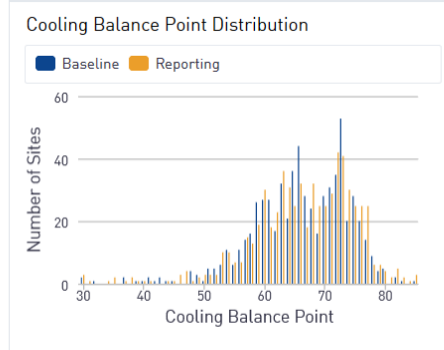
Future Participant Group

Heating Balance Point Distribution









4. Methodology

CalTRACK and Comparison Group Methods
Documentation: docs.caltrack.org
Code: https://github.com/energy-market-methods/caltrack

Data Preparation

Baseline period: Since the predicted baseline may be unstable with different baseline period lengths, which may, in turn, affect calculated savings, the consensus of the CalTRACK 2.0 working group was to set the maximum baseline period at 12 months, since the year leading to the energy efficiency intervention is the most indicative of recent energy use trends and prolonging the baseline period increases the chance of other unmeasured factors affecting the baseline. In addition, CalTRACK uses a minimum 12-month baseline by default.

Blackout period: The blackout period refers to the time period between the end of the baseline period and the beginning of the reporting period. In this analysis, it is specified to coincide with the project installation time period, meaning that the billing period that contains the project installation date is dropped from the analysis.

Analysis periods: Different portions of the analysis used different time periods of consumption data, therefore, it is useful to clearly define these time periods and where they were used. Consider a project with an installation date on a particular day d in a particular month m in a particular program year y. The year before the program year is labelled as y-1, the year prior to that as y-2 and so on, while the years following the program year are labelled y+1, y+2 etc. In all cases, the billing period that contains the project installation was dropped from the analysis. Other sections of the analysis use the following time periods:
- *Treatment and site-level matched groups:* Baseline period includes the 12 months preceding the installation billing period. Reporting period includes the 12 months following the installation billing period.
- *Future participant group:* Baseline period is the calendar year preceding the program year (Year *y-1*). Reporting period is the program year itself (Year *y*).
- Site-level consumption matching was performed using the 12 months of data immediately prior to the project installation date.
- Equivalence tests were performed using data from the previous calendar year (y-1).

Modeling

Weather Normalization: Weather normalization of billing data in CalTRACK follows certain model foundations in literature (PRISM, ASHRAE Guideline 14, IPMVP Option C and the Uniform Methods Project for Whole Home Building Analysis). Building energy use is modeled as a combination of base load, heating load, and cooling load. Heating load and cooling load are assumed to have a linear relationship with heating and cooling demand, as approximated by heating and cooling degree days, beyond particular heating and cooling balance points. A number of candidate OLS models are fit to the consumption data using different combinations of heating and cooling balance points (ranging from 30 to 90 F) and different sets of independent variables. The model with the highest adjusted R-squared that contains strictly positive coefficients is selected as the final model and used to calculate normalized energy usage.

Model Types: CalTRACK specifies a linear relationship between energy use and temperature as reflected in the building consumption profile. In the most generic case, a model would include an intercept term, a heating balance point and heating slope coefficient, and a cooling balance point and a cooling slope coefficient. Depending on the fuel a building uses for heating or cooling or its consumption patterns, models with a single temperature coefficient and balance point (i.e., heating or cooling) may be more appropriate.

Difference in Normalized Annual Consumption (DNAC): The DNAC is calculated by using two CalTRACK regression models in conjunction with Typical Meteorological Year (TMY3) weather data, as follows:
- Two models are fit to the consumption data - one model for the baseline (pre-intervention) period and one for the reporting (post-intervention) period.
- Long-term heating and cooling degree days based on TMY3 data are substituted in both regression equations to calculate the Normalized Annual Consumption (NAC) for each period. TMY3 data is maintained by NREL and includes weather averages for 1020 locations in the US between 1991-2005.
- DNAC is determined by subtracting the two NACs (DNAC = Baseline NAC - Reporting NAC).

Disaggregation: Disaggregated loads are calculated from the different components of the statistical model fit. The weather sensitive components (heating and cooling load) are calculated by multiplying the relevant model coefficients (beta_hdd or beta_cdd) by the total degree days in a normal weather year (total HDD or CDD). For each site, the total HDD or CDD can be calculated using that site's estimated degree day balance points (also an output of the model) and the temperature for its closest weather station. The base load is estimated by multiplying the intercept of the statistical model by the number of days (365 for a full year).

Savings calculation: Savings are calculated by subtracting the DNAC for either comparison group from the DNAC for the treatment group.

Savings Uncertainty: Uncertainty presented in this analysis is calculated using the ASHRAE Guideline 14 formulation for aggregating the prediction uncertainty of point estimates in a time series. It is calculated at a 90% confidence level. The total uncertainty at the site-level is calculated using the sum of squares of the baseline and reporting models. Other aggregate uncertainty values (e.g. for a portfolio or for a difference-in-differences estimate) are also aggregated using the square root of the sum of squares.

Comparison Group Generation

Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting n matches (n=5 in this analysis) from the comparison group pool with the shortest distance d to the treatment group customer under consideration. The pool is limited to non-participants within the same zipcode as the treatment group customer. The distance d is, in essence, a way to reduce 12 monthly consumption differences between any two customers to one metric (see Figure). In the present analysis, we selected (without replacement) five nearest neighbors for each treatment site based on the Euclidean distance of monthly consumption.

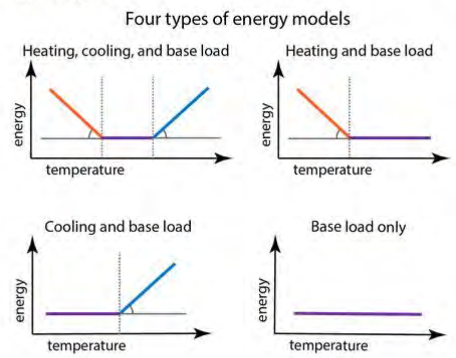
Future Participant Groups: Comparison groups comprising future participants are considered to be representative of participants in most aspects (observable and non-observable). For example, future participants are known to be eligible to receive the measure, and for some measures, they may have the same baseline equipment as the participants. Future participants have the same propensity to participate in the program as participants, thus reducing or eliminating self-selection bias, something that is otherwise difficult to control for in a quasi-experimental study. More comprehensive data is typically collected for future participants, allowing for potentially better matching and more insightful analysis. From a practical perspective, future participant groups may be difficult to construct for all measures, unless a program has been running for multiple years and is considered stable with sufficient data collection over the analysis period. Sample sizes for the comparison group may also be constrained if using future participants.

Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding bin in the comparison group pool of future participants.

Geographical screen: For the site-level matched group, only sites within the same zipcode as the treatment site were considered as potential comparison group matches.

Sampling method: In all cases where sampling was required from the comparison group, sampling was performed without replacement.

Model Types



Euclidean Distance Matching