MEMO

Date: March 23, 2020
To: Board of Directors
From: Dan Rubado, Evaluation Project Manager
Andrew Shepard, Sr. Project Manager – Residential
Jackie Goss, Sr. Planning Engineer
Subject: Wrap-up of the Extended Capacity Heat Pump Pilot

SUMMARY

Energy Trust launched a pilot to learn more about Extended Capacity Heat Pumps (ECHPs) in 2018. ECHPs are a quickly emerging technology at the top end of the variable capacity heat pump (VCHP) market, often referred to as cold climate units. The primary pilot goal was to learn more about the performance of these systems compared with standard VCHP models and validate preliminary energy savings estimates. The pilot investigated sources of savings, including standby energy use, low temperature capacity, auxiliary heat use, defrost cycles, and cooling. The pilot also investigated optimal sizing, installation, and commissioning strategies to increase savings. Pilot activities included providing incentives for ECHPs, recruiting and training contractors, and conducting market research, a power metering study, and electricity billing analysis.

The pilot showed that ECHPs had improved energy performance at cold outdoor temperatures and produced 1,300 kWh per year of energy savings, on average, over standard VCHPs installed in comparable scenarios across all heating zones. This estimate is reasonably precise and was corroborated by power metering of a small sample of ECHP and VCHP systems. ECHPs also provide a small amount of peak demand savings in the winter. The pilot did not identify any standby mode or cooling savings. Average installed costs of ECHPs were $15,790, resulting in an incremental cost above standard VCHPs of roughly $1,100.

Best practices for sizing, commissioning, and installation have not been established for VCHPs, but some recommendations are beginning to emerge. It may be possible to further improve the energy performance of ECHP systems over time by encouraging contractors to disconnect auxiliary heat in some cases, developing commissioning and setup best practices, providing guidance on proper system sizing, and working with manufacturers to develop energy saving features like improved crank case heaters and defrost cycles. If extended capacity technology is successful and Energy Trust can increase its penetration in the top end of the heat pump market, then the technology may work its way into less expensive heat pump tiers. Extended capacity technology has the potential for broad market transformation, improving the energy performance of heat pumps across the board.

Energy Trust is currently wrapping up its coordinated research activities on EHCP systems and winding down the pilot. Energy Trust will create a new deemed savings measure and incentive for ECHPs based on the pilot findings. The Residential program plans to begin supporting ECHP systems more broadly with a market-based incentive by July 2020. Energy Trust will conduct research and evaluation on ECHPs in the coming years as needed and encourage manufacturers to bring additional technology improvements to market.

Introduction

Heating systems are one of the largest contributors to household energy use in the Pacific Northwest and have a long replacement cycle. Heat pump technology has evolved significantly over the past several decades and has become increasingly efficient. Energy Trust currently offers incentives to convert from an electric resistance forced air furnace to a heat pump but does not offer incentives to upgrade an existing heat pump to a high efficiency heat pump due to cost-effectiveness limitations. However, there is significant energy savings potential in increasing the efficiency of the heat pump market.

Within the top tier of efficient heat pump technology—variable capacity heat pumps—certain premium models, referred to hereafter as "extended capacity heat pumps" or ECHPs, may save additional energy when compared with their standard, variable capacity counterparts. Extended capacity systems are defined as high efficiency, variable capacity, ducted heat pumps that maintain at least 85% of their capacity at 17 degrees Fahrenheit, compared with their full capacity at 47 degrees. The assumed baseline equipment type for an ECHP system is a similar, high efficiency, variable capacity heat pump (VCHP) system that aligns with the Regional Technical Forum's top tier heat pump designation but doesn't perform as well as an ECHP at low outdoor temperatures. Both ECHP and VCHP systems stand out from the less efficient portion of the market by utilizing variable capacity, inverter driven compressors, with a Heating Season Performance Factor (HSPF) rating of 10.0 or greater.

During Energy Trust's spring 2018 Trade Ally Forums, trade allies suggested Energy Trust could play a valuable role in creating market differentiation for high efficiency heat pump technologies and improving installation practices. Trade allies also expressed they wanted Energy Trust to introduce new prescriptive incentives for ECHPs to help push the heat pump market. ECHP products were already in the market at that point and appeared to be a commercially viable technology. Interviews with installers later in 2018 indicated a market share of approximately 20% for ECHP units within the high efficiency, variable capacity segment of the market. However, some installers expressed uncertainty about installing this equipment because they were unsure of the best practices. Given that this technology is relatively new to the market, and the optimal sizing, installation and commissioning practices are not widely known, there may be room to improve ECHP performance, energy savings and cost-effectiveness over time.

In 2018, Energy Trust's Residential Program Management Contractor, CLEAResult, completed a preliminary analysis of AMI[1] data obtained from PGE for a small sample of homes that installed heat pumps in past years that were identified as either ECHP or VCHP systems. CLEAResult compared post-installation hourly electricity usage between these two groups. This initial analysis suggested ECHP systems performed significantly better during cold periods and might also save energy during mild temperatures when systems were likely in standby mode. However, the findings of this analysis were inconclusive because it was based on a convenience sample and only post-installation energy usage data were available. The impact of ECHPs compared with VCHPs could not be isolated from other factors, such as differences in baseline energy usage and home characteristics. The results suggested further research was warranted but that ECHP energy savings could be borderline cost-effective if corroborated. The analysis also raised additional research questions, best investigated through field data collection and large sample utility billing analysis.

In quarter four of 2018, Energy Trust launched a pilot to learn more about ECHP systems. The primary goal was to learn more about the performance of ECHP systems compared with standard

[1] Advanced Metering Infrastructure, or AMI, enables the collection of short-interval electricity usage data from homes and businesses.

VCHPs and to validate the preliminary energy savings estimate developed by CLEAResult. The pilot also investigated the potential sources of savings (e.g., standby energy usage, low temperature capacity, auxiliary heat control, defrost cycles, cooling) for extended capacity units and whether there are ways to further optimize installations to yield additional savings. The pilot was intended to provide a better understanding of this technology, determine if it could be a new source of cost-effective savings in the heat pump market and inform future measure design.

If ECHP cold weather energy performance and savings are proven, they may have significant future savings potential in Oregon. Energy Trust could help increase the adoption and market share of ECHPs within the high efficiency, variable capacity heat pump category. That could help prime the market and accelerate the adoption of higher efficiency heat pump technology across all tiers of the heat pump market, not just the high end, resulting in broad market transformation.

Research Questions

The overarching goals of the pilot were to assess the viability of ECHPs as an efficiency technology that Energy Trust could support to achieve additional energy savings compared with standard VCHP systems. The pilot had the following specific research questions:

- What are the energy savings for extended capacity heat pumps over other variable capacity heat pumps?
- What are the operational characteristics of these extended capacity heat pumps that can provide additional energy savings compared to other variable capacity heat pumps? Do they vary by manufacturer? (e.g. standby usage, aux heat usage, defrost cycles.)
- What are the sizing, commissioning, and setup practices for best energy performance while not negatively affecting occupant comfort?
- How does sizing, commissioning and installation differ between extended capacity and other variable capacity heat pumps?
- What is the incremental cost of extended capacity heat pumps over other variable capacity heat pumps?

Summary of Pilot Research Activities

Pilot implementation activities were conducted by the Residential program team at CLEAResult. These activities included developing a pilot incentive to promote ECHPs, recruiting trade ally contractors, holding trainings and processing incentives. The Residential program introduced an incentive offer for ECHPs in quarter one of 2019 to correspond with the pilot time period. This was an incremental savings measure for heat pump conversions or upgrades above an assumed baseline of a standard VCHP system. The program began recruiting trade ally contractors to install ECHPs across Energy Trust electric service territory, representing a spectrum of heat pump manufacturers. Uptake of the pilot measure was slow at first but gained momentum after the program worked with distributors to hold lunch-and-learn events to help educate and recruit contractors. In addition, CLEAResult and Energy Trust conducted several research activities to learn about ECHP systems and their differences with standard VCHP systems to answer the pilot research questions. This research is summarized in this memo and described in more detail in the attached reports.

1. <u>Market Research:</u> CLEAResult conducted interviews with heat pump installers, distributors and manufacturers and attended several installation site visits with contractors to learn more about how systems were being sized, installed and commissioned and best practices to

achieve optimal energy performance. They also reviewed available installation guidelines from manufacturers and collected and summarized ECHP project data related to the pilot. In addition, they summarized ECHP installed costs from the pilot period and compared these with the installed costs for VCHP systems.

2. Power Metering Study: Energy Trust leveraged a field study being conducted by Bonneville Power Authority (BPA) to conduct power metering in a small sample of homes with heat pump systems to better understand the energy performance of ECHP and standard VCHP systems. Energy Trust hired SBW Consulting to recruit homes in BPA's study sample that had ECHP and VCHP systems to do circuit level monitoring for a period of roughly eight months. Only two ECHP and six VCHP homes with electric resistance backup heat were successfully recruited and monitored. The monitored homes were west and east of the Cascades. SBW installed metering equipment in February 2019 and retrieved it at the end of August 2019. Heat pump compressor and air handler power were recorded at one-minute intervals. Refrigerant vapor temperature and outside air temperature were also recorded. These data were used to model heating, cooling, standby and total annual electricity usage for ECHP and VCHP systems in a typical weather year and to compare their energy performance. While the results are somewhat anecdotal, they do provide insight into the operation and performance of these systems in a range of weather conditions.

3. Utility Billing Analysis: Energy Trust conducted a utility billing analysis, using an analysis tool built by Recurve Analytics, to evaluate the incremental electricity savings of ECHP systems installed in single-family homes compared with similar homes that installed standard VCHP systems. We analyzed electricity usage for heat pump systems installed from 2015 to 2018. The heat pump projects included in this analysis were a combination of conversions from electric forced air furnaces and upgrades from less efficient heat pumps. Homes heated with gas, propane and other fuels were excluded to the extent possible. The intent was to isolate the electricity impact of an ECHP system in electrically-heated homes. Monthly utility billing data were used to conduct pre/post analyses of whole home energy usage. Energy usage data were weather-normalized using typical meteorological year data. Changes in normalized annual energy usage were then evaluated against changes in a comparison group. The comparison group was created by matching each ECHP project to similar VCHP projects, based on monthly electricity usage. Using these methods, we estimated the average annual electricity savings resulting from an ECHP project compared with a standard VCHP project.

Key Findings

Market research

Most market actors refer to ECHPs as "cold climate" heat pumps and don't use the term "extended capacity." Most manufacturers also define this class of variable capacity equipment slightly differently than Energy Trust's working definition. They define them as variable capacity units that can maintain their maximum 47 degree heating capacity down to 25 degrees or lower, although Mitsubishi sets this threshold at 5 degrees. During interviews, contractors noted they did not sell variable capacity heat pump systems, including cold climate units, based on efficiency or energy savings, but rather based on comfort benefits. Due to their longer run times at low speeds, they can produce more consistent temperatures.

Interviews with market actors revealed there is no consensus on proper sizing, commissioning or installation of variable capacity heat pump systems, particularly ECHPs. Variable capacity systems are more difficult to size than single speed systems. Contractors must consider the estimated heating balance point at both maximum and minimum capacity to ensure a heat pump produces enough heat at cold temperatures while not overproducing heat at mild temperatures and inducing short cycling. Some contractors use manufacturer heating capacity tables to estimate system balance points and determine the optimal system size. Manufacturers provide sizing tools and installation specifications for their equipment, but there is little consistency and many contractors do not use them. Many ECHP systems have automatic testing procedures and may collect performance data, which contractors can use to help commission systems. It is unclear if these tools are intended to achieve efficient operation or simply ensure that systems function properly.

There was widespread interest in additional training opportunities for contractors so they can properly size, install and commission ECHPs to operate efficiently. There was significant concern about making sure systems don’t run at full capacity too often or short cycle during mild temperatures, both of which can cause inefficient operation. Manufacturers and contractors noted ECHPs can be installed without backup heat in many cases. Installing systems without backup resistance heat could potentially result in large demand reductions in the winter and additional energy savings. Some contractors noted they prefer to install the backup heat but not wire it in so that it is available for emergencies.

CLEAResult looked at several potential ways ECHP energy performance might be improved. They found all ECHP systems require a proprietary thermostat to properly control system operations. Nearly all proprietary controls are capable of setting an auxiliary heat lockout temperature, above which the backup heat will not operate. EHCP controls typically have comfort and efficiency modes that dictate how quickly the compressor will transition from low to high speed and may have an impact on energy performance. There may also be an opportunity to improve how ECHP defrost cycles function. Some systems monitor refrigerant pressure to determine when defrost is needed, rather than simply running it on a regular time interval, which may produce additional savings. Some ECHPs also have variable wattage crank case heaters that are only engaged when needed. This improvement could also save significant energy.

Analysis of the total installed cost of 25 ECHPs installed in 2019 resulted in an average cost of $15,790. The assumed installed cost for standard VCHP systems was $14,690. Thus, the incremental cost of an ECHP above the baseline of a standard VCHP system is $1,100.

Power metering study

The two ECHP systems that were metered were both Carrier models that retained an average of 99% of their 47 degree maximum heating capacity down to 17 degrees. These systems both had HSPF ratings of 12. According to SBW, one ECHP was oversized and the other was sized correctly. Both had control settings for auxiliary heat lockout that were in line with best practices, although the best practices were designed for standard capacity systems and do not account for the cold temperature capacity of ECHPs. The six VCHP systems metered were a mix of Carrier, Mitsubishi, Trane and Lennox models that retained an average of 77% of their 47 degree maximum heating capacity down to 17 degrees. They had an average HSPF of 9.9. According to SBW, three VCHPs were oversized, two were undersized and one was sized correctly. Four of these had control settings for auxiliary heat lockout that were in line with best practices, while two did not. One of these had a relatively high auxiliary heat lockout temperature and the other did not have a lockout temperature set, meaning both may employ their backup heat at more mild temperatures when it is not needed.

Metering results showed the ECHP systems performed better than the VCHPs at cold temperatures during the heating season and used less energy per unit of heat produced. This effect was slightly larger east of the Cascades where more hours of heating occurred at temperatures below 30 degrees. However, there was some evidence that at very low temperatures (below 15 degrees), the energy performance of the ECHPs declined and was more similar to that of the VCHP systems. This may be due to a drop in heating capacity at low temperatures that required both the ECHP and VCHP systems to engage their auxiliary resistance heat and run their defrost cycles more frequently. ECHPs saved energy throughout the heating season but savings tapered off during the shoulder season as temperatures became warmer. There was no discernable difference in the energy performance of ECHP and VCHP systems during mild to hot outdoor temperatures that required either minimal space conditioning or cooling. Thus, there was no evidence ECHPs saved energy when they were in standby or cooling modes.

Although results from so few sites are unlikely to be representative, SBW estimated the expected annual energy and demand savings an ECHP compared with a VCHP system. West of the Cascades, energy savings were estimated to be 1,450 kWh per year. East of the Cascades, they were estimated to be as high as 3,350 kWh per year. West of the Cascades, average winter peak demand savings were estimated to be 0.6 kW during the morning peak and 0.4 kW during the evening peak. East of the Cascades, average winter peak demand savings were estimated to be 0.8 kW during both the morning and evening peaks.

Billing analysis

The Recurve utility billing analysis showed overall average electric savings of 1,300 kWh per year (+/- 271) or 6% of baseline electricity usage, in electrically-heated homes installing an EHCP versus a VCHP system from 2015 to 2018. There were 394 ECHP treatment homes analyzed, which had baseline annual electricity usage of 20,391 kWh on average. They were distributed across Western and Central Oregon but concentrated in the Portland metro area. The VCHP comparison group provided a good representation of the baseline electricity usage and geographic distribution of the treatment group. This made it a reasonable point of comparison to homes that installed ECHP systems. The large sample size, relatively good precision and close match between groups give us relatively high confidence in the overall results.

For heating zone 1, during the same time period, average incremental electric savings were 1,239 kWh per year (+/- 285) or 6%. Heating zone 1 results were nearly identical to the overall results because 90% of homes in the treatment group were in heating zone 1. For heating zone 2, average incremental electric savings were 1,425 kWh per year (+/- 1,203) or 7%. These homes were nearly all in Central Oregon. Although a minor difference was observed in savings between heating zones, the existence and magnitude of this difference is very uncertain due to the low sample size of homes in heating zone 2 and the low precision of the estimate.

Heat pump commissioning and advanced controls incentives are intended to improve the performance of heat pumps, but the impact of such measures on ECHPs is unknown. Differences in the uptake of commissioning incentives may represent a true difference in installation and setup practices, or simply a difference in the contractor's familiarity with Energy Trust's program. Commissioned ECHP projects had average electric savings of 1,092 kWh per year (+/- 294) or 5%. Non-commissioned ECHP projects had average electric savings of 1,612 (+/- 769) or 7%. Although the non-commissioned ECHP savings results were notably higher than for commissioned projects, they are based on a much smaller sample size with lower precision. The power of this analysis was further limited by an uneven split in treatment homes, with commissioned projects accounting for 84% of installations. It is unclear what could be driving this difference or whether it will persist with a larger sample of homes.

We also analyzed ECHP electric savings by manufacturer. Carrier systems had average electric savings of 1,519 kWh per year (+/- 545) or 7%. Trane systems had average electric savings of 893 kWh per year (+/- 353) or 5%. For all other heat pump manufacturers, average electric savings were 1,953 kWh per year (+/- 755) or 9%. Trane ECHP models appear to produce significantly lower electric savings than equivalent Carrier and other ECHP manufacturer models. The source of these differences is not clear, and we do not know if they will persist with a larger sample of homes.

The table below summarizes the ECHP analysis results. Results represent incremental kWh per year savings of ECHP systems compared with standard VCHP systems that were installed from 2015 to 2018. Additional details are provided in the Recurve snapshot reports attached to this memo.

Table 1: Summary of ECHP incremental energy savings results, 2015-2018

Fuel Analyzed	Heating Zone	Cx*	Make	N**	Baseline Energy Usage	Average Savings	Absolute Precision†	Percent Savings	Conf. Level
kWh	All	All	All	394	20,391	1,300	271	6%	High
kWh	1	All	All	356	20,448	1,239	285	6%	High
kWh	2	All	All	31	21,314	1,425	1,203	7%	Low
kWh	All	Yes	All	332	20,015	1,092	294	5%	Moderate
kWh	All	No	All	62	22,401	1,612	769	7%	Moderate
kWh	All	All	Carrier	122	21,763	1,519	545	7%	Moderate
kWh	All	All	Trane	212	19,582	893	353	5%	Moderate
kWh	All	All	Other	60	20,905	1,953	755	9%	Moderate

Note: Savings and precision values are based on a comparison between ECHP projects and a matched comparison group of similar VCHP projects.

* "Cx" is short for heat pump commissioning, which includes heat pump projects that received additional incentives for commissioning and/or installation of a heat pump advanced control.

** N is the final treatment group sample size in the analysis. Matched comparison group sample sizes were roughly five times the treatment group sample sizes, in most cases.

† Absolute precision of the mean, or margin of error, at the 90% confidence level.

Answers to Pilot Research Questions

Below, we apply the findings from our research activities to answer each of the original pilot research questions.

- What are the energy savings for extended capacity heat pumps over other variable capacity heat pumps?

  Based on the billing analysis and corroborating findings from the power metering study, it appears overall energy savings for ECHPs are about 1,300 kWh per year above standard VCHP systems. Savings in heating zone 2 may be higher than that, but we need larger sample sizes to make a reasonably precise estimate.

- What are the operational characteristics of these extended capacity heat pumps that can provide additional energy savings *compared to* other variable capacity heat pumps? Do they vary by manufacturer? (e.g. standby usage, aux heat usage, defrost cycles.)

  Based on the limited sample of homes in the power metering study, it appears that ECHPs have improved energy performance during the heating season, especially during cold periods below 35oF. The billing analysis corroborates this finding, as nearly all of the savings were observed during heating season. However, at very cold temperatures, below 15oF, ECHP and VCHP energy performance may converge again, likely due to increased auxiliary heat use. We found no evidence of ECHP savings in standby mode or during the summer cooling season. We still have more to learn about how ECHPs operate, including their use of auxiliary heat, crank case heaters, and defrost cycles. These are all areas where there may be opportunities to further increase ECHP savings. Operational characteristics are likely to vary somewhat between manufacturers and we did see some evidence that Carrier models and smaller ECHP makes had better energy performance than Trane models.

- What are the sizing, commissioning and setup practices for best energy performance while not negatively affecting occupant comfort?

  The industry is still developing guidance and contractors are still trying to determine best practices for sizing, commissioning and setup of variable capacity systems, including ECHPs. Contractors are primarily concerned about proper system operation and comfort but are also unsure how to maximize energy performance. Proper sizing for all variable capacity systems is more complicated than for single speed systems and there is not yet consensus on this topic from manufacturers or contractors. More sophisticated contractors reported using manufacturer capacity tables to select a system with an estimated maximum capacity balance point between 15 and 25 degrees. Contractors also considered the possibility of short cycling at mild temperatures if the minimum capacity balance point was too low.

  From the power metering study, we saw heat pump control auxiliary heat lockout temperatures appeared to be set according to industry best practices, even though these settings may reduce the savings potential of ECHP systems. Manufacturers and contractors noted auxiliary heat is not always required for ECHP systems to maintain comfortable temperatures and some contractors reported installing but not connecting the backup resistance heat in certain cases. This change in practice has the potential to significantly increase ECHP energy and demand savings in Western Oregon. Revisiting the best practice auxiliary heat lockout settings for ECHPs could yield energy and demand savings in homes where backup heat is still required. Other potential equipment improvements were identified with defrost cycles and crank case heaters that could lead to further energy savings.

- How does sizing, commissioning and installation differ between extended capacity and other variable capacity heat pumps?

  At this point, there seem to be few differences in sizing, commissioning and installation practices between ECHP and VCHP systems. Contractors tended to group both types of systems together as variable capacity. According to contractors, all variable capacity systems tended to be sized based on their actual capacity values at different outdoor temperatures. The metering study showed that heat pump controls on ECHP and VCHP systems were generally set up similarly, according to industry best practices.

- What is the incremental cost of extended capacity heat pumps over other variable capacity heat pumps?

  Based on the cost data reported to Energy Trust through incentive applications for 25 ECHP projects installed in 2019, total installed ECHP costs averaged $15,790. Compared with the assumed average standard VCHP installed cost of $14,690, ECHP systems have a roughly $1,100 incremental cost.

Conclusions

ECHPs are a quickly emerging technology at the top end of the variable capacity heat pump market, commonly referred to as cold climate models. ECHPs appear to have improved energy performance at cold outdoor temperatures and offer significant energy savings over standard VCHPs installed in comparable scenarios. The overall incremental electricity savings were 1,300 kWh per year across heating zones. This savings estimate is reasonably precise and was corroborated by direct power metering of a small sample of systems where similar savings levels were observed. In addition, it is likely ECHPs provide a small amount of peak demand savings in the winter above and beyond standard VCHPs. The pilot did not identify any standby mode or cooling savings from ECHPs. The average installed cost of ECHPs was approximately $15,790, indicating a roughly $1,100 incremental cost above standard VCHPs.

Although best practices for sizing, commissioning and installation for variable capacity systems have not yet been established, and there is no consensus among manufacturers or contactors, some recommendations are beginning to emerge. Even in the absence of best practices, ECHP systems appear to have good energy performance and significantly outperformed their standard capacity counterparts. Whether this difference is due to ECHPs having outstanding low temperature performance or VCHPs having poor performance is not yet known, but it is worth further investigation. Additional research on optimal sizing for variable capacity systems and commissioning and installation practices may be needed before market actors agree upon a set of best practices.

It may be possible to further improve the energy performance of ECHP systems over time by encouraging contractors to disconnect auxiliary heat in some cases, developing commissioning and setup best practices, providing guidance on proper system sizing and working with manufacturers to develop energy saving features like improved crank case heaters and defrost cycles. If extended capacity technology is successful and Energy Trust can increase its penetration in the top end of the heat pump market, then the technology may work its way into less expensive heat pump tiers. Extended capacity technology has the potential for broad market transformation, improving the energy performance of heat pumps across the board.

Next Steps

Energy Trust is currently wrapping up its coordinated research activities on EHCP systems and winding down the pilot. Energy Trust will adopt the overall incremental electricity savings of 1,300 kWh per year as the deemed savings value across heating zones above a baseline of a standard VCHP system. We will assume an average incremental cost of $1,100. It is expected that cost-effectiveness screening based on these values will result in a cost-effective new ECHP measure. The Residential program plans to develop a new measure and begin supporting ECHP systems more broadly with a market-based incentive by July 2020. There will also be some work involved in developing and maintaining an updated qualified products list that provides options for contractors while maximizing energy savings. It is expected that volume will increase as contractors who would have otherwise installed standard VCHPs convince customers to switch to ECHPs. Energy Trust may place additional requirements on these incentives, including stipulating new heat pump control lockout settings that are more closely aligned with ECHP capabilities. Energy Trust will also consider guidance on disconnecting auxiliary resistance heating coils when feasible.

While research is needed on optimal sizing, commissioning and installation practices and their impacts on energy performance, this work is not currently a high priority for Energy Trust and may be taken on by other entities. The ECHP market is relatively new and it is expected that best practices will naturally emerge as manufacturers, distributors and installers all gain experience with the technology. Once best practices for maximizing system efficiency do emerge, Energy Trust will provide recommendations to contractors and may adjust measure requirements accordingly. While Energy Trust has no plans to conduct a second phase coordinated research project on ECHPs at this point, we will conduct research and evaluation in the coming years as needed. We believe there are improvements manufacturers could make to achieve additional EHCP energy savings, including improvements to heat pump defrost cycles and crank case heaters. Energy Trust will encourage manufacturers to bring these improvements to market.

After Energy Trust's new, expanded ECHP offer has been in the market for a year or two, and we see significant project volumes, we will conduct additional evaluation activities. The most basic of these will be monitoring the costs of ECHP and VCHP equipment. We will rerun the billing analysis of ECHP versus VCHP systems using Recurve to obtain more precise savings estimates and monitor potential changes in savings. At that point, we may investigate savings against other baseline equipment types. If enough installs have occurred in heating zone 2, then we will be able to obtain more precise estimates of energy savings by heating zone as well. We may also investigate the impact of duct location and condition on savings, which has been hypothesized to be a factor. And we may conduct a short survey with program participants installing ECHP and VCHP systems to research customer satisfaction and comfort, control settings and home and occupancy characteristics that could impact performance.

In addition to this ongoing evaluation research, Energy Trust will identify applications for lower cost ECHP equipment, such as in manufactured homes. Energy Trust will investigate savings in different scenarios, including whether ECHPs may have viable, cost-effective applications with different baselines. For instance, a future ECHP measure may be an alternative to single speed heat pumps, have a full market baseline or work as a retrofit in homes with electric forced air furnaces. Energy Trust will also collaborate with regional entities to expand the ECHP market, accelerate efficiency improvements and market transformation across the heat pump market and support the development of best practices.

List of Attachments

1. CLEAResult. 2020. Extended Capacity Heat Pump Pilot Report: Learnings from the Field.
2. Rodriguez-Anderson S and Bertini D. 2019. Memo for Add-On Metering at BPA Heat Pump Study Sites. SBW Consulting.
3. Recurve Analytics. 2020. Extended Capacity Heat Pump Impact Analysis Snapshot Reports, 2015-2018.
    a. Overall electricity savings
    b. Heating zone 1 electricity savings
    c. Heating zone 2 electricity savings
    d. Commissioned heat pump electricity savings
    e. Non-commissioned heat pump electricity savings
    f. Carrier heat pump electricity savings
    g. Trane heat pump electricity savings
    h. Other heat pump manufacturer electricity savings

Attachment 1:

CLEAResult. 2020. Extended Capacity Heat Pump Pilot Report: Learnings from the Field.

CLEAResult®

Extended Capacity Heat Pump Pilot Report-
Learnings from the Field

Introduction and Pilot Description

Over the past several years, installations of variable capacity heat pumps (VCHPs) and extended capacity heat pumps (ECHPs) in Energy Trust of Oregon territory have become increasingly prevalent. It is expected that systems with extended capacities at lower outdoor temperatures (i.e., ECHPs) will produce additional savings, however there is a lack of research quantifying the expected savings from ECHPs in Oregon's climate. Additionally, commissioning best practices such as sizing, installation and control of electric strip heat, appropriate airflow, and optimal thermostat settings are relatively unknown for these systems.

In 2019, Energy Trust launched an Extended Capacity Heat Pump Pilot (hereafter referred to as "the Pilot") in order to better understand whether there are additional savings for ECHP systems compared to standard VCHPs. Additionally, the Pilot was also intended to learn more about the operational characteristics of ECHPs, as well as sizing, commissioning, and installation best practices for these systems. Sizing in the context of heat pumps refers to finding the best match between the heating and cooling loads of the house and the capacity of a particular unit, whereas commissioning refers to a set of startup procedures that ensures that a system is correctly installed and operating at its engineered performance levels. Installation best practices for heat pumps refers to procedures such as wiring and piping, sealing penetrations, brazing, charging system refrigerant, evacuating line sets, etc.

Energy savings for ECHPs compared to VCHPs were estimated using billing analysis and through a small power metering study, and are presented in separate reports. This report presents learnings and insights gained through conversations in the field with heat pump manufacturers, distributors and installers during the Pilot period.

Background

The Pilot was developed and executed by Energy Trust's Residential program, which is implemented by CLEAResult. CLEAResult program staff, referred to here as program staff, were responsible for contract management, budget and delivery oversight, pilot design assistance, quality assurance, and coordination of forms & systems tracking.

Measure Approval Document # 227 was published for use in the Extended Capacity Heat Pump Pilot. The MAD contained an estimated average incremental savings of 930 kWh per unit for ECHP systems, compared to standard VCHPs. Savings were estimated using preliminary billing analysis results from ECHP systems installed in PGE territory in 2014 and 2015. A $400 Pilot incentive, in addition to Energy Trust's standard $700 heat pump incentive (or $1,000 for Savings within Reach customers), was provided for qualifying ECHP systems submitted through the Pilot. Energy Trust's Heat Pump Advanced Controls incentive ($250) could also be applied to projects submitted through the pilot, resulting in $1,350 of potential incentives for qualifying systems ($1,650 for Savings within Reach customers)

A list of qualifying ECHP models (qualified products list, or QPL) was developed by program staff using system capacity information gathered from AHRI (Air Conditioning, Heating and Refrigeration Institute). Systems that have at least 85% of their rated capacity at 17 degrees Fahrenheit were considered ECHP systems for the purposes of this pilot.

The Pilot was approved to run through June 2020, and initially expected to see approximately 20-40 installations submitted[2].

As of December 31st, 2019, nine ECHP systems have been submitted and recorded for the Pilot. 16 additional projects were recorded in 2019 for Energy Trust's standard heat pump incentive for systems that qualify as ECHPs but were not submitted through the Pilot.

Outreach

Installer Recruitment

Recruitment for participation in the pilot was initially limited to a select group of installers in order to limit the potential for oversubscription for the pilot, due to the relatively high incentive amount available for qualifying projects. Installers that had previously submitted applications for ECHPs through Energy Trust's standard heat pump upgrade incentive were invited to a series of pilot kick-off webinars starting in May 2019. Installers from 19 different companies attended these initial Pilot outreach events.

Unfortunately, there were no ECHP Pilot incentive applications submitted until late-August, nearly 3 months into the pilot. This suggested the need for a different recruitment approach. At this point, a conversation with Mar-Hy distributors led to the suggestion that distributor-hosted "Lunch and Learn" events, where program staff could present the details of the pilot in a familiar environment, may be a more effective way to engage and recruit HVAC installers.

Mar-Hy distributors hosted the first Lunch and Learn event on October 15th, 2019, which saw installers from 7 different companies in attendance. Following this event, four additional distributor-hosted events were held in the Portland and Bend areas, with over 17 companies attending in total. These distributor-hosted events were very successful in recruiting additional installers for participation in the pilot, and proved useful in gathering program design input and technical information about ECHPs. However, despite the relative success of the Lunch and Learn events, Pilot project volume remains below the expected number of installations, which has limited the ability of program staff to draw robust insights regarding sizing, installation and commissioning best practices for ECHPs.

Additionally, participating installers initially agreed to allow program staff to attend or audit the first 3 ECHP installations. The goal of attending and observing installations of ECHPs was to better understand how installers were performing the following aspects of heat pump sizing, commissioning and installation;

- Compressor and auxiliary heat lockout settings
- Use of auxiliary heat during defrost
- Auxiliary heat staging and power draw
- Estimate heat loss of house
- Dehumidification settings
- Fan speed settings
- Measured CFM during test mode
- Duct static pressure issues
- Obtaining access to remote thermostat monitoring

However, it proved difficult to schedule installation visits with installers during the 2019 program year, and only 3 site visits have been successfully completed to date.

[2] The expected number of ECHP pilot applications described in MAD # 227 is based on the historical volume of standard heat pump incentive applications for systems that meet the definition of ECHPs

Manufacturer Engagement

CLEAResult program staff also met with the seven major heat pump manufacturers during the initial pilot period to gain further insights about the operation characteristics and sizing/ commissioning of extended capacity heat pump equipment. Program staff met with the following heat pump manufacturers;

- Carrier
- Rheem
- Mitsubishi
- Daikin
- Trane
- Johnson Controls (York/Coleman)
- Lennox
- York

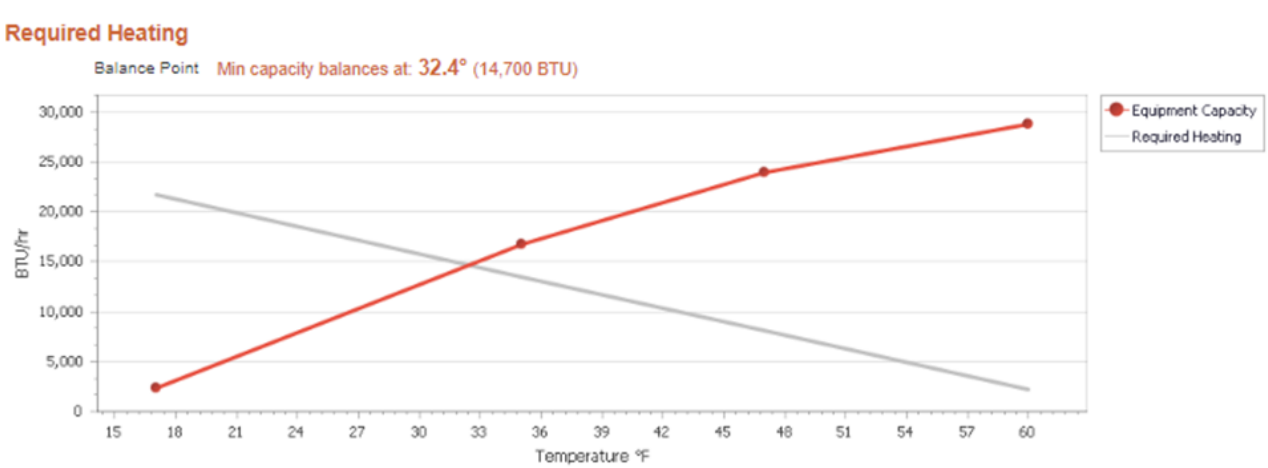
Sizing, Commissioning, and Installation Findings

Learnings from the field confirmed our understanding that VCHPs (including ECHPs) are materially different from single speed heat pumps in that they are not standardized pieces of equipment and cannot be sized, commissioned, and controlled across all brands and models using universal "one-size-fits-all" specifications. This represents a significant departure from sizing and commissioning for single speed heat pumps, where systems are more standardized and similar and can better utilize standardized sizing procedures.

In the Northwest, energy efficiency programs (in particular PTCS) require a balance point of 30F or below. The balance point is best described as the lowest outdoor temperature at which the compressor alone, without the aid of electric resistance auxiliary heat, can meet the load of the house. Historically, programs in the Northwest have ignored cooling loads when selecting system balance points due to typical heating loads being higher than cooling loads. Installers also mentioned that for new construction applications this may not be the case and it is possible that heating loads may end up being smaller than cooling loads in certain situations. ACCA's Manual S is generally regarded as an acceptable method of selecting the best fit for single and dual speed systems, however it does not yet provide clear guidance about sizing for variable speed equipment, include ECHPs.

The figure below illustrates an example single speed Heat Pump with a balance point of roughly 32F. The point at which the lines cross indicates the temperature at which the equipment's capacity is equal to the expected heating load.

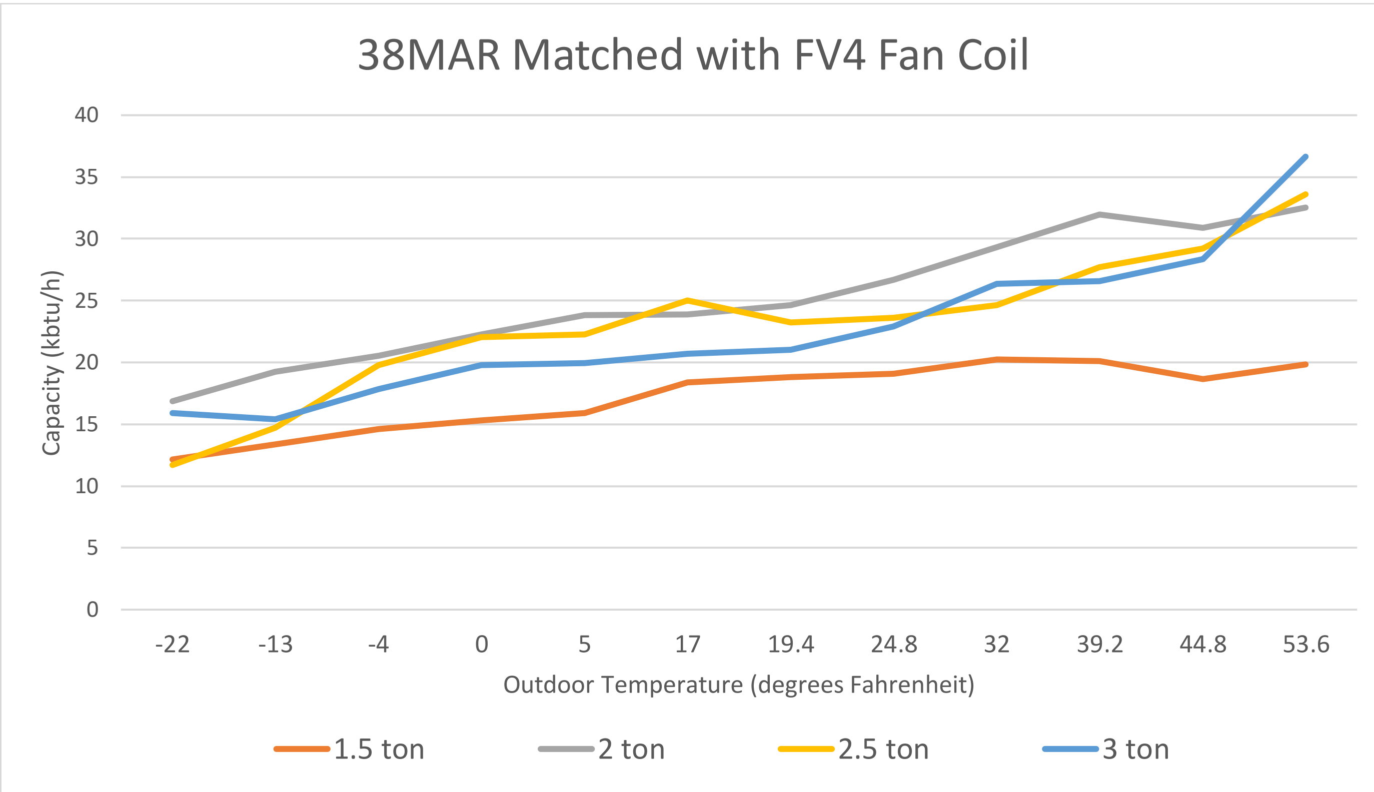
Figure 1: Single Speed Heat Pump Balance Point Example



Source: https://hvac.betterbuiltnw.com/Account/Lo

The figure above is typical for single speed systems, where capacity is fairly linear with respect to outdoor temperature. A three-ton single speed systems will always have a greater capacity than a two-ton system, across all temperature bins. However, for variable speed equipment (including ECHPs) the nominal capacity rating cannot be interpreted in the same straightforward manner. It is instead best to consider the nominal capacity rating as the "marketed capacity". We've seen that within a single model line, actual capacities at various outdoor temperatures can differ significantly. The chart below demonstrates this situation for a Carrier model 38MAR matched with FV4 air handler. As the graphs indicates, the capacity at 17F for the two-ton and 2.5-ton systems is greater than the capacity of the 3-ton system. This would not be the case with single-speed heat pumps.

Figure 2: System Capacity vs Outdoor Temperature- Carrier



source: https://www.carrier.com/commercial/en/us/products/ductless-systems/38mar/

With regards to auxiliary heat, both manufacturers and installers stated that in many cases ECHP systems can be installed without any auxiliary heat source. The conditions where it was possible to install systems without an auxiliary heat source are determined by the winter design temperature at the site, and the heat rate loss of the house. Generally, conditions that allowed for installation of ECHPs without auxiliary heat were found West of the Cascades where weather is more mild.

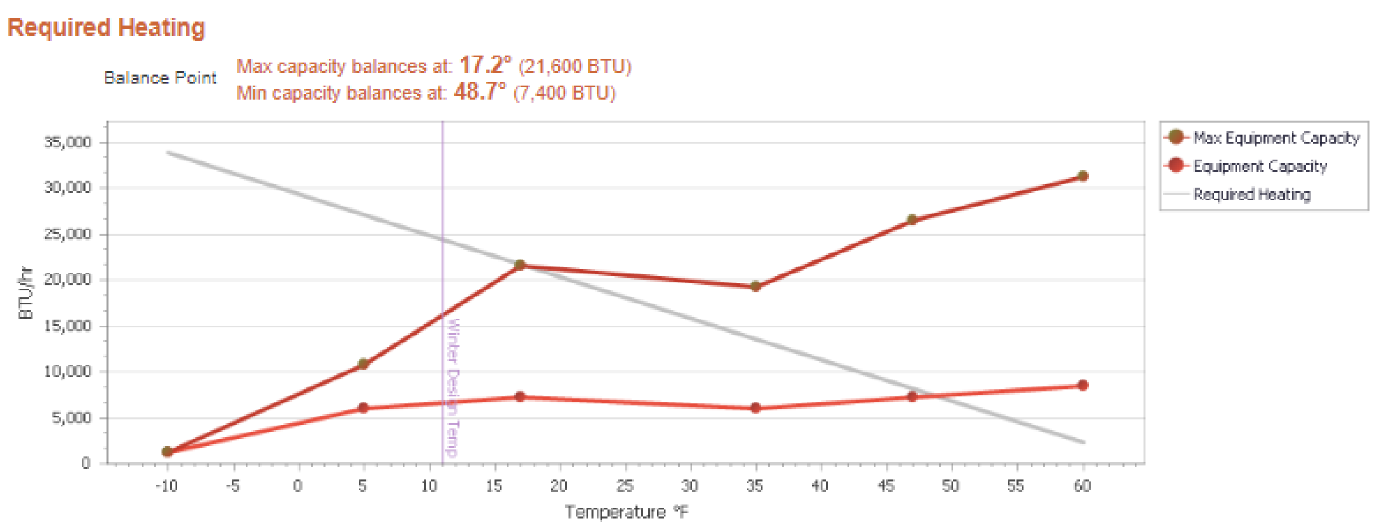
Some installers indicated that they do install the auxiliary heat, but do not wire them in. The auxiliary heat is then utilized when the compressor needs repair or replacement, giving the installer a fast and easy way to provide emergency heat to the house. Installing ECHP units without aux heat may significantly lower a house's peak demand and also significantly improve the heat pump system's energy performance. The average house in the Northwest has 10 to 15 kW of auxiliary heat. Removing this back-up heat source could significantly lower a utility's future peak demand if these units were to gain significant market share, in addition to saving a significant amount of energy.

Installer Findings

As described above, proper sizing of variable speed equipment and in particular ECHPs is more complicated than single speed systems and can also vary by system manufacturer. Discussions with installers similarly revealed very little uniformity in the sizing practices currently being used in the market. Installers that appeared to have a more sophisticated understanding of the equipment stated that they do not rely on the AHRI ratings[3], but instead utilize the manufacturer's capacity tables or charts to better understand a system's capabilities over a wider temperature range. The most common approach appears to use the maximum capacity of the system to determine the balance point. Using this method, installers reported aiming for a maximum capacity balance point of between 15F and 25F.

Other installers reported using their experience to select a system that, in their judgment, is the best fit utilizing both the maximum and minimum capacities. The graph below shows a system with a balance point of 17F using the maximum capacity and a balance point of 49F at low capacity.

Figure 3: Variable Speed Heat Pump Balance Point Example



Source: [*https://hvac.betterbuiltnw.com/Account/Lo*](https://hvac.betterbuiltnw.com/Account/Lo)

Another sizing consideration often mentioned by installers during discussions in the field is the issue of unintended 'short-cycling'. If the minimum capacity balance is too low, say 35 F, there is a concern that the system will then cycle on and off too frequently, compromising comfort and perhaps equipment longevity. Evidence from the Next Step Homes program has indicated that when variable speed single head ductless systems have low minimum capacity balance points, the short-cycling causes the systems to react in ways that cause higher energy consumption. It is unknown if this effect is universal across all inverter driven, variable speed heat pumps, or if this issue was unique to ductless systems.

One installer reported sizing systems based on the ability to operate at 60% of maximum capacity at a 40F outside temperature. This installer reported that their sizing strategy is intended to keep the system's noise level low and maximize homeowner comfort.

Installers generally felt that commissioning for single speed heat pumps is much more straightforward than for variable speed equipment (including ECHPs). Installers described using an equipment 'test mode' that most ECHPs systems have during the commissioning process. The test mode runs the system at its maximum capacity (or close to it depending on outside temperatures), which allows the installer to test the performance at maximum capacity and compares that to published factory specifications. Installers stated that without this test mode it would be difficult to

[3] AHRI publishes data on the capacity of every combination of outdoor and indoor unit at 17 and 47 degrees.

determine what the correct commissioning values for a system should be since the operating speed would be unknown.

Conversations with installers also revealed that some ECHP systems have on-board temperature and pressure sensors that allow viewing of critical performance data, which helps guide the installer through the commissioning process (e.g., informs whether to add or subtract refrigerant). Installers expressed general satisfaction with these automated guides, however it is not known the extent to which those guides focus on achieving energy efficiency versus sufficient system capacity.

According to installers, VSHPs (including ECHPs) are typically sold based on their comfort benefits. Installers reported that the higher run times associated with variable capacity equipment produces higher levels of comfort than traditional heat pumps, by delivering an even temperature and not turning on and off frequently. Additionally, installers also asserted that they have learned to not sell heat pump upgrades on potential energy savings, since operation costs are a small part of the overall system cost, and also to protect themselves from the risk of overstating the benefits the customer will actually realize. While energy savings might be realized on average across all customers, guaranteeing that a given site will realize those energy savings is extremely risky from the perspective of installers.

Manufacturer and Distributor Findings

Conversations with manufacturers and distributors led to some general recommendations regarding sizing, commissioning and installation best practices. However, these market actors also felt that additional research and guidance on best practices is also needed for all three areas.

Manufacturers stated that sizing tools are available, however few installers use them. Manufacturers suggested following their published installation specifications to assure the best combination of proper function, comfort, and efficiency. However, there does not seem to be any industry consensus around proper sizing of ECHPs.

Discussions with manufacturers also indicated that further research about each manufacturer's algorithms used for airflow and capacity is warranted, to better inform potential differences in proper sizing, commissioning and installation across different manufacturers. However, manufacturers also warned that it will likely not be possible to fully understand exactly how the various control algorithms work, which will limit the ability to develop detailed best practice specifications for ECHP systems.

All manufacturers felt that duct size should be reviewed before sizing a system, since restrictive ducts may result in too little airflow for the capacity of the system, thereby increasing the fan motor's current draw. However, manufacturers also acknowledged that the minimal time spent at high capacity, for properly sized systems, should limit the negative impacts of restrictive ducts. Furthermore, manufacturers also suggested that ECHP systems may be a good solution for homes where ducts are small and/or restrictive, since a lower capacity ECHP system may be sufficient where a higher capacity standard VCHP would have been required instead.

In homes where ducts are located in unconditioned spaces (i.,e., crawlspaces and attics), there will be greater conductive and air leakage losses when compared to homes with single speed heat pumps. This is due to two reasons; 1) the system runs longer more hours than a single speed system, by design, and 2) when operating at a lower capacity, airflow velocity is also decreased, which conductive losses per CFM of delivered air.

Overall Findings

All respondents agreed that it is very important to size and commission systems such that compressor is not being run at full capacity for extended periods of time. Unfortunately, there does not seem to be any consensus across market actors about how to best achieve that.

Both installers and manufacturers agreed that an additional level of training, beyond standard heat pump installation training, is needed for installers to be able to properly install ECHPs. However, respondents stated that these trainings typically focus on mitigating comfort and noise issues and do not place much emphasis on energy savings or efficiency. These trainings are typically provided by the manufacturers. In some cases, we heard that additional training is required before installers can even purchase ECHP systems, since it is also in the manufacturer's best interests to ensure the units are installed properly.

Conversations with manufacturers and installers also provided insights about unintended short-cycling of systems at mild temperatures. ECHP systems are inverter-driven, meaning they can operate across a range of speeds/ capacities rather than just a single speed. However, at moderate outdoor temperatures (50-60 degrees) the heat loss of the house might be significantly less than the lowest heating capacity of the unit, depending on how the system was sized. In that case, the heat pump will cycle on and off more than intended, reducing the system's overall efficiency and negating the primary benefits of having an inverter-driven system. If a large percentage of the heating takes place above 50 degrees, the minimum capacity of the equipment should be a primary consideration when sizing the system in order to avoid short-cycling and maximize system performance. Market actors felt that short-cycling issues are more prevalent for systems that can be sized with very low balance points (0-5 degrees). Additionally, there is generally consensus across market actors that unintended short-cycling may also reduce the life of the equipment.

When manufactures and installers were asked about installation best practices for achieving the greatest energy savings, they responded that installation guidance currently does not include suggestions or practices for maximizing savings. Respondents did have some suggestions, though, about how energy savings might be achieved.

- Most systems offer the ability to stage auxiliary heat in increments of roughly 5kW
- Some *Defrost* settings may provide energy savings, but could also lead to the defrost not functioning under cold and/or humid conditions.

Operational Characteristics

Most heat pump manufacturers interviewed had similar definitions of what is considered an "Extended Capacity Heat Pump", however Mitsubishi provided a slightly more restrictive definition than other manufacturers;

- Maintain similar 47-degree capacity down to 25 degrees or less **(Carrier, Rheem/RUUD, York)**
- Maintain similar 47-degree capacity down to 5 degrees or less **(Mitsubishi)**

Conversations with manufacturers also revealed that the term "Extended Capacity Heat Pump" is not used industry wide, and that most manufacturers refer to these types of systems as "Cold Climate Heat Pumps". While these definitions do not align exactly with the definition of an Extended Capacity Heat Pump employed by Energy Trust[4], they are all meant to identify systems that are likely to be able to meet the entire heating load of a home (or the majority of its load) at outdoor temperatures of around 20F (or 5F for Mitsubishi).

Controller settings- All qualifying systems require a proprietary controller (i.e., thermostat). Historically, controllers/thermostats for single speed Heat Pumps have essentially functioned as an on/off switch. However, VCHPs and ECHPs are now often programmed and settings are adjusted through the controller or thermostat. Across the various controllers that were observed through the pilot, there were some common characteristics that emerged;

Controlling Auxiliary Heat- With the exception of the Mitsubishi controllers, all allow for the lockout of auxiliary heat at selected outdoor temperatures. Mitsubishi controllers require an external lock-out for strip heat, however they stated that their control algorithm is designed such that auxiliary heat is not needed. Program staff were not able to assess typical default temperature settings for auxiliary heat lock-out for these systems, and further investigation is needed.

Comfort Vs Efficiency mode- Controllers often have a "switch" between comfort and efficiency. When comfort mode is selected, the transition time between compressor stages is allowed to decrease, increasing the rate the house warms up or cools down. In other words, the heat pump is allowed to move from the lower capacity stages to higher capacity stages more quickly when comfort mode is selected. However, all market actors that mentioned this controller setting reported that if an auxiliary heat lockout temperature has

[4] Systems that maintain at least 85% of their nominal rated capacity at 17 degrees F.

been programmed, the system will keep the auxiliary heat off until the lockout temperature is reached, regardless of whether comfort or efficiency is selected.

Compressor Limiting- This feature allows the technician to limit the maximum RPM of the compressor. Adjusting this setting impacts the system's maximum capacity and is used primarily to limit the noise level of the system. Installers indicated that they do not use this setting unless there is a potential noise issue (e.g., sound ordinances).

Defrost controls- Most of the qualifying systems have various options on how to accomplish defrost. These range from the more traditional timed defrost method to a system that looks at refrigerant pressure to optimize the defrost cycle. The installers interviewed leave the setting in default mode, which is the optimized option. This may be a source of savings for these systems over more traditional heat pumps

Crank Case heaters- Traditional heat pump systems control the crankcase heater with resistance heat, typically of between 40 and 90 watts. This heater is typically engaged at temperatures below 70F. At least one of the qualifying systems (York) engages the crank case heater at 55F and deploys a variable wattage heater that varies its wattage between 5 watts and 55 watts. This control strategy may provide savings on the order of 200-300 kWh per year. It is not known if other manufactures have modulating crank case heaters as well.

Equipment Costs

As of January 27, 2020, nine applications for the ECHP Pilot had been received and recognized in Energy Trust's tracking system. However, 16 additional incentive applications were received in 2019 for systems meeting the ECHP specification, however those projects did not apply for the additional $400 Pilot incentive and received only Energy Trust's standard heat pump incentive amount.

Total cost and system characteristics for the nine projects submitted through the ECHP Pilot are shown in the following table;

Table 1: System Characteristics for ECHP Pilot Projects

Installer Company	Total Cost	Cost per Ton	System Location	System Manufacturer	Model	HSPF	EER	SEER	Nominal Capacity
Home Heating & Cooling, Inc.	$20,601	$10,301	Bend	York	YZV24B21	10.50	14.25	21.00	2.0
Specialty Heating & Cooling INC	$17,925	$5,975	Lake Oswego	Carrier	25VNA036A003	11.50	13.50	20.00	3.0
Specialty Heating & Cooling INC	$20,278	$6,759	Lake Oswego	Carrier	25VNA036A	11.50	13.50	20.00	3.0
Hendrix Heating & Air Conditioning LTD	$9,974	$3,990	Corvallis	Mitsubishi	PUZ-HA30NHA5	9.70	12.50	17.00	2.5
Hendrix Heating & Air Conditioning LTD	$12,645	$4,215	Corvallis	Mitsubishi	PUZ-HA36NHA5	11.00	12.50	17.80	3.0
Bend Heating & Sheet Metal, Inc.	$25,828	$6,457	Bend	Carrier	25VNA048A003	11.00	12.50	17.50	4.0
Mill Creek Heating	$14,684	$3,671	Salem	TRANE	4TWV8048A1	10.00	12.50	18.00	4.0
Hendrix Heating & Air Conditioning LTD	$18,903	$7,561	Corvallis	Mitsubishi	PUZ-HA30NHA5	9.70	12.50	17.00	2.5
Hendrix Heating & Air Conditioning LTD	$11,695	$3,898	Philomath	Mitsubishi	PUZ-HA36NHA5	11.00	12.50	17.80	3.0
All Installers Average	**$16,948**	**$5,870**				**10.66**	**12.92**	**18.46**	**3.0**

The systems shown in Table 1 above were installed between June 19th and December 30th, 2020 and showed a large range in pricing practices. System costs, including installation costs, ranged from a low of $3,600 per ton to a high of over $10,000 per ton, with an average rated capacity of 3 tons. Carrier units were the most prevalent, with four installed units, and showed an average cost of approximately $6,400 per ton.

Average total system cost for projects submitted through the Pilot was approximately $17,000. However, due to the relatively small number of systems submitted for the Pilot, it is difficult to draw conclusions or identify patterns with regard to how cost relates to system location, installer, or manufacturer. However, it does appear from the limited data available Carrier units have surprising consistency in their cost per ton, even when multiple installers are considered.

The following table presents summary statistics for the 16 systems additional submitted through Energy Trust's standard heat pump incentive, which would have qualified for the ECHP pilot but were not submitted for under the Pilot incentive application.

Table 2: System Characteristics for non-Pilot ECHP Projects

Installer Company	Total Cost	System Location	System Manufacturer	Model	HSPF	EER	SEER
AccuAir Inc.	**$12,814**	Bend	American Standard	4A6V8048A1	10	12.5	18
Melton Heating & Air Conditioning, Inc	**$18,436**	Sublimity	Trane	4twv8037a1	10	13	18
Ben's Heating & Air Conditioning	**$18,500**	Portland	Bryant	280ANV036*0**A*	11.5	13.5	20
Anctil Heating & Cooling INC	**$19,405**	Hillsboro	CARRIER	25VNA060A003	12	12	18
Central Oregon Heating & Cooling, Inc.	**$15,480**	Bend	Carrier	25VNA024A0030040	10.5	14.5	19
Bull Mountain Heating AC & Insulation	**$14,780**	Beaverton	Mitsubishi	PUZ-HA36NHA5	11	12.5	17.8
First Call Heating & Cooling	**$12,694**	Wood Village	Trane	4TWV0036A1000BA	10	13	18
Climate Control, Inc.	**$3,045**	Portland	Carrier	25VNA024A003	10.5	14.5	19
Ben's Heating & Air Conditioning	**$21,146**	West Linn	Bryant	280ANV024	10.5	14.5	19
All Phase Remodeling Inc.	**$11,334**	Cornelius	Ruud	UP2024AJVCA	11	15	21
Doug Woodward Heating Inc	**$12,253**	Salem	Coleman	HC20B2421S	10.5	14	21
Heating Solutions LLC	**$10,980**	Astoria	Mitsubishi	PUZ-HA36NHA5	11	12.5	17
Deluxe Heating & Cooling	**$15,628**	Portland	Mitsubishi	PUZ-HA42NKA	9.3	10.1	18
Sun Glow Inc	**$18,897**	Welches	CARRIER	25VNA036A003	10.5	12	18.5
First Call Heating & Cooling	**$11,135**	Portland	Trane	4TWV8036A1000B	10	13	18
A&E Heating and Air Inc	**$17,633**	Hood River	TRANE	4TWV8048A1000B	10	12.5	18
All Installers Average	**$14,635**				**10.37**	**13.12**	**18.83**

As shown in Tables 1 and 2 above the average total system cost for ECHPs, including both pilot and non-pilot installations, is $15,791. Average HSPF, SEER, and EER ratings were very similar for the two groups of projects[5].

The average system cost for projects submitted through Energy Trust's standard heat pump incentive is $14,692, which suggests an incremental cost of $1,100 for ECHPs compared to standard VSHPs. This is close to the assumed incremental cost of $769 that is shown in the ECHPs Pilot MAD #227[6].

Next Steps & Future Research Questions

Learnings from the field during the first phase of the pilot provided some broadly applicable learnings with regards to sizing, commissioning and installation best practices for ECHPs. However, it has also become clear that there are still several areas requiring further study in order to effectively inform development of best-practice guidelines that apply across all ECHP systems.

One important area of study that was identified during the initial pilot period is the added duct losses of VCHP systems over conventional heat pump systems. A potential approach to better understanding this issue would be to compare savings in homes with ducts outside the conditioned space to homes with ducts inside the conditioned space. Collecting information about duct location for systems incentivized through the pilot may allow for future analysis of how savings relate to duct location.

Extended capacity heat pumps have the ability to heat well insulated structures under design conditions without the use of auxiliary heat. Some installers are currently installing ECHPs without auxiliary heat. Additional research on the specific installation scenarios where it is possible to eliminate the need for auxiliary heat with ECHPs, and thereby achieve deeper energy and demand savings, would be valuable to inform future program activities.

Another possible area of future research is regarding the ECHP systems in manufactured homes. A certain system combination meeting ECHP requirements was identified during the pilot period that when paired with an air handler, appears to be well suited to manufactured homes due to its small physical size and relatively low cost. This system also appears to provide sufficient capacity at design temperatures west of the Cascades to operate without auxiliary heat. This provides an additional opportunity to study whether ECHPs systems can eliminate the need for auxiliary heat in certain situations.

The pilot was not able to determine typical default lock-out temperature settings for ECHPs. Program staff suggested that this information should be collected in future phases of the pilot.

Thermostat data sharing for ECHP systems, facilitated by the installer, was an initial goal of the pilot. It became clear early in the pilot that few contractors were willing and able to provide access to the thermostat data from the system. Additional exploration of alternate options for accessing thermostat data would be beneficial.

Program staff also feel that on-site monitoring of more ECHP systems would likely yield valuable insights. Primary areas of interest for future on-site monitoring include a better understanding of savings potential from the defrost and crank case heater strategies these systems use, and also sizing best practices to avoid short-cycling during mild weather in homes with low balance points.

Lastly, conversations in the field with market actors indicated that best practices tend to emerge only after a technology has been available in the market for a few years (or even decades). This suggests that it may still be too early to definitively establish best practices for ECHPs, and that continued monitoring of evolving market trends and installation practices is warranted.

[6] Cost data for MAD #227 is based on Energy Trust Heat Pump projects from 2014-2015

Attachment 2:

Rodriguez-Anderson S and Bertini D. 2019. Memo for Add-On Metering at BPA Heat Pump Study Sites. SBW Consulting.

ENERGY • WATER • EFFICIENCY

Memorandum

FROM: Santiago Rodriguez-Anderson and Dan Bertini, SBW Consulting

TO: Dan Rubado, Energy Trust of Oregon

DATE: November 15, 2019

RE: Memo for Add-On Metering at BPA Heat Pump Study Sites

Executive Summary

Energy Trust contracted with SBW Consulting for a comparison of extended capacity heat pumps to standard variable capacity heat pumps in a field study in the Pacific northwest. The cohort included two extended capacity units, and six variable capacity units. The SBW team audited each site's HVAC equipment, thermostat settings, and building envelope. While auditing the homes, the team installed power metering and temperature sensors. The data loggers collected data from January to July 2019. The team collected the data and created load-normalized energy use models for each type of heat pump.

The modeled results show an annual savings of 1,107 kWh/Rated Ton[1] and 434 kWh/Rated Tons in Redmond, OR and Portland, OR respectively. The savings occur during heating season, and there is negligible difference between the heat pumps in standby or cooling mode. The team's models show peak winter morning savings in Pacific Power and PGE service territories as 0.8 kW/Ton and 0.55 kW/Ton respectively. Extended capacity heat pumps are intended to maintain 85% of their nameplate capacity at 17°F. This study shows that increased heat pump operation at lower temperatures translates to savings in heating mode. While the results are encouraging, only one of the extended capacity heat pumps operated in temperatures below 17°F during the study. The results are neither statistically significant nor conclusive.

1. Background

SBW Consulting conducted a field study of recently installed central ducted air source heat pump systems in homes across the Pacific Northwest for the Bonneville Power Administration (BPA). The goals of the BPA study were to better understand the heat pump market, the equipment they are replacing, and the current installation practices, particularly controls, commissioning, and sizing, regardless of program participation. These data gaps were identified in a previous BPA study and are key components in updating the Regional Technical Forum's (RTF) suite of heat pump measures. Study participants were recruited using a random sample of recently installed heat pumps across the Northwest region, primarily in Oregon and Washington, identified from permit data. The sample was stratified based on geographic location, so that an equal number of heat pumps were included from West and East of the Cascades.

[1] The models were created based on heating demand, and the final results were normalized to weighted heat pump heating capacities.

Energy Trust learned of the BPA heat pump study while developing a study that aimed to measure the electricity consumption of two types of high efficiency heat pumps over time when installed in Oregon homes and subjected to various weather conditions. The two types of heat pumps of interest are standard variable capacity and "extended capacity". The latter are variable capacity air source heat pumps designed to maintain most of their heating capacity at very low outside air temperatures. This dramatically reduces or eliminates the need for these heat pumps to engage less energy efficient back-up electric resistance heating during cold spells. Lab testing and preliminary analysis performed by CLEAResult have shown that these extended capacity units may save a significant amount of energy, especially during very cold periods, beyond other high efficiency, variable capacity heat pumps (VCHP). Energy Trust's working definition of extended capacity heat pumps (ECHP) is: variable capacity central heat pumps that maintain at least 85% percent of their rated heating capacity at 17°F compared to their rated capacity at 47°F. Most of the standard VCHPs are comparable, top tier, high efficiency heat pumps with inverter driven compressors and rating of 10.0 HSPF or greater, that do not maintain their heating capacity at low temperatures.

Energy Trust wished to collect information to better understand the performance of ECHPs and their energy and demand savings, compared to the base case variable capacity units. They approached SBW to explore adding circuit-level power monitoring at sites in Oregon being visited for the BPA study which have standard variable or extended capacity heat pumps. Results of this study may be used to inform a new heat pump measure and incentive so that Energy Trust can support this technology.

The metering was left in place January through July 2019. This memo documents the methodology and results of this add-on metering study.

The primary research goals of the add-on metering study were to:

- Learn about the energy performance of ECHP systems, especially during the coldest days and during "standby" periods, when there are no calls for heating or cooling
- Establish heating, cooling, standby, and total annual electricity savings estimates for ECHPs, compared to base case variable capacity units, for a typical weather year
- Establish electricity demand savings, especially during the coldest days, and a savings shape compared to base case variable capacity units
- What are the operational characteristics of these extended capacity variable capacity heat pumps that can provide additional energy savings compared to standard variable capacity heat pumps, and do they vary by manufacturer?
- What are the sizing, commissioning and setup practices for best energy performance while not negatively affecting occupant comfort?

- How does sizing, commissioning and installation differ between extended capacity and non - extended capacity variable speed heat pumps?

2. Site Selection and Recruitment

Energy Trust provided SBW with program tracking data of recent (2015-2018) incentivized heat pump installations which included characteristics of the homes and heat pump makes and models. SBW compared the program participant heat pumps to the list of ECHPs provided by CLEAResult to characterize the population. SBW identified approximately 400 program participants which appear to have installed ECHPs out of 5,672 incentivized heat pump installations. The majority of these were installed in site-built single-family homes with basement or crawlspace foundations supplied with electric heat only (no gas backup). Approximately 90% of the ECHP installations were in homes west of the Cascades. SBW could not readily identify standard (base case) VCHPs in the program tracking data without extensive effort to look up each model on each manufacturer's website.

For a population of 400 ECHPs and assuming a coefficient of variation of 0.5, a simple random sample design targeting 10% precision at 90% confidence would require a sample of 58 ECHPs. A sample size of 10 would be sufficient to achieve 20% precision at 80% confidence, depending on observed variance in performance.

The team recruited sites for this study concurrently with the BPA heat pump study. For the BPA study, SBW targeted 24 site visit completions in western Oregon and up to 28 site visits in eastern Oregon[8]. SBW identified and selected a subsample of homes with ECHPs and a subsample of homes with VCHPs. During the recruiting call or email, the SBW field engineer requested the heat pump nameplate model number, informed the participant about the add-on metering study, and offered an incentive to participate. If the participant expressed interest in the add-on study, the field engineer asked the participant to also supply a photograph of their breaker panel(s). The field engineer compared the heat pump model number to a list of ECHP models. The site was recruited for metering installation during the BPA study visit if the following criteria were met:

1. If the heat pump is an extended capacity model and the target number of ECHP sites had not yet been reached OR the heat pump is a "standard" model and the target number of VSHP sites had not yet been reached.
2. The home did not have gas-fired back-up heat.
3. There was sufficient space around the breaker panel to install power metering.
4. The site was in Oregon[9]

SBW provided a $25 gift card to households where metering equipment was installed.

3. Metering Installations

SBW recruited two sites with ECHPs, one east of the Cascades and one west of the Cascades; and six sites with VCHPs, two east of the Cascades and four west of the Cascades. The field engineer installed equipment at each site to collect and record the heat pump and air handler demand at 1-minute intervals and the outside air temperature and refrigerant vapor temperature at 15-minute intervals. The field engineer tested all metering equipment prior to deploying to ensure that it was properly calibrated,

[8] The east of Cascades quota covers eastern Oregon and southern Idaho.

[9] Due to low recruitment rate in Oregon, the team also recruited from similar climates in Washington.

accurately recording data, and had enough capacity to collect the requisite data for the period deployed in the field.

Field engineers installed metering equipment beginning January 2019 through February 2019 during the planned site visits for the BPA heat pump study. The field team retrieved the metering equipment in August 2019 such that data was recorded for up to eight months, covering portions of the heating, cooling, and shoulder seasons.

4. Analysis Methodology

The team developed a model from the collected site data. Specifically, SBW modeled heating, cooling, standby, and total annual electricity savings for ECHPs, compared to base case variable capacity units, for a typical weather year. The analysis controlled for weather and differences in home characteristics. In addition, SBW estimated electricity demand savings, during the coldest days, and modeled savings shape compared to base case heat pumps for a typical weather year. SBW modeled heat pump savings with data from homes located west and east of the Cascades.

Processing the data included rolling up 1-minute interval kW data and 15-minute interval temperature data to hourly interval and identifying the dominant mode in each hour (heating, cooling, or standby). Additionally, SBW used data collected for the BPA study to determine the heating load of each home. The team created regression models for each site with kW/ton as the dependent variable and difference between the balance point and the outdoor air temperature as the independent variable with the mode as categorical variable (i.e., a submodel per mode). SBW developed these site models into generalized models for the base case, or control group (standard VCHPs) and treatment group (ECHPs), respectively. The team calculated average values for overall house heat transfer coefficients and balance points for a typical house from the cohort of houses studied. The team applied typical weather to model the annual heating, cooling, standby, and total savings as well as demand savings during peak hours by season and on coldest days.

5. Results

The team summarized results in two parts. The equipment and settings section describes heat pump nameplate information, thermostat settings, and select findings from the BPA study which may impact energy efficiency. The data results section examines data collected for Energy Trust to compare and evaluate extended capacity heat pumps to standard variable capacity heat pumps.

5.1. Equipment and settings

Table 1 shows the manufacturer, whether the heat pump is considered an extended capacity heat pump or a standard capacity heat pump, and the AHRI certificate for the combination of indoor and outdoor units.

Table 1 Equipment Brand, Reference Number, and Heat pump Type

Site	Brand	Heat Pump Type	AHRI Certificate Number
SE115	Carrier	Extended	6938465
SW49	Carrier	Extended	9892751

SE147	Mitsubishi Electric	Variable[a]	201754323 [b]
SE141	Carrier	Variable	7175587
SW46	Trane	Variable	6750234
PSE3	Trane	Variable	10093505
SW64	Lennox	Variable	5947679
SW92	Carrier	Variable	9893367

[a] This is not on the list of approved extended capacity units, but the manufacturers specification indicates that this meets the extended capacity criteria.

[b] This AHRI reference number is the closest match we could find to this indoor/outdoor unit pairing. All specifications shown in this memo come from the manufacturer.

The capacity ratio between 47°F and 17°F determines heat pump effectiveness at low outdoor air temperatures. Table 2 shows the capacity ratios for studied heat pumps, along with some heating performance metrics. Site SE147 did not use an AHRI certified combination of indoor and outdoor units. The only specification available for this site was the manufacturer specifications which listed efficiency and capacity for the outdoor unit only. A similar AHRI certified combination was chosen as shown in Table 1 to go with the manufacturer specifications in Table 2 to show a range of possible performance metrics for a less well documented outdoor unit. Performance information for site SE147 indicates not all extended capacity heat pumps may be captured in preferred make and model lists for untested combinations. Capacity information at site SW49 counterintuitively indicates that more heat may be delivered at lower temperatures. Nonetheless, all heat pumps have lower COP efficiencies at the lower temperatures. The first three heat pumps shown, have higher overall HSPF ratings than the other heat pumps in addition to greater capacity.

Table 2 Equipment Heating Capacity, Efficiency

Site	Heating Capacity (BTU @ 47°F)	Heating Capacity (BTU @ 17°F)	17°F /47°F - Heating Capacity Ratio	HSPF (Region IV)	COP @ 47°F	COP @ 17°F
SE115	33,400	31,000	93%	12	4.3	2.5
SW49	45,500	47,500	104%	12	4.1	2.2
Avg ECHP[a]	***39,450***	***39,250***	***99%***	***12***	***4.2***	***2.3***
SE147	40,000	38,000	95%	11	3.5	2.6
SE141	34,800	23,600	68%	9.5	Not Available	Not Available
SW46	32,200	25,200	78%	10	Not Available	Not Available
PSE3	32,200	25,200	78%	10	Not Available	Not Available

SW64	32,000	21,000	66%	10	3.3	2.3
SW92	60,000	44,500	74%	9	3.1	2.5
Avg VCHP[a]	***38,533***	***29,583***	***77%***	***9.9***	***3.3***	***2.4***

[a] Average (mean) performance metrics for group of heat pumps

The field team also noted thermostat settings and schedules. A selection of those settings is shown in Table 3. The team considers the heating setpoint during "Sleep Mode" an indicator of how hard a heat pump will need to work during heating season. "Sleep Mode" schedules coincide with electrical peak loads for utility districts. Thermostats may be set manually or scheduled based on homeowner preference to leverage comfort or savings. While the setpoints don't factor into the savings analysis directly, but they do impact balance points which were used in the analysis. Thermostat settings will affect overall energy consumption and will also impact homeowner comfort if set too aggressively to save energy. Setpoint dead bands also impact overall comfort and savings.

Table 3 Thermostat Settings

Site	Scheduled or Manual setpoints	Heating Setpoint During "Sleep Mode" (°F)	Heating Setpoint During "Away Mode" (°F)	Cooling Setpoint During "Away Mode" (°F)	"Away" Heating/Cooling Dead Band (°F)
SE115	Manual	70	70	76	6
SW49	Scheduled	60	68	70	2
SE147	Scheduled	64	62	78	16
SE141	Manual	73	73	84	11
SW46	Manual	74	74	78	4
PSE3	Scheduled	62	70	78	8
SW64	Scheduled	70	70	74	4
SW92	Scheduled	65	68	78	10

Heat pump sizing impacts how often the compressor will run in its most efficient range. Table 4 shows the sizing findings from the heat pump study, and some other results that impact overall system performance. Sizing will be discussed more later. Auxiliary heating lockouts are critical to maximize heat pump savings overall. Auxiliary lockouts will determine when an HVAC system may start combining electric backup heat with the heat pump. Auxiliary heat will never operate above a COP of 1, so using resistance heat when it is not needed will reduce the effectiveness of a heat pump as an energy saver. Compressor lockouts dictate the minimum operating temperature for a heat pump before the entire heating load must be delivered with backup heat. Site SW46 auxiliary lockout was very close to an acceptable temperature, so the data results from this site are likely suitable for comparison. Sites SW64 and PSE3 both had comparatively high or disabled auxiliary lockouts, and Site SW64 had a compressor lockout well above an approved temperature. The results from these two sites may be outside of best practices, but given a mild winter for these sites, the impact of the settings is insignificant so SBW decided to include them in its analysis.

Table 4 Heat Pump Sizing and Lockouts

Site	Installed Heating Capacity (kbtu/hr)	Required Heating Capactity (kbtu/hr)	Sizing (6 kbtu/hr error band)	Auxiliary Heat Lockout	Compressor Lockout
SE115	33	23	Oversized	Correct	Correct
SW49	46	39	Right-sized	Correct	Correct
SE147	38	36	Oversized	Correct	Correct
SE141	35	23	Oversized	Correct	Correct
SW46	32	40	Undersized	Incorrect	Correct
PSE3	32	29	Right-sized	Incorrect	Correct
SW64	32	59	Undersized	Incorrect	Incorrect
SW92	60	33	Oversized	Correct	Correct

Energy Trust wanted to know what operational characteristics of extended capacity heat pumps would provide additional savings compared to variable capacity heat pumps. They also wanted to know which installation, and commissioning practices will leverage energy efficiency while not affecting comfort.

The study of unit characteristics found that only the capacity ratios and the overall unit efficiencies may have an impact on energy savings. The team discusses some observations from standby operation in the Data Results section.

Homeowners choose their setups based on their preferences and contractor decisions, but they may not have a full understanding of how these choices impact their HVAC performance. Table 3 shows that setpoint choices may vary a lot between homes. These choices should be considered when comparing heat pumps with billing data, especially during peak periods. Heat pumps with narrower dead bands may see more energy use during shoulder season weather.

The team found that there are no differences in best practices for installation of ECHPs compared to VCHPs. The best practices for sizing, duct-sealing and settings apply regardless of heat pump type chosen. The senior scientist on the BPA heat pump study team mentioned that correctly sizing ducts to heat pumps may impact user comfort. If ducting systems are sized for forced air furnaces, the supply air at the registers will exit faster which may cause sensations of “cold drafts”. Studies by the DOE also show a marginal performance improvement for any variable capacity heat pump if the unit is somewhat oversized[10]. Some of these best practices are critical to realizing savings between extended capacity units and standard variable capacity units. Lockout temperatures must be set correctly[11] to ensure that

[10] https://www1.eere.energy.gov/buildings/publications/pdfs/building_america/variable-capacity-heatpumps-indoor-ductwork.pdf

[11] Typically 35 °F for auxiliary heat, and 5 °f or not at all for the compressor

heat pumps generate heat during the most advantageous outdoor air temperature ranges. BPA's PTCS guideline describes the lockout settings in detail[12][13].

5.2. Data Results

In August 2019, SBW returned to all eight of the metered sites to retrieve all data loggers and download data for final analysis. All electrical metering data was downloaded successfully. Outside air temperature data could not be downloaded from one site due to a logger failure, however historical weather data from a nearby airport was inserted as a proxy for the site-specific outside air temperature data.

Figure 1 shows modeled hourly kW/Ton as a function of outside air temperature for the total system (HP, air handler, and auxiliary heat) for all eight sites in the study. The Tons[14] are the modeled heating and cooling loads for each house and not the heat pump nameplate Tons. Lower kW/Ton to deliver a heating or cooling ton at a particular temperature shows greater unit efficiency. The span of temperatures from 45 °F to 75 °F shows an increase in the required kW to meet space conditioning requirements. This was a result of short and low capacity compressor cycles to deliver small amounts of useful heat. Over short periods, this resulted in high ratios of kW/Ton. In contrast, more extreme outdoor air temperatures allowed for longer compressor operation at high capacity to deliver larger amounts of useful heat. The latter kind of operation yields more favorable kW/Ton.

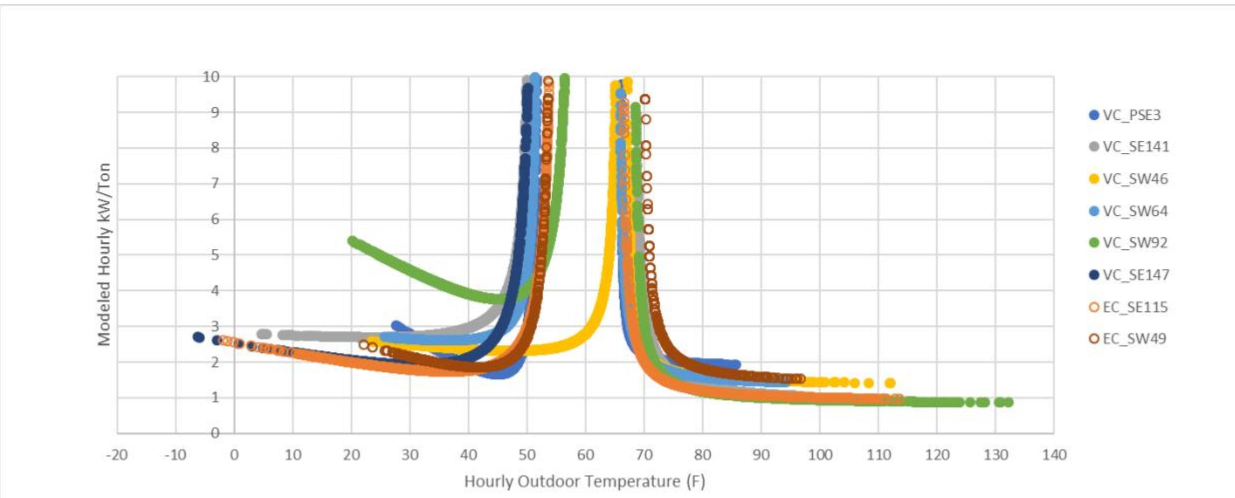


Figure 1 Modeled Hourly kW/Ton by Site

Figure 2 and Figure 3 show the generalized model savings results for Typical Meteorological Years (TMY) in Redmond, OR and Portland, OR, respectively. These figures show the accumulation of savings and TMY dry-bulb temperatures throughout the year.

[12] https://www.bpa.gov/EE/Sectors/Residential/Documents/Notes_on_Auxiliary_Heat_Controls_and_Thermostat.pdf

[13] https://www.bpa.gov/EE/Sectors/Residential/Documents/ASHP_Specifications.pdf

[14] kWh/Ton where [Tons of system demand =((OAT- Balance point) * house UA value)].

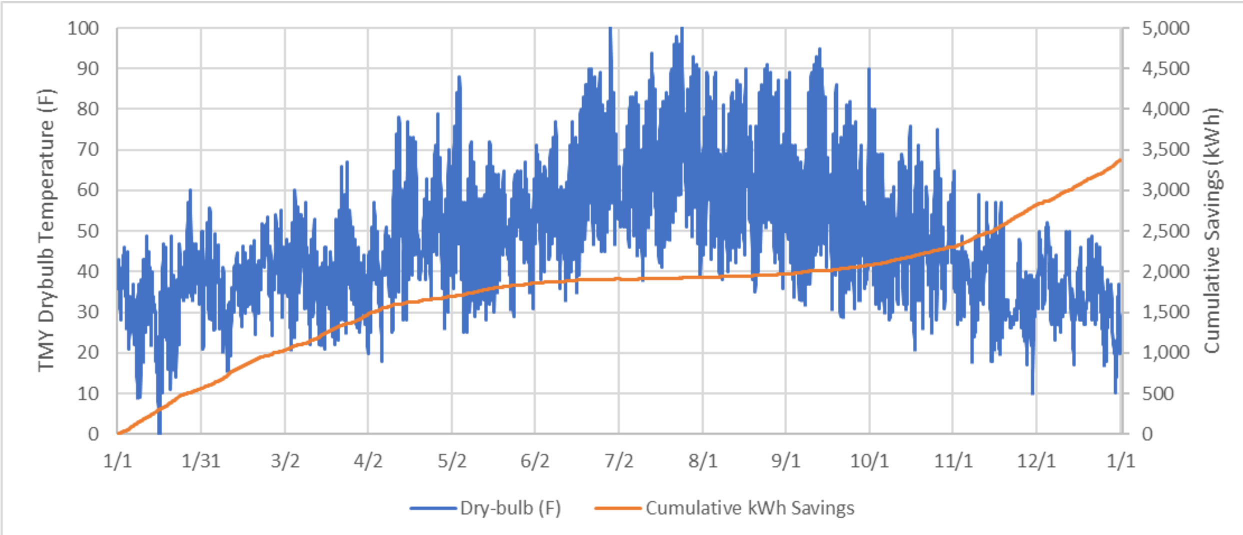


Figure 2 Modeled savings in Redmond, OR for Extended Capacity Heat Pumps Over Variable Capacity Heat Pumps

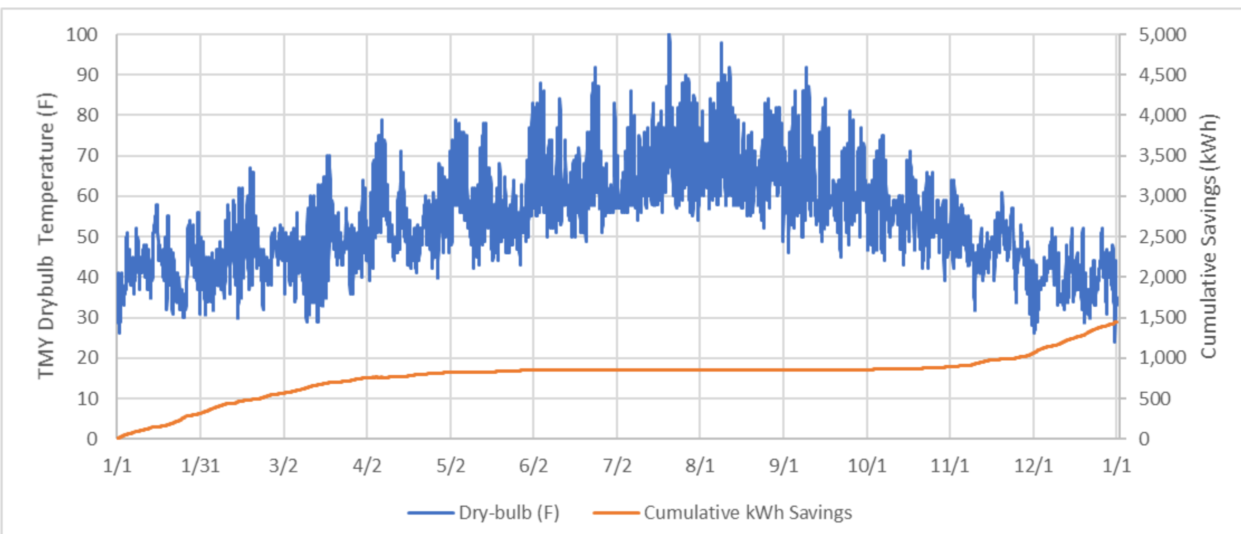


Figure 3 Modeled savings in Portland, OR for Extended Capacity Heat Pumps Over Variable Capacity Heat Pumps

Figure 4 and Figure 5 show the modeled demand results with respect to TMY dry-bulb temperatures for Redmond and Portland respectively. The charts show some difference between the extended capacity and standard capacity heat pump operation in the colder months and a moderate difference in hotter months. The demand drops to a very low standby rate as the dry-bulb temperatures are within the heating and cooling balance points for the models.

The hatched points on the secondary vertical axis show the number of collected data points represented within each point of modeled demand. Only three of eight heat pumps operated at sub 17°F during the study, and one of those three was an extended capacity heat pump. Fewer available data points resulted in some uncertainty and modeling scatter at the coldest temperatures in Figure 4. Normalized kW/Ton models display discontinuity when TMY dry-bulb temperatures approach heating and cooling balance points. For the range of temperatures where the houses were between the balance points, a mean (flat) kW for each heat pump type was used instead of the model. The extended capacity units used marginally more power, but not enough to verify a difference in operating efficiency. This represents a period where the heat pumps run very intermittently, effectively in standby mode. The number of hours in this mode is known for each model and shown in the supporting documentation, but not plotted in these charts.

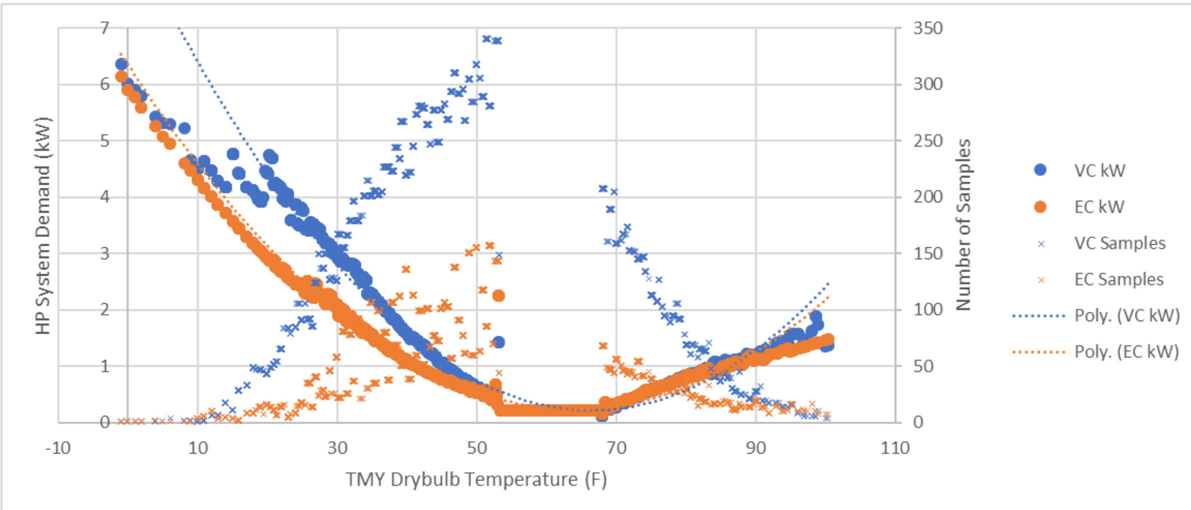


Figure 4 Normalized Hourly Demand – Redmond, OR

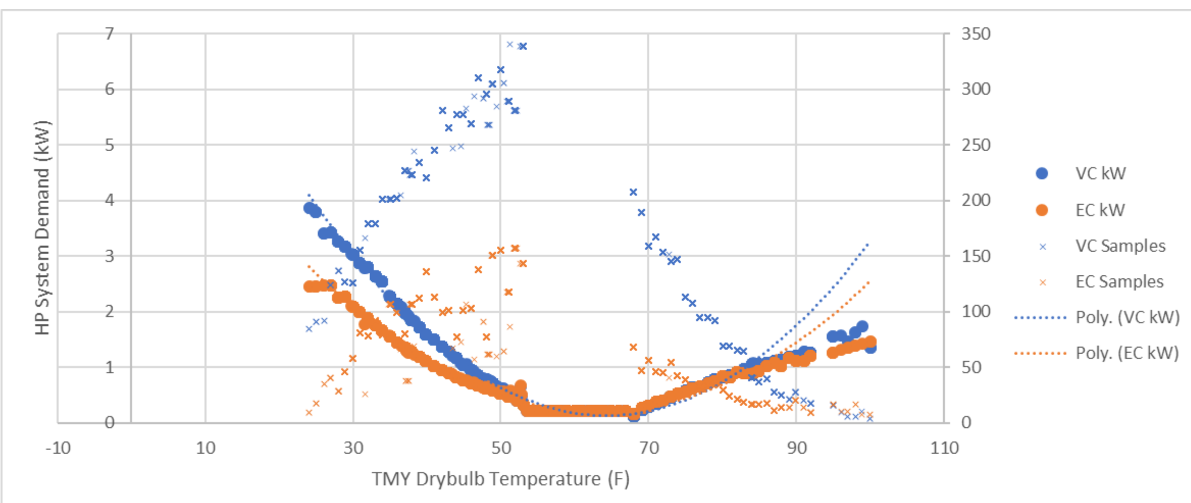


Figure 5 Normalized Hourly Demand – Portland, OR

Table 5 shows overall observed savings and normalized savings results for east (Redmond, OR) and west (Portland, OR) climate zones. Table 6 and Table 7 show modeled demand per ton during the peak seasonal hours for Redmond, OR (Pacific Power) and Portland, OR (PGE) respectively. The peaks are the mean kW savings during the peak schedules specified by the utility. The results during the utility peaks do not necessarily correspond with the coldest or hottest dry-bulb temperatures shown in Figure 4 and Figure 5.

Table 5 Total Savings in Redmond, OR and Portland, OR Models

	Redmond, OR		Portland, OR	
	Savings	Model Hours	Savings	Model Hours
Heating (kWh)	3,353	5,870	1,447	4,463
Heating (kWh/Ton[a])	1,007	NA	434	NA
Cooling kWh	21	1,183	4	1,242
Cooling (kWh/Ton[a])	6	NA	1	NA

[a] Weighted rated heating tons from studied heat pumps.

Table 6 Demand Savings During Pacific Power Peak Periods (Redmond, OR TMY)

	Peak Savings (Avg kW)	Peak Savings (Min kW)	Peak Savings (Max kW)	Peak Savings (Avg kW/ Tons[a])	Peak Month(s)	Peak Time
Winter Morning Peak	0.80	0.12	1.51	0.24	Jan	6:00-10:00
Winter Evening Peak	0.79	0.00	1.51	0.24	Jan	17:00-19:00
Summer Peak	0.04	-0.11	0.32	0.01	Jul	11:00-20:00

[a] Weighted rated heating tons from studied heat pumps.

Table 7 Demand Savings During Portland General Electric Peak Periods (Portland, OR TMY)

	Peak Savings (Avg kW)	Peak Savings (Min kW)	Peak Savings (Max kW)	Peak Savings (Avg kW/ Tons[a])	Peak Month(s)	Peak Time
Winter Morning Peak	0.55	-0.01	1.41	0.17	Dec-Jan	7:00-11:00

Winter Evening Peak	0.45	-0.01	0.94	0.13	Dec-Jan	17:00-21:00
Summer Peak	0.01	-0.11	0.32	0.00	Jul-Aug	13:00-21:00

[a] Weighted rated heating tons from studied heat pumps.

Table 8 shows site inputs for house UA values, peak temperatures difference, peak modeled heating loads, and measured peak electric loads. House UA values are a combination of area and heat loss coefficients for all house surfaces. Peak heating temperature difference is the difference between the coldest observed outdoor air temperature and the observed balance points for the house. These two values combined create a modeled peak heating load.[15] Peak kW comes from a model correlating the inside-outside temperature difference and measured HVAC power. The peak kW value was the largest hourly observed HVAC power consumption for each site. The savings results will vary proportionally to the UA values estimated for the normalized house exposed to TMY conditions. Two sites (SW46 and SW92) were thought to be undersized in Table 4. The modeled peak heating load in Table 8 surpassed the 32 kBTU (2.66 Ton) installed capacity for site SW46. This was the only case where sizing may have negatively impacted energy consumption.

Table 8 Site model characteristics

Site	UA (Btu/°F)	Peak Heating Temp Difference °F	Peak Modeled Heating (Tons)	Rated Heating (Tons)	Peak Heating (kW)
SE115	361.3	57	1.71	2.78	4.47
SW49	802.0	32	2.17	3.79	5.39
SE147	592.2	58	2.85	3.33	7.70
SE141	332.2	46	1.28	2.90	3.58
SW46	767.8	42	2.72	2.68	7.01
PSE3	454.8	25	0.93	2.68	2.81
SW64	606.3	26	1.33	2.67	3.59
SW92	385.6	38	1.23	5.00	6.64

At face value Figures 2-5 indicate a significant difference in energy use between ECHPs and VCHPs, especially at outside air temperatures that require heating. Figure 2 and Figure 3 indicate that most of the energy savings are in colder months. Portland, OR and Redmond, OR show little or no savings in cooling dominated months. The rate of savings accumulation slows in the shoulder seasons, which is a result of warmer outdoor air temperatures. Extended capacity heat pumps may perform better than standard variable capacity heat pumps at colder outdoor air temperatures. The models have heating savings of 3,353 kWh in Redmond, OR and 1,447 kWh and Portland, OR, and the normalized savings are

[15] Peak Load (Tons) = UA * Peak Heating Temperature difference /12000

1,107 kWh/Rated Ton and 434 kWh/Rated Ton respectively (Table 5). This is an expected result from a technology designed to operate better at colder temperatures.

At the lowest temperatures in the Redmond, OR-modeled demand savings (Figure 4), the cluster of modeled points appear to converge. This may result from similar operations between standard and extended capacity heat pumps at extremely low temperatures (defrost, more auxiliary heat, etc.), but this is a small sample size both in number of sites and number of cold days. There may also be some overall savings on hot days (Figure 4 and Figure 5) but it appears negligible. The cooling demand may increase during utility peak hours (Table 6

Table 5 Total Savings in Redmond, OR and Portland, OR Models

	Redmond, OR		Portland, OR	
	Savings	**Model Hours**	**Savings**	**Model Hours**
Heating (kWh)	3,353	5,870	1,447	4,463
Heating (kWh/Ton[a])	1,007	NA	434	NA
Cooling kWh	21	1,183	4	1,242
Cooling (kWh/Ton[a])	6	NA	1	NA

[a] Weighted rated heating tons from studied heat pumps.

and Table 7 **Error! Reference source not found.**) if we consider the minimum savings bound of observed results. Annual savings accumulation seen in Figure 2 and Figure 3 also shows negligible cooling energy savings. Table 6

Table 5 Total Savings in Redmond, OR and Portland, OR Models

	Redmond, OR	Portland, OR

	Savings	Model Hours	Savings	Model Hours
Heating (kWh)	3,353	5,870	1,447	4,463
Heating (kWh/Ton[a])	1,007	NA	434	NA
Cooling kWh	21	1,183	4	1,242
Cooling (kWh/Ton[a])	6	NA	1	NA

[a] Weighted rated heating tons from studied heat pumps.

and Table 7 show some potential savings during cooling peak hours on average although the uncertainty from those peak periods is high. Shoulder season operation does show savings accumulation in Figure 2 and Figure 3, but this is primarily due to the greater effectiveness of heating at frequent low dry-bulb temperatures in spring and fall.

The results indicate both kW and kWh savings for the two studied extended capacity heat pumps in heating mode. The results for cooling and standby modes are uncertain and any observed savings or losses are small to negligible. The models created in this study come from six variable capacity units and two extended capacity units. Heating load estimates at each site add another risk factor to the small sample. Estimated heating loads depend on UA values calculated by the field team (Table 8). To create representative savings estimates, representative rated household heating tons may be multiplied with the normalized results[16] for eastern and western climate zones. Rated tons yield an approximation of the UA values and peak temperature differences used in savings models[17]. The results for savings during the heating season are compelling, but the sample is not representative of a cross section of Energy Trust's program participants. We recommend use of any results from this study with caution and collection of direct metering data from more sites to build a robust data set.

[16] Savings (kWh) = Normalized kWh Savings [kWh/Ton] *representative heating tons [Ton]

[17] Rated [Tons] ≈ UA [BTU/°F] * Peak Heating Temperature Difference [°F] * $\frac{1\ [Ton]}{12000\ [BTU]}$

Attachment 3:

Recurve Analytics. 2020. Extended Capacity Heat Pump Impact Analysis Snapshot Reports, 2015-2018.

a. Overall electricity savings
b. Heating zone 1 electricity savings
c. Heating zone 2 electricity savings
d. Commissioned heat pump electricity savings
e. Non-commissioned heat pump electricity savings
f. Carrier heat pump electricity savings
g. Trane heat pump electricity savings
h. Other heat pump manufacturer electricity savings

a. Overall electricity savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	
Meter Data Filters:	DNAC: <100%	DNAC Percentile: All	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: All
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

394 Treatment Meters	1922 +/- 250 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	9 +/- 1 % Percent Normal Year Pre-Post Difference in Consumption per Participant	20,391 Mean Baseline Consumption (Electricity)	73% Realization Rate
1,938 Site-level Matched Meters	1300 +/- 271 kWh Average Savings Relative to Site-level Matched Comparison Group	6 +/- 1% Percent Savings Relative to Site-level Matched Comparison Group	20,117 Mean Baseline Consumption (Electricity)	49% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

1 Report Date: March 2, 2020

RECURVE

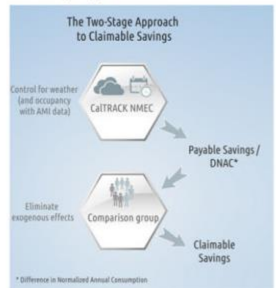
1. Introduction

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).

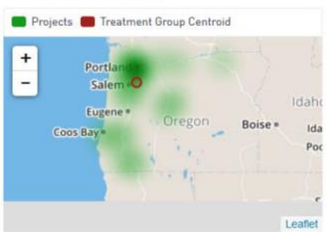
The report includes the following sections:
Result Summary - Includes the overall portfolio results
Section 1. Introduction - Overview of report and the different groups included in the analysis
Section 2. Data Preparation - Data cleaning and sample attrition
Section 3. Modeling Results - CalTRACK model 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



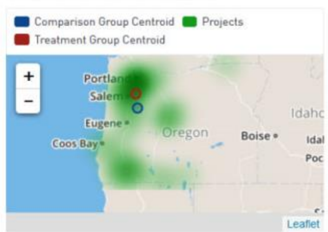
97.1 miles

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

394	20,391
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



35.4 miles

Distance between treatment and comparison group centroids

1,938	20,117
Meters	Mean Baseline Consumption (Electricity)

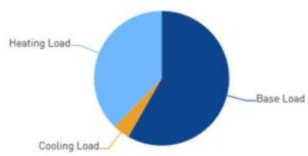
Future Participant Group
The pool of sites that was used to create this group was composed of sites that installed the same measure in the year following the specified program year. The final sites were selected by stratified sampling using deciles of annual energy consumption.

Future Participant Site Locations

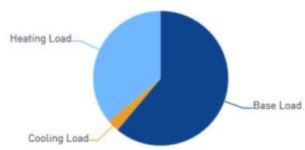
Distance between treatment and future participant group centroids

0	
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.

817	394	48%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

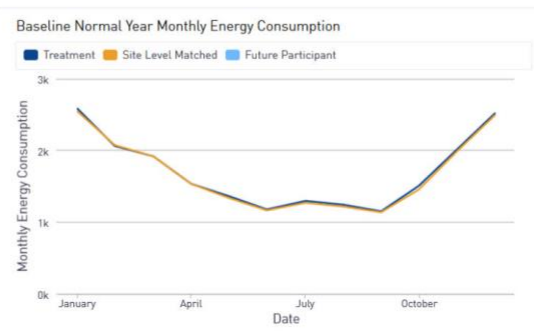
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	167	609
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	145	464
Meters with at least 5 site-level matched meters from the comparison group pool.	--	5	459
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	6	453

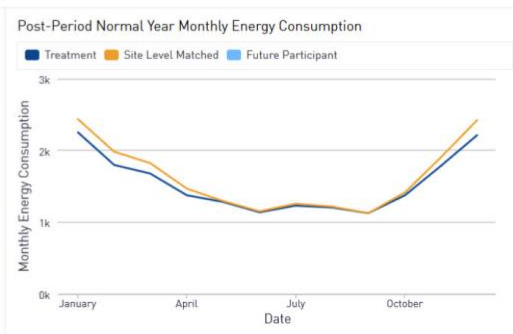
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	453
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	2	451
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	57	394
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	394
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	394
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	394
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	394

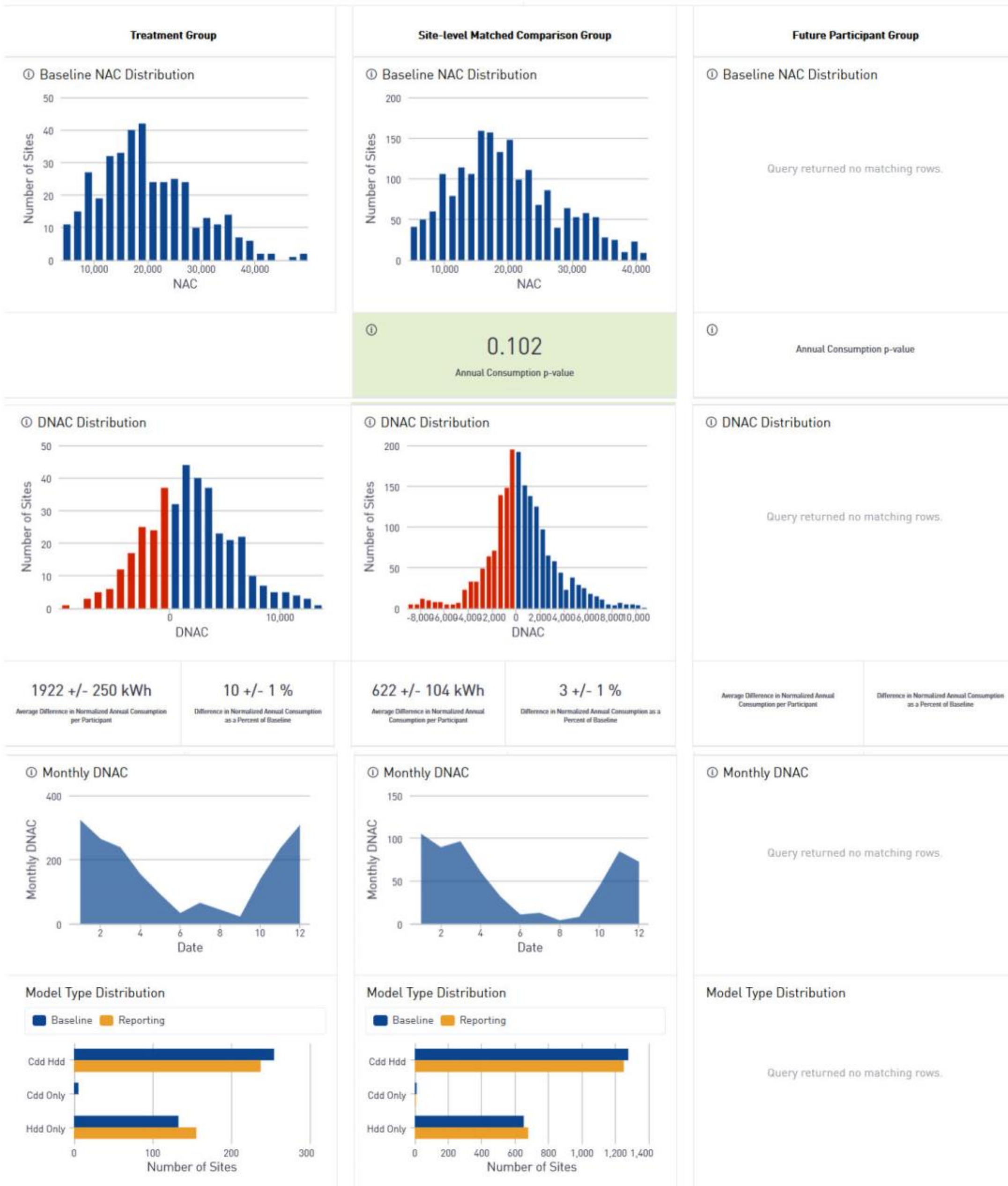
3. Modeling Results

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.

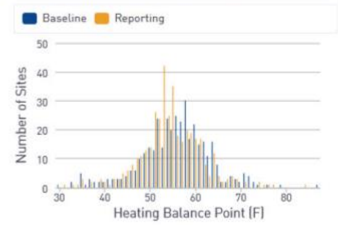




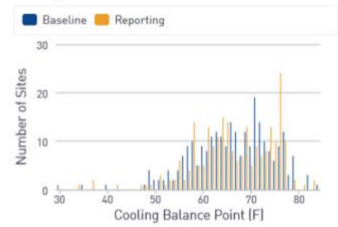

Treatment Group
Site-level Matched Comparison Group
Future Participant Group
Baseline NAC Distribution
Number of Sites
NAC
Query returned no matching rows.
0.102
Annual Consumption p-value
Annual Consumption p-value
DNAC Distribution
DNAC
1922 +/- 250 kWh
Average Difference in Normalized Annual Consumption per Participant
10 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline
622 +/- 104 kWh
Average Difference in Normalized Annual Consumption per Participant
3 +/- 1 %
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
Date
Model Type Distribution
Baseline
Reporting
Cdd Hdd
Cdd Only
Hdd Only
Number of Sites

Treatment Group

Heating Balance Point Distribution

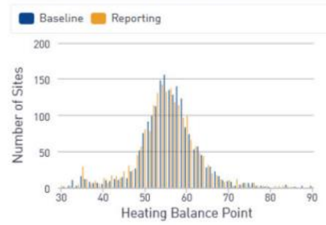


Cooling Balance Point Distribution

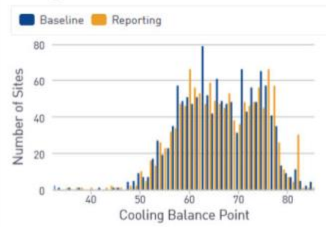


Site-level Matched Comparison Group

Heating Balance Point Distribution



Cooling Balance Point Distribution



Future Participant Group

Heating Balance Point Distribution

Query returned no matching rows.

Cooling Balance Point Distribution

Query returned no matching rows.

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.

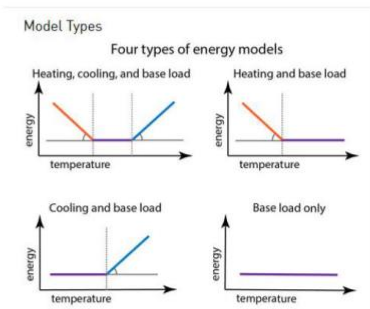
Comparison Group Generation

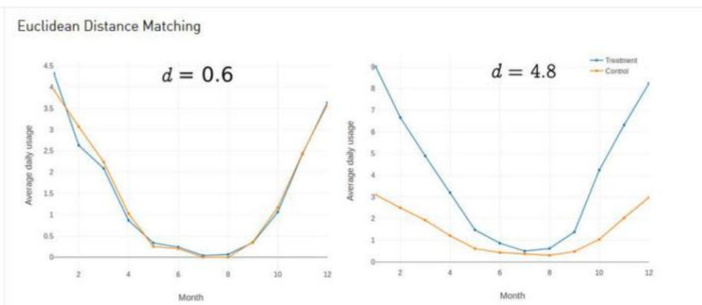
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





b. Heating Zone 1 Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	
Meter Data Filters:	DNAC: <100%	DNAC Percentile: All	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 Zone(s): 1 - Hdd <= 6000	Heating Fuel: Electricity	Heat Pump Manufacturer: All
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

356 Treatment Meters	1964 +/- 263 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	10 +/- 1 % Percent Normal Year Pre-Post Difference in Consumption per Participant	20,448 Mean Baseline Consumption (Electricity)	76% Realization Rate
1,736 Site-level Matched Meters	1239 +/- 285 kWh Average Savings Relative to Site-level Matched Comparison Group	6 +/- 1% Percent Savings Relative to Site-level Matched Comparison Group	20,173 Mean Baseline Consumption (Electricity)	48% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

1 RECURVE Report Date: March 2, 2020

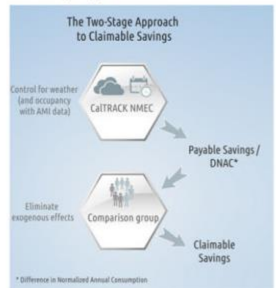
1. Introduction

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

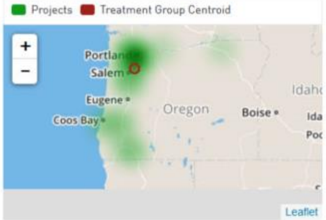
Two-Stage Approach



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



76.2 miles

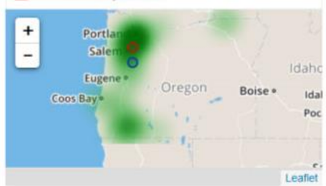
80% of projects lie within this distance from treatment group centroid

356	20,448
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



35.9 miles

Distance between treatment and comparison group centroids

1,736	20,173
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

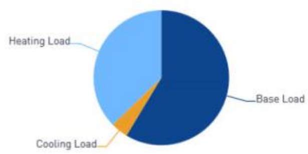
Comparison Group Centroid
Treatment Group Centroid

Distance between treatment and future participant group centroids

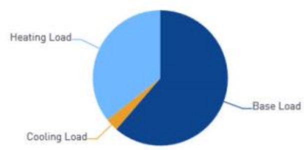
0	
Meters	Mean Baseline Consumption (Electricity)

RECURVE

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.

817	356	44%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

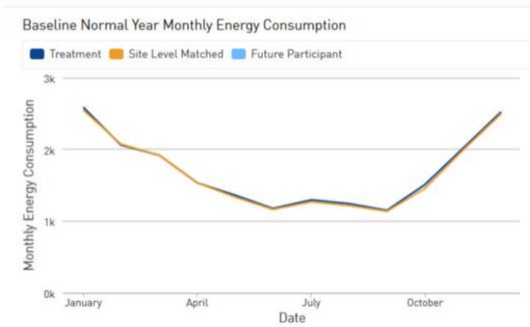
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 1 - Hdd <= 6000 -- Cooling Zone: All	77	699
Other measure-specific filters.	--	161	538
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	121	417
Meters with at least 5 site-level matched meters from the comparison group pool.	--	3	414
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	6	408

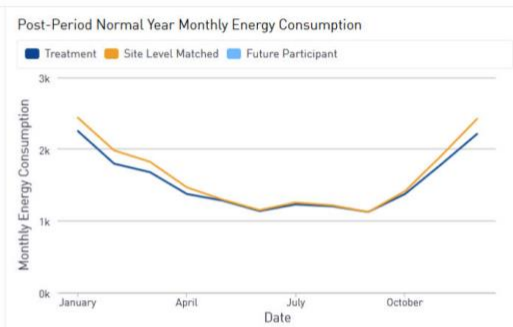
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	2	406
R2_Threshold: Meters with valid model R-squared for the baseline and reporting periods that meet a specified threshold. Models may have invalid R-squared due to data issues.	R-Squared: >0.5	50	356
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	356
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	356
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	356
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	356

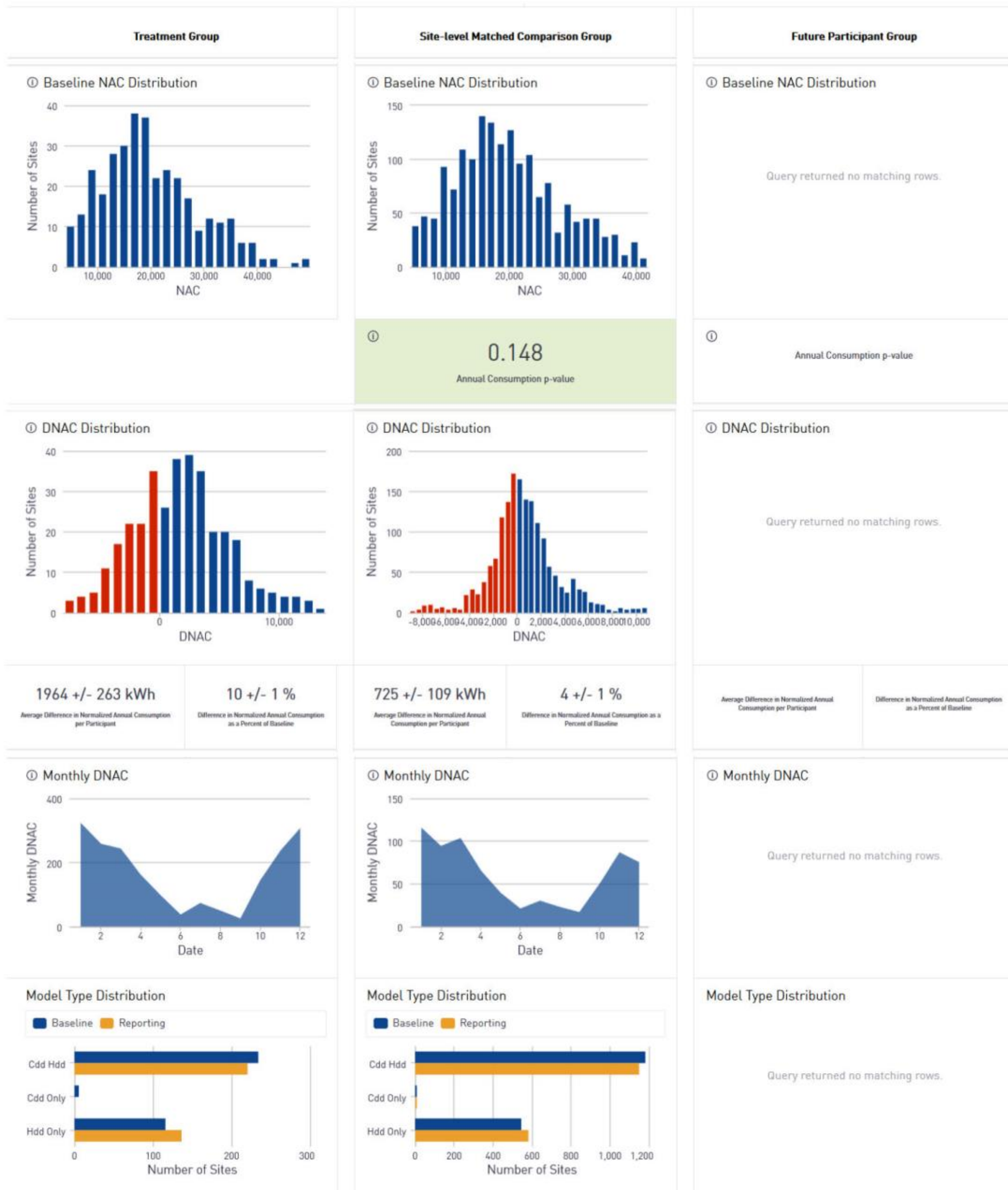
3. Modeling Results

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.

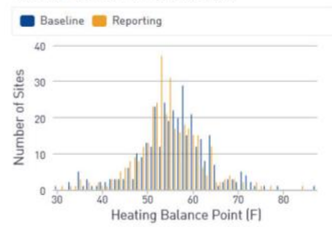




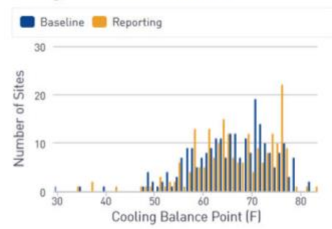

Treatment Group
Site-level Matched Comparison Group
Future Participant Group
Baseline NAC Distribution
Number of Sites
NAC
Query returned no matching rows.
0.148
Annual Consumption p-value
DNAC Distribution
DNAC
1964 +/- 263 kWh
Average Difference in Normalized Annual Consumption per Participant
10 +/- 1 %
Difference in Normalized Annual Consumption as a Percent of Baseline
725 +/- 109 kWh
4 +/- 1 %
Monthly DNAC
Date
Model Type Distribution
Baseline
Reporting
Cdd Hdd
Cdd Only
Hdd Only
Number of Sites

Treatment Group

Heating Balance Point Distribution

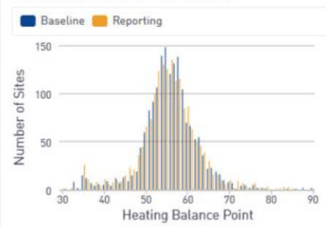


Cooling Balance Point Distribution

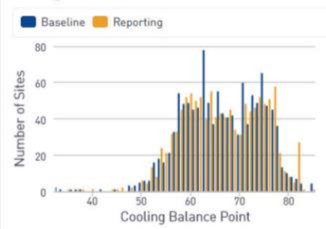


Site-level Matched Comparison Group

Heating Balance Point Distribution



Cooling Balance Point Distribution



Future Participant Group

Heating Balance Point Distribution

Query returned no matching rows.

Cooling Balance Point Distribution

Query returned no matching rows.

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.

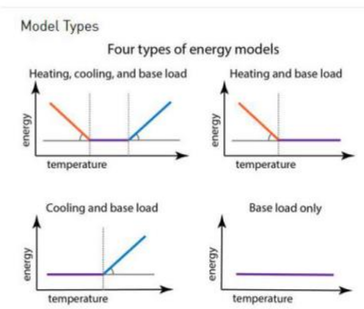
Comparison Group Generation

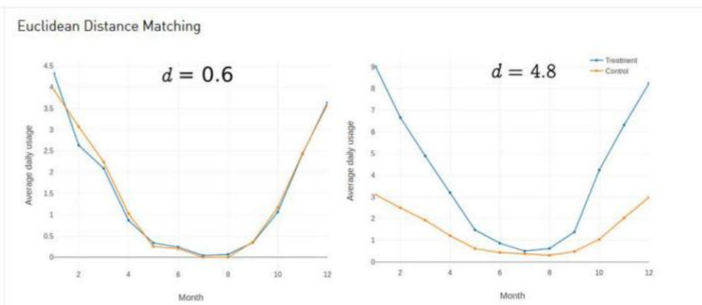
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





c. Heating Zone 2 Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	
Meter Data Filters:	DNAC: <100%	DNAC Percentile: All	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 Zone(s): 2 - 6000 < Hdd < 7500, 3 - Hdd >= 7500	Heating Fuel: Electricity	Heat Pump Manufacturer: All
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

31 Treatment Meters	1797 +/- 946 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	8 +/- 4 % Percent Normal Year Pre-Post Difference in Consumption per Participant	21,314 Mean Baseline Consumption (Electricity)	60% Realization Rate
44 Site-level Matched Meters	1425 +/- 1203 kWh Average Savings Relative to Site-level Matched Comparison Group	7 +/- 6% Percent Savings Relative to Site-level Matched Comparison Group	22,277 Mean Baseline Consumption (Electricity)	47% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

1 Report Date: March 2, 2020

RECURVE

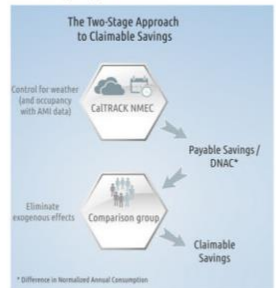
1. Introduction

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

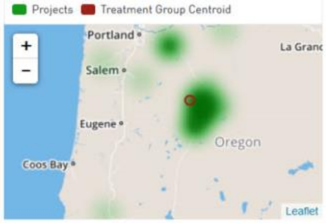
Two-Stage Approach



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



68.9 miles

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

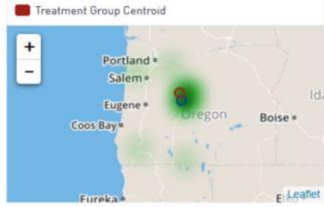
Meters

31	21,314
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



17.5 miles

Distance between treatment and comparison group centroids

44	22,277
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.

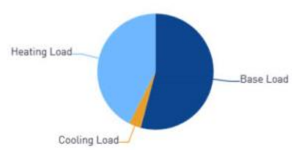
Future Participant Site Locations

Comparison Group Centroid | Treatment Group Centroid

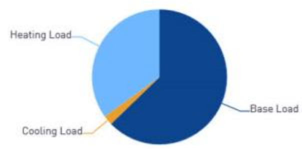
Distance between treatment and future participant group centroids

0	
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.

817	31	3.8%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

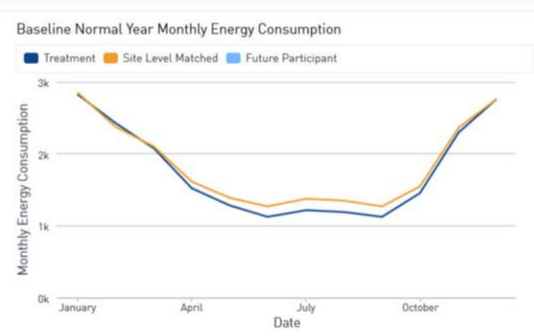
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: 2 - 6000 < Hdd < 7500, 3 - Hdd >= 7500 -- Cooling Zone: All	714	62
Other measure-specific filters.	--	4	58
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	20	38
Meters with at least 5 site-level matched meters from the comparison group pool.	--	0	38
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	0	38

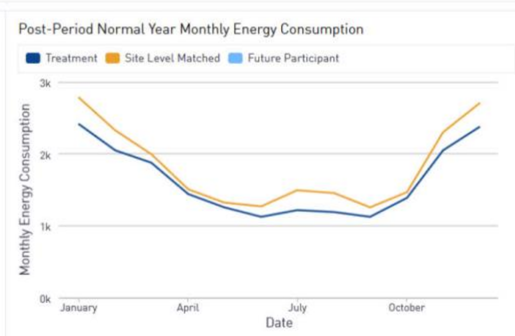
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	38
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	38
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	7	31
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	31
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	31
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	31
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	31

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.

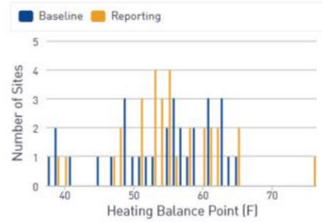




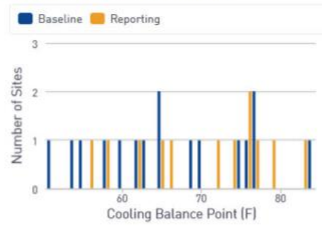

Treatment Group
Site-level Matched Comparison Group
Future Participant Group
Baseline NAC Distribution
Number of Sites
NAC
Query returned no matching rows.
0.497
Annual Consumption p-value
Annual Consumption p-value
DNAC Distribution
DNAC
1797 +/- 946 kWh
Average Difference in Normalized Annual Consumption per Participant
8 +/- 4 %
Difference in Normalized Annual Consumption as a Percent of Baseline
373 +/- 744 kWh
Average Difference in Normalized Annual Consumption per Participant
2 +/- 3 %
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
Date
Model Type Distribution
Baseline
Reporting
Cdd Hdd
Hdd Only
Number of Sites

Treatment Group

Heating Balance Point Distribution

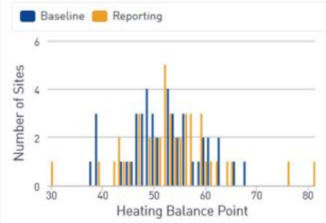


Cooling Balance Point Distribution

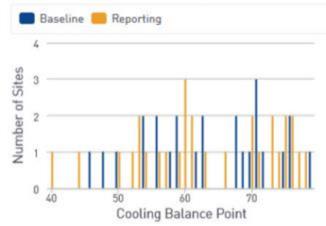


Site-level Matched Comparison Group

Heating Balance Point Distribution



Cooling Balance Point Distribution



Future Participant Group

Heating Balance Point Distribution

Query returned no matching rows.

Cooling Balance Point Distribution

Query returned no matching rows.

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.

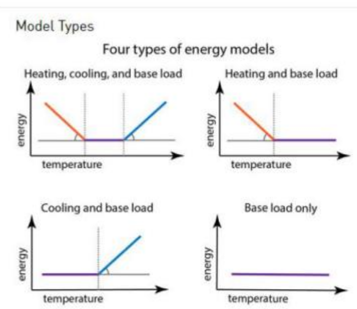
Comparison Group Generation

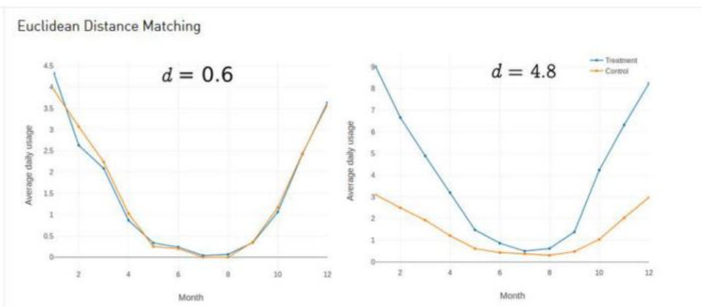
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





d. Commissioned Heat Pump Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <100%	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: All
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: Yes
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

332 Treatment Meters	1753 +/- 271 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	9 +/- 1 % Percent Normal Year Pre-Post Difference in Consumption per Participant	20,015 Mean Baseline Consumption (Electricity)	67% Realization Rate
1,603 Site-level Matched Meters	1092 +/- 294 kWh Average Savings Relative to Site-level Matched Comparison Group	5 +/- 1% Percent Savings Relative to Site-level Matched Comparison Group	20,083 Mean Baseline Consumption (Electricity)	42% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

RECURVE

Report Date: March 2, 2020

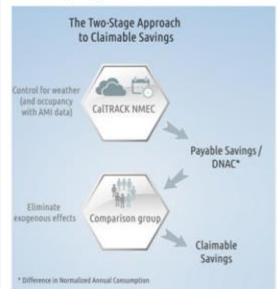
1. Introduction

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

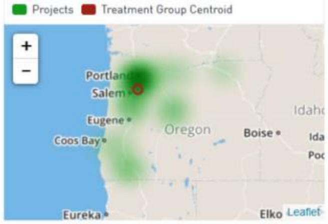
Two-Stage Approach



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



97.1 miles

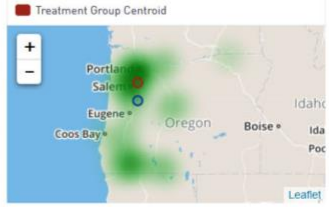
80% of projects lie within this distance from treatment group centroid

332	20,015
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



43.0 miles

Distance between treatment and comparison group centroids

1,603	20,083
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.

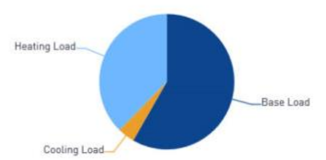
Future Participant Site Locations

Comparison Group Centroid
Treatment Group Centroid

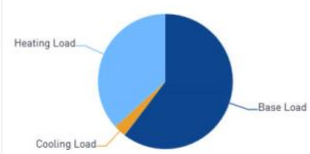
Distance between treatment and future participant group centroids

0	
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.

817	332	41%
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: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	268	508
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	118	390
Meters with at least 5 site-level matched meters from the comparison group pool.	--	2	388
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	6	382

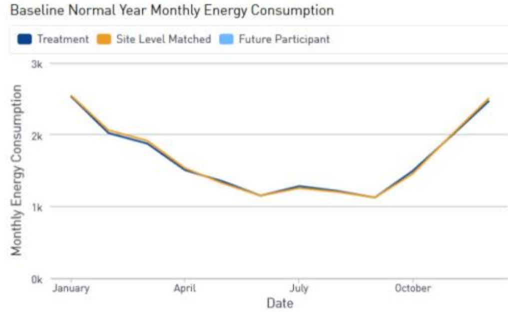
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	382
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	1	381
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	49	332
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	332
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	332
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	332
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	332

3. Modeling Results

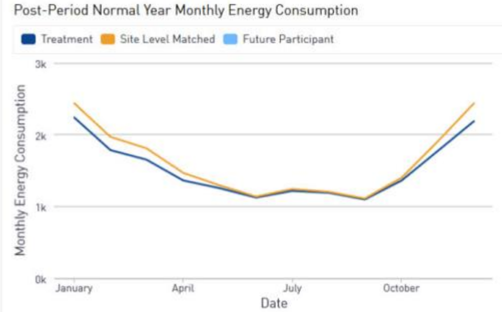
This section includes summaries of the Difference in Normalized Annual Consumption (DNAC) results for the treatment and comparison groups. The time series of monthly energy consumption illustrates the similarities and/or differences in energy consumption for the different groups in the baseline and reporting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

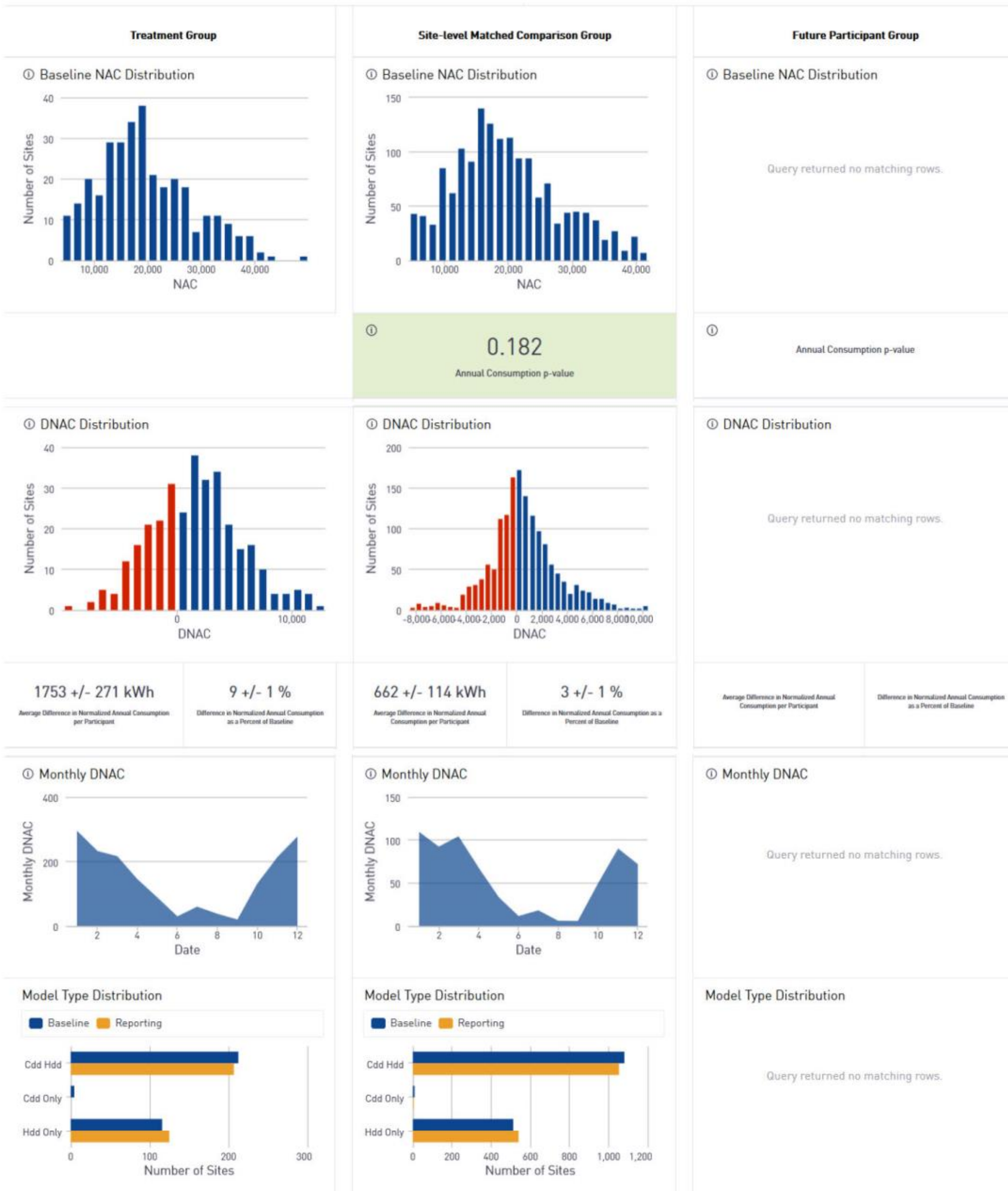
Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional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

Baseline Normal Year Monthly Energy Consumption



Post-Period Normal Year Monthly Energy Consumption





Report Date: March 2, 2020

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.

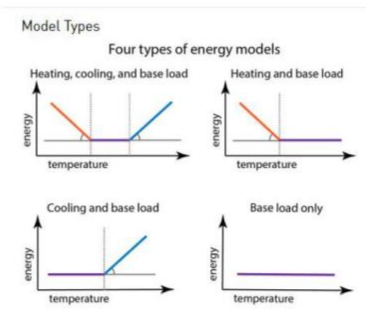
Comparison Group Generation

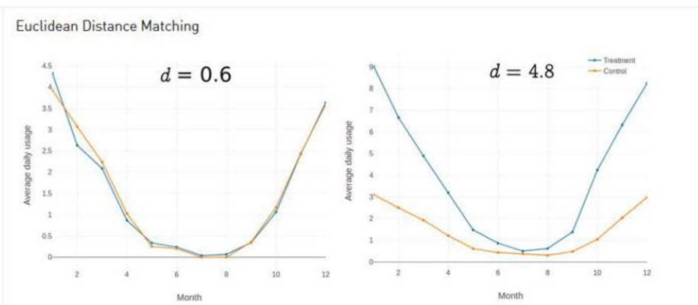
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





e. Non-Commissioned Heat Pump Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	
Meter Data Filters:	DNAC: <100%	DNAC Percentile: All	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: All
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: No
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

62 Treatment Meters	2823 +/- 649 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	13 +/- 3 % Percent Normal Year Pre-Post Difference in Consumption per Participant	22,401 Mean Baseline Consumption (Electricity)	99% Realization Rate
169 Site-level Matched Meters	1612 +/- 769 kWh Average Savings Relative to Site-level Matched Comparison Group	7 +/- 3% Percent Savings Relative to Site-level Matched Comparison Group	23,344 Mean Baseline Consumption (Electricity)	56% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

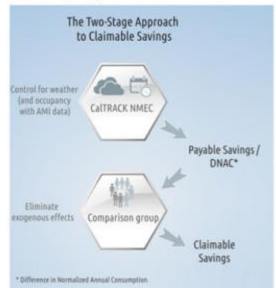
1. Introduction

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

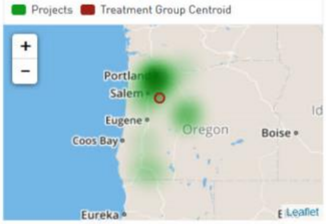
Two-Stage Approach



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



165.7 miles

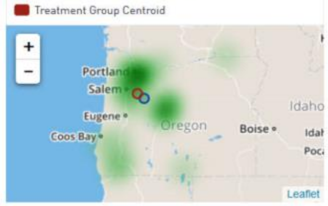
80% of projects lie within this distance from treatment group centroid

62	22,401
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



17.8 miles

Distance between treatment and comparison group centroids

169	23,344
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.

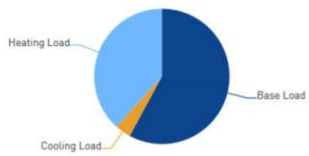
Future Participant Site Locations

Comparison Group Centroid
Treatment Group Centroid

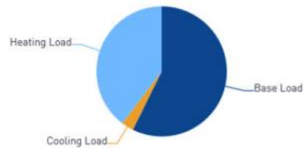
Distance between treatment and future participant group centroids

0	
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.

817	62	7.6%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

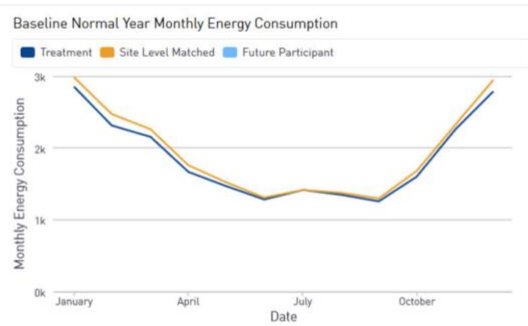
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	675	101
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	27	74
Meters with at least 5 site-level matched meters from the comparison group pool.	--	3	71
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	0	71

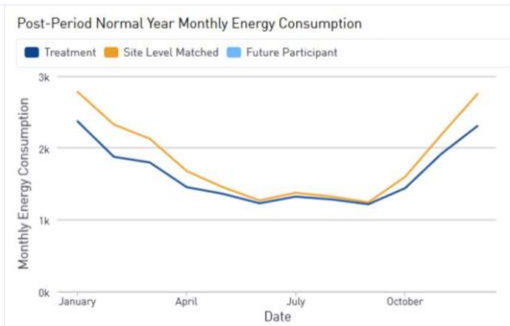
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	71
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	71
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	9	62
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	62
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	62
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	62
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	62

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.

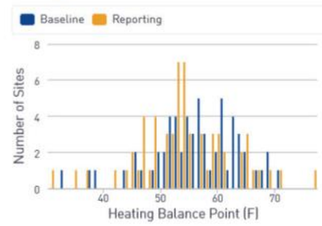




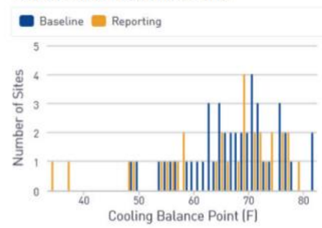

Treatment Group
Site-level Matched Comparison Group
Future Participant Group
Baseline NAC Distribution
Number of Sites
NAC
Query returned no matching rows.
0.298
Annual Consumption p-value
DNAC Distribution
DNAC
2823 +/- 649 kWh
Average Difference in Normalized Annual Consumption per Participant
13 +/- 3 %
Difference in Normalized Annual Consumption as a Percent of Baseline
1211 +/- 414 kWh
5 +/- 2 %
Monthly DNAC
Date
Model Type Distribution
Baseline
Reporting
Cdd Hdd
Cdd Only
Hdd Only
Number of Sites

Treatment Group

Heating Balance Point Distribution

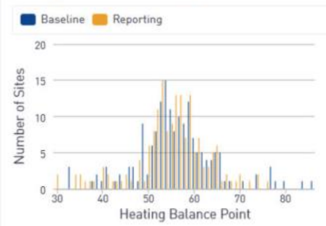


Cooling Balance Point Distribution

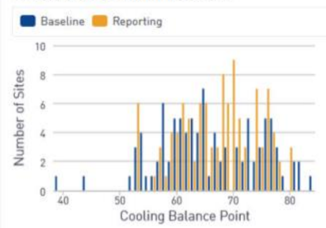


Site-level Matched Comparison Group

Heating Balance Point Distribution



Cooling Balance Point Distribution



Future Participant Group

Heating Balance Point Distribution

Query returned no matching rows.

Cooling Balance Point Distribution

Query returned no matching rows.

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.

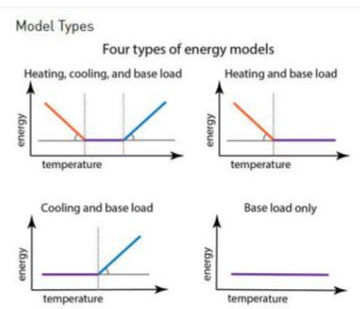
Comparison Group Generation

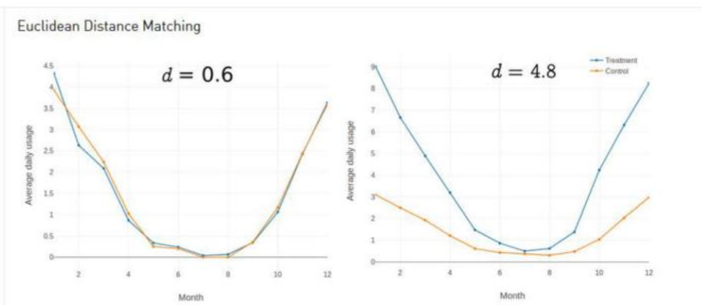
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





RECURVE

f. Carrier Heat Pump Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	ⓘ Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <100%	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: Carrier, Carrier
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

122 Treatment Meters	2140 +/- 510 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	ⓘ 10 +/- 2 % Percent Normal Year Pre-Post Difference in Consumption per Participant	21,763 Mean Baseline Consumption (Electricity)	92% Realization Rate
605 Site-level Matched Meters	1519 +/- 545 kWh Average Savings Relative to Site-level Matched Comparison Group	7 +/- 3% Percent Savings Relative to Site-level Matched Comparison Group	21,327 Mean Baseline Consumption (Electricity)	66% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

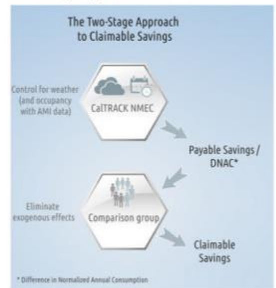
1. Introduction

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

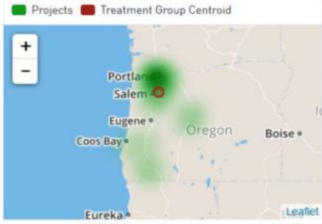
Two-Stage Approach



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



103.2 miles

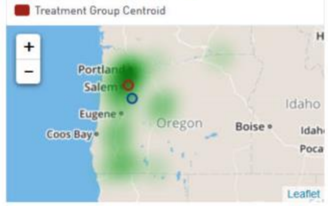
80% of projects lie within this distance from treatment group centroid

122	21,763
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



32.7 miles

Distance between treatment and comparison group centroids

605	21,327
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.

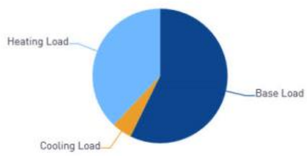
Future Participant Site Locations

Comparison Group Centroid · Treatment Group Centroid

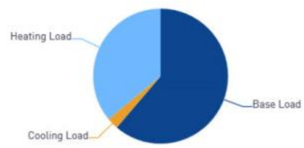
Distance between treatment and future participant group centroids

0	
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.

817	122	15%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

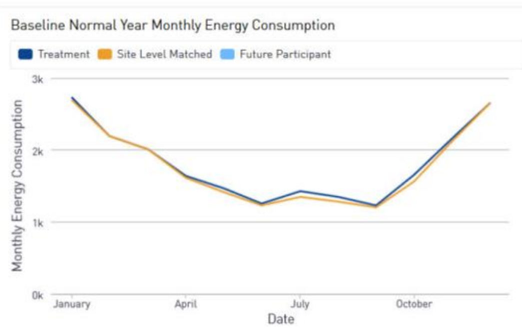
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	584	192
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	49	143
Meters with at least 5 site-level matched meters from the comparison group pool.	--	2	141
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	1	140

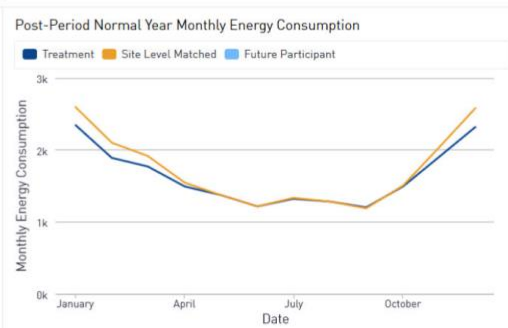
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	140
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	140
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	18	122
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	122
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	122
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	122
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	122

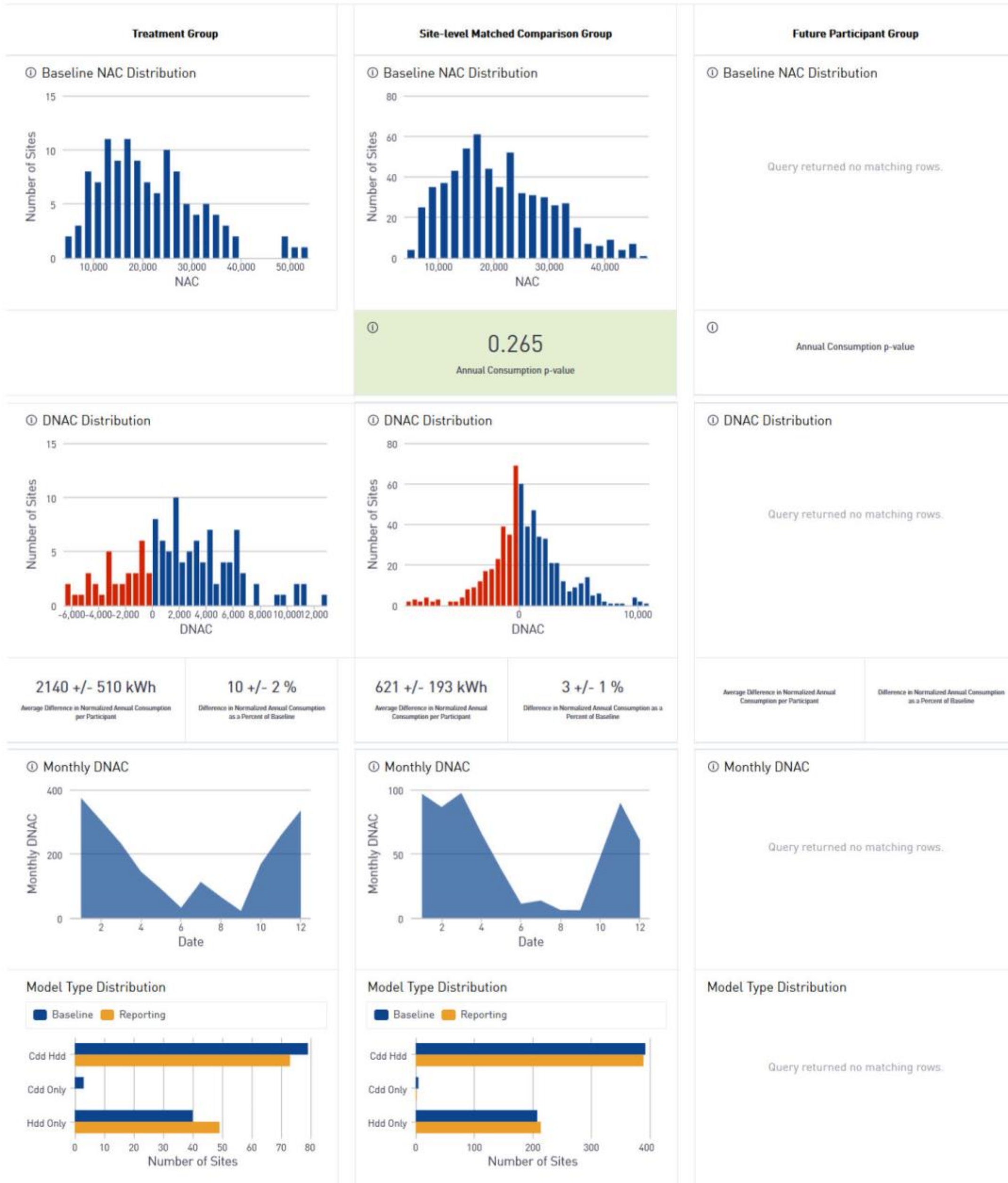
3. Modeling Results

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.

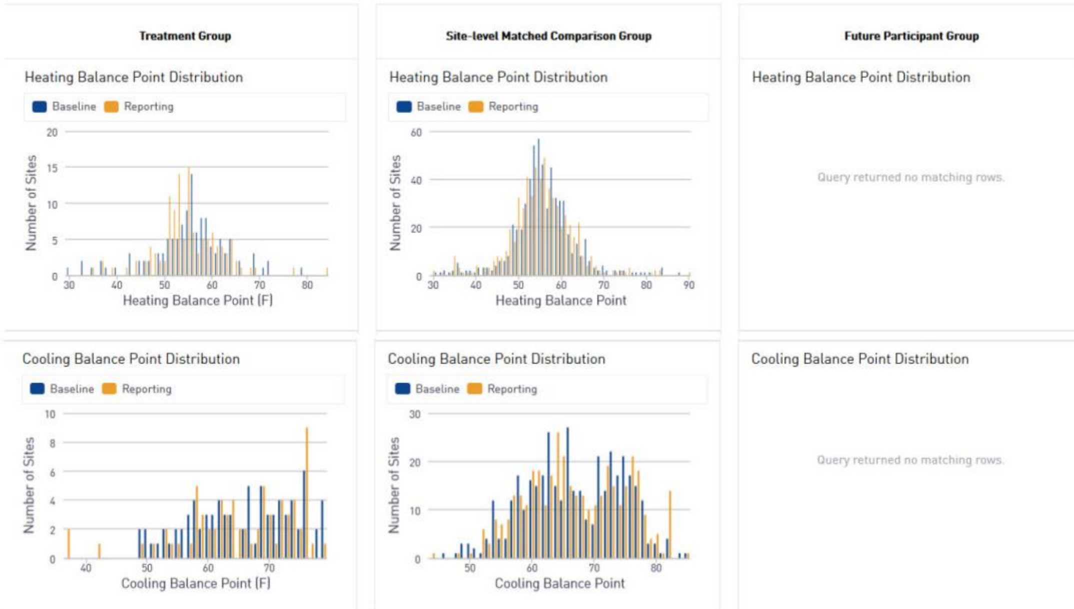
Below, you will find a breakdown of the DNAC results by group, showing the histograms of DNAC as well as the mean value expressed in raw units and as a percent of baseline annual consumption. Finally, the distribution of model types in the baseline and reporting periods are also provided as an additional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







Treatment Group
Site-level Matched Comparison Group
Future Participant Group
Baseline NAC Distribution
Number of Sites
NAC
Query returned no matching rows.
0.265
Annual Consumption p-value
DNAC Distribution
DNAC
2140 +/- 510 kWh
Average Difference in Normalized Annual Consumption per Participant
10 +/- 2 %
Difference in Normalized Annual Consumption as a Percent of Baseline
621 +/- 193 kWh
3 +/- 1 %
Monthly DNAC
Date
Model Type Distribution
Baseline
Reporting
Cdd Hdd
Cdd Only
Hdd Only
Number of Sites



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.

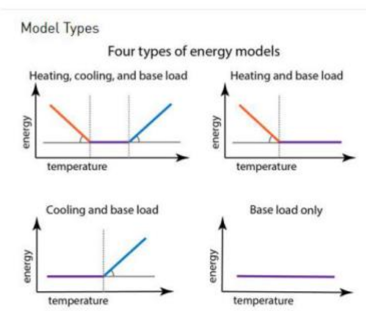
Comparison Group Generation

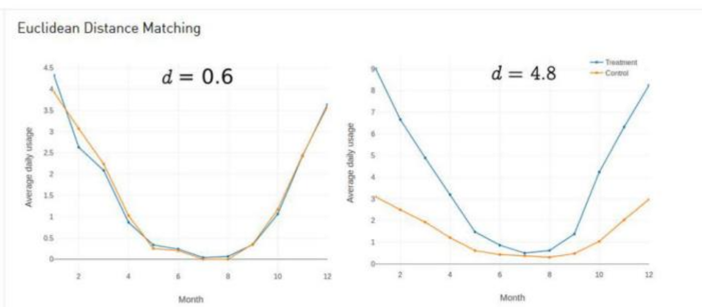
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





g. Trane Heat Pump Electricity Savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	*Last Consumption Data Update:* October 1, 2019 *Last Participation Data Update:* October 1, 2019 *CalTRACK Version:* 2.0
Meter Data Filters:	DNAC: <100%	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: Trane, Trane
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

212 Treatment Meters	1426 +/- 324 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	7 +/- 2 % Percent Normal Year Pre-Post Difference in Consumption per Participant	19,582 Mean Baseline Consumption (Electricity)	49% Realization Rate
1,043 Site-level Matched Meters	893 +/- 353 kWh Average Savings Relative to Site-level Matched Comparison Group	5 +/- 2% Percent Savings Relative to Site-level Matched Comparison Group	19,635 Mean Baseline Consumption (Electricity)	31% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

RECURVE

Report Date: March 2, 2020

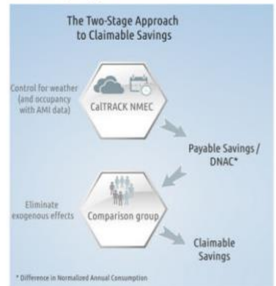
1. Introduction

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

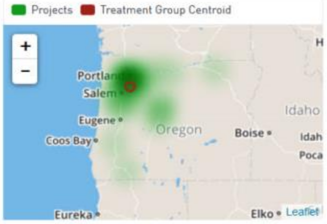
Two-Stage Approach



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



71.6 miles

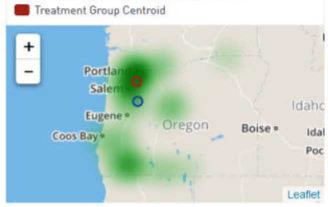
80% of projects lie within this distance from treatment group centroid

212	19,582
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



44.8 miles

Distance between treatment and comparison group centroids

1,043	19,635
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.

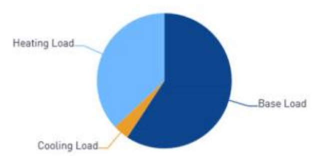
Future Participant Site Locations

Comparison Group Centroid / Treatment Group Centroid

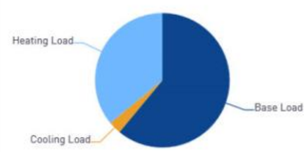
Distance between treatment and future participant group centroids

0	
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.

817	212	26%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

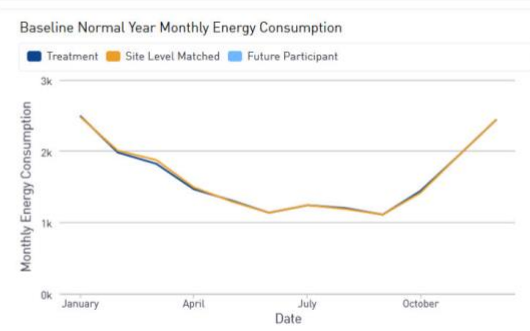
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	472	304
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	60	244
Meters with at least 5 site-level matched meters from the comparison group pool.	--	1	243
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	3	240

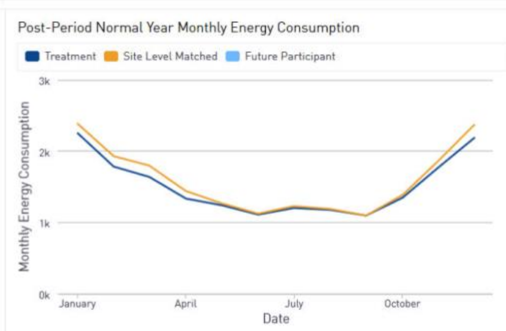
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	240
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	1	239
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	27	212
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	212
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	212
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	212
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	212

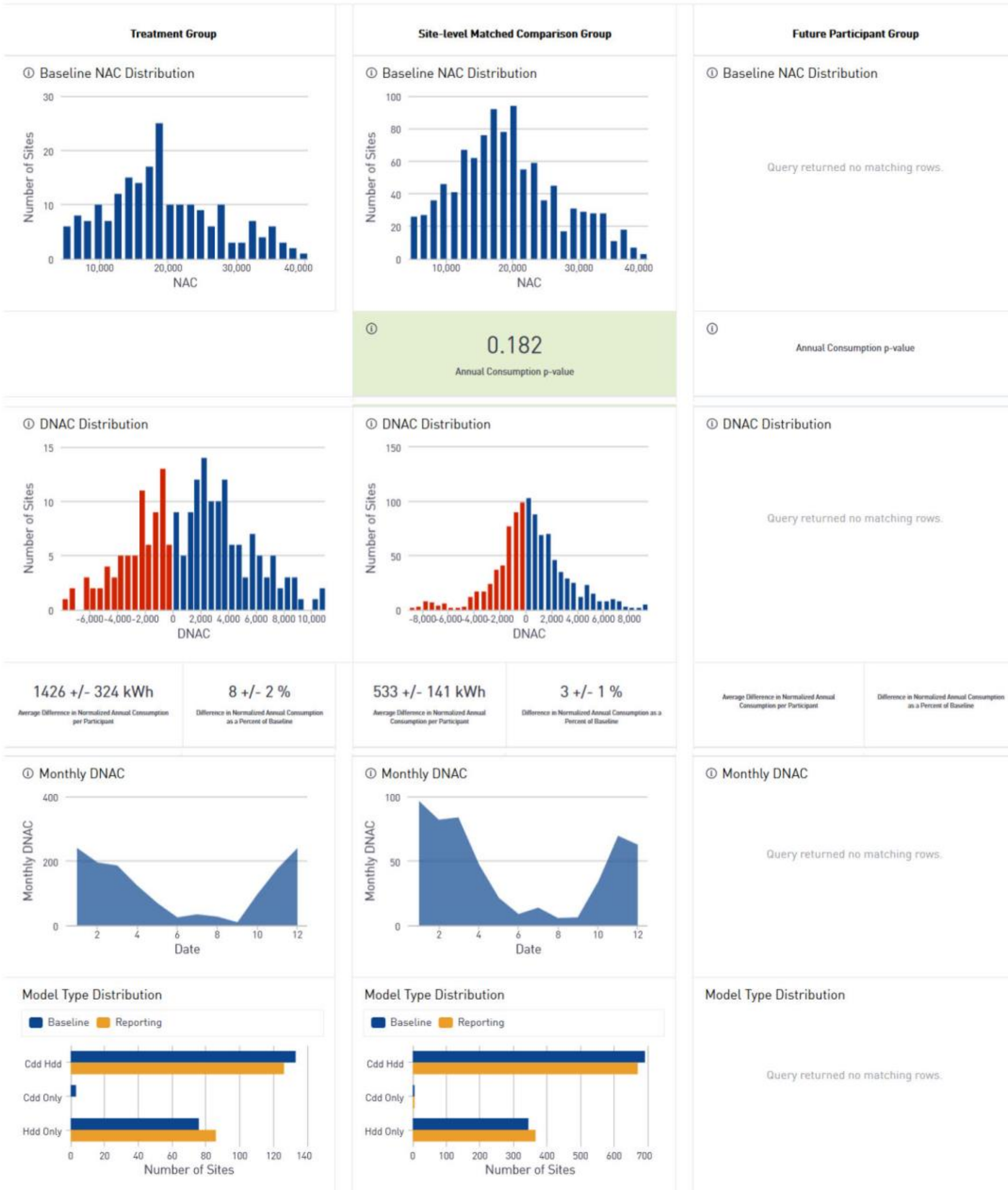
3. Modeling Results

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.





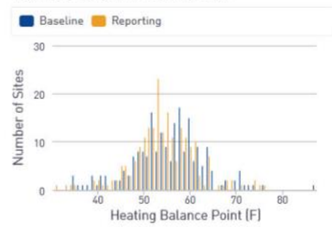


Report Date: March 2, 2020

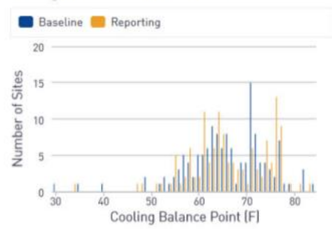
Treatment Group	Site-level Matched Comparison Group	Future Participant Group

Treatment Group

Heating Balance Point Distribution

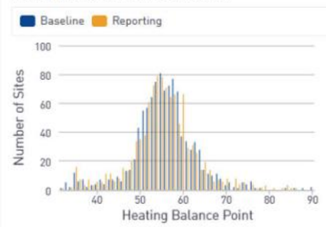


Cooling Balance Point Distribution

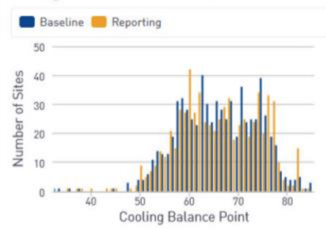


Site-level Matched Comparison Group

Heating Balance Point Distribution



Cooling Balance Point Distribution



Future Participant Group

Heating Balance Point Distribution

Query returned no matching rows.

Cooling Balance Point Distribution

Query returned no matching rows.

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

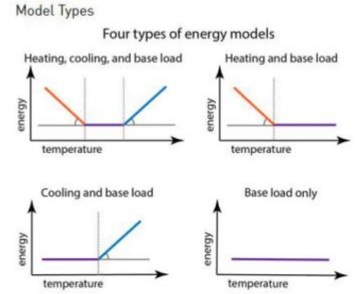
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

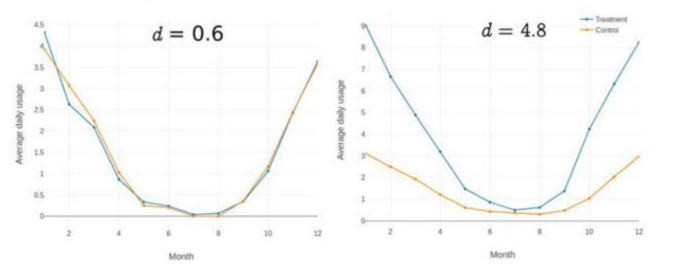
Stratified sampling is applied to future participant groups to attempt to replicate the distributions of the underlying variable (annual consumption) in the comparison group. Annual consumption of all treatment sites is first split into deciles, then a random sample is selected from within each corresponding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

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

Model Types



Euclidean Distance Matching



h. Other heat pump manufacturer electricity savings

Impact Evaluation Report

Electricity Impact of Extcapheatpump in Program Year 2014, 2015, 2016, 2017, 2018

Result Summary

Measure: Extcapheatpump	Program Year: 2014, 2015, 2016, 2017, 2018		Fuel: Electricity	
Meter Data Filters:	DNAC: <100%	DNAC Percentile: All	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 Zone(s): All	Heating Fuel: Electricity	Heat Pump Manufacturer: American Standard, Bryant, Coleman, Mitsubishi, Ruud
	Thermostat Name: All	Heat Pump Baseline: Eleboardht, Elebsbrd, Electfurnace, Electricfurnace, Elefurnace, Electric Baseboard Heat, Electric Furnace	Multi Measure Filter: No Filtering Based on Measures	Heat Pump Adv. Controls or Commissioning: All
	Air / Duct type: No Filtering Based on Air/duct Type	Home size: No Filtering Based on Home Size	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	

60 Treatment Meters	2925 +/- 711 kWh Average Normal Year Pre-Post Difference in Consumption per Participant	14 +/- 3 % Percent Normal Year Pre-Post Difference in Consumption per Participant	20,905 Mean Baseline Consumption (Electricity)	122% Realization Rate
294 Site-level Matched Meters	1953 +/- 755 kWh Average Savings Relative to Site-level Matched Comparison Group	9 +/- 4% Percent Savings Relative to Site-level Matched Comparison Group	19,892 Mean Baseline Consumption (Electricity)	81% Realization Rate
0 Future Participant Meters	Average Savings Relative to Future Participant Group	Savings Relative to Future Participant Group	Mean Baseline Consumption (Electricity)	Realization Rate

1 Report Date: March 2, 2020

RECURVE

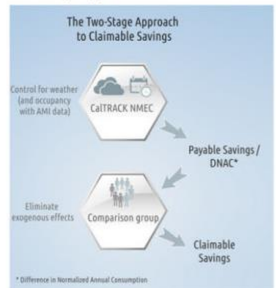
1. Introduction

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

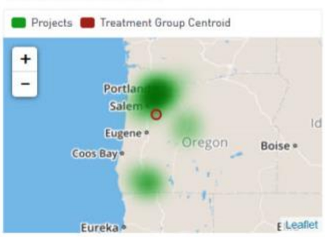


Treatment Group
The treatment group consists of sites that participated in the specified energy efficiency projects in the specified program year. Only sites that installed single measures are included in the treatment group. And this group includes the subset of sites that had sufficient data quality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

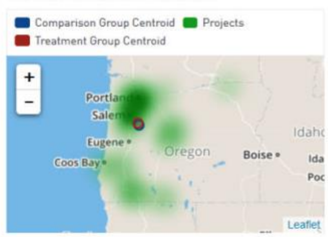
Site-level Matched Comparison Group
This group includes comparison group sites that were matched at the site-level to treatment group sites. Each treatment group site is matched to five comparison group sites from the same zipcode, but only the sites with sufficient data quality were included in the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

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.

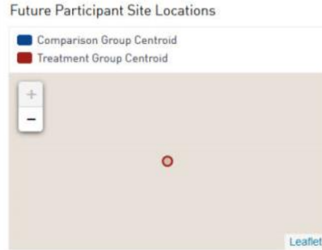
Treatment Site Locations



Site-level Matched Site Locations



Future Participant Site Locations



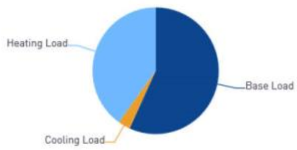
159.9 miles
80% of projects lie within this distance from treatment group centroid

10.9 miles
Distance between treatment and comparison group centroids

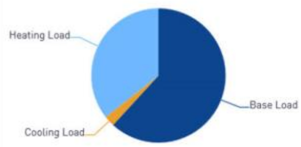
Distance between treatment and future participant group centroids

Group	Meters	Mean Baseline Consumption (Electricity)
Treatment Group	60	20,905
Site-level Matched Comparison Group	294	19,892
Future Participant Group	0	

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.

817	60	7.3%
Meters in Treatment Population	Final Sample Size	Percent of Treatment Population Represented by Sample

Sample Attrition Table

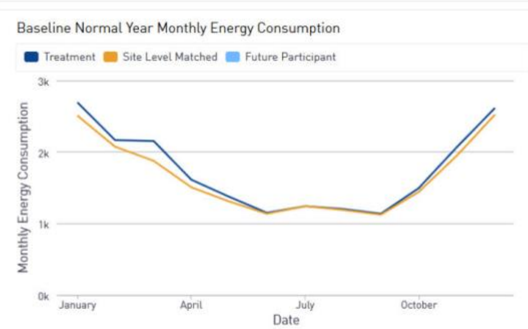
Filter	Selected Filter Value (if applicable)	Number of Dropped Meters	Sample Size after Applying Filter
Measure: Meters associated with a particular measure in program participation data. **Year:** Program year. **Fuel:** Type of metered fuel.	Measure: Extcapheatpump -- Year: 2014, 2015, 2016, 2017, 2018 -- Fuel: Electricity	--	817
Meters with valid consumption data in baseline and/or reporting periods.	--	32	785
MultiMeasure_Filter: Meters with single/multiple measure installations in baseline and/or reporting periods.	Multi Measure Filter: No Filtering Based on Measures	0	785
HeatingFuel: Meters with a valid heating fuel that corresponds to the selected filter value.	Heating Fuel: Electricity	9	776
HeatingZone, CoolingZone: Meters in selected heating and/or cooling climate zones.	Heating Zone: All -- Cooling Zone: All	0	776
Other measure-specific filters.	--	663	113
PeriodLength_Threshold: Meters meeting a threshold number of months of valid consumption data.	Period Length: 11 Months or Longer	36	77
Meters with at least 5 site-level matched meters from the comparison group pool.	--	2	75
DNAC_Threshold: Meters with normalized change in annual energy consumption under a specified threshold.	DNAC: <100%	2	73

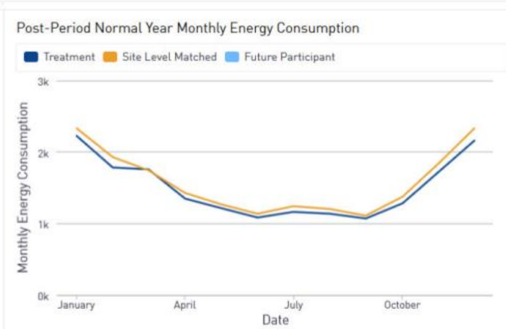
DNACPercentile_Threshold: Meters within specified percentile bands of normalized change in annual consumption.	DNAC Percentile: All	0	73
ConsumptionPercentile_Threshold: Meters within specified percentile bounds of annual energy consumption.	Annual Consumption Percentile: Remove Top and Bottom 0.5%	0	73
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	13	60
CVRMSE_Threshold: Meters with valid model CV(RMSE) for the baseline and reporting periods that meet a specified threshold.	CV(RMSE): < 1	0	60
home_size: Meters with manufactured home size meeting a specific criteria (single-wide, double-wide, or triple-wide).	Home Size: No Filtering Based on Home Size	0	60
complex_duct_sealing: Meters with the 'MH Complex Add-On' measure.	Complex Duct Sealing: No Filtering Based on Complex Duct Sealing	0	60
airduct_type: Meters that used specific measures relevant to Air and Duct Sealing programs.	Air/duct Type: No Filtering Based on Air/duct Type	0	60

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.





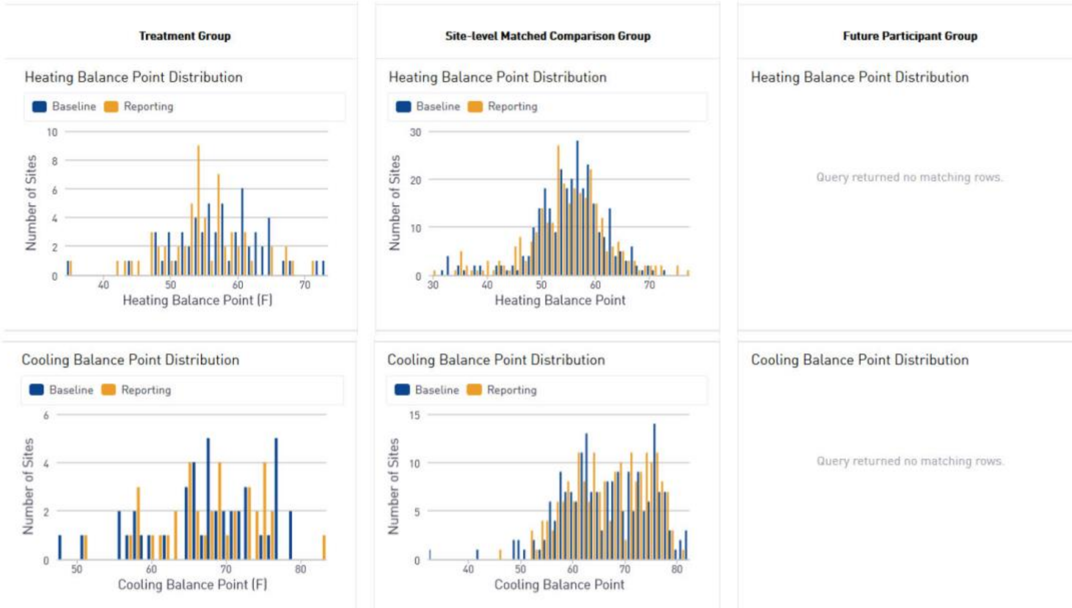
RECURVE

Report Date: March 2, 2020

5

RECURVE

Report Date: March 2, 2020



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.

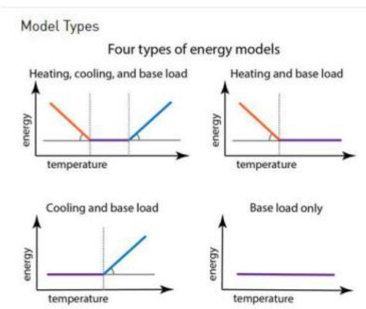
Comparison Group Generation

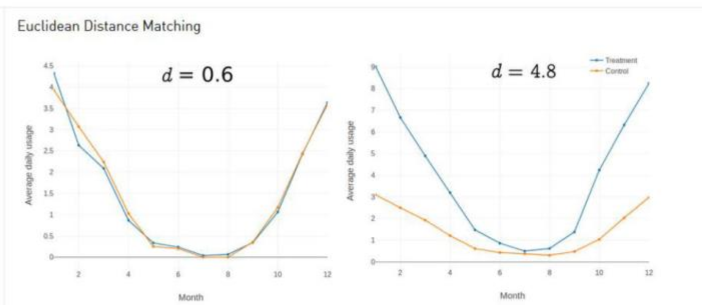
Site-level Matching: In monthly consumption matching, a comparison group is constructed by selecting 20 matches from the comparison group pool with the shortest distance d to the treatment group customer under consideration. After applying the selected filters on the comparison group, the comparison group is filtered down to the closest 5 matches to each treatment group member. 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 twenty 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.

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





RECURVE