March 02, 2015

REPORT #E15-306

HEAT PUMP WATER HEATER MODEL VALIDATION STUDY

Prepared by:

Ecotope Inc.
4056 9th Avenue NE
Seattle, WA 98105

Northwest Energy Efficiency Alliance
PHONE
503-688-5400
FAX
503-688-5447
EMAIL
info@neea.org

Acknowledgements:

We would like to acknowledge BPA, EPRI, and CLEAResult for collecting and sharing the results of their respective HPWH field studies. The addition of these sources to our own data greatly improved the quality and clarity of the results and insight contained herein.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Glossary of Acronyms and Abbreviations

AC	air conditioning
aCOP	average annual coefficient of performance
aMW	average megawatts
ANOVA	analysis of variance
ANCOVA	analysis of covariance
ATI	Air Tap Integrated
BPA	Bonneville Power Administration
Btu	British thermal unit
Btu/hr	British thermal unit per hour
CDD	cooling degree days
COP	coefficient of performance
DHW	domestic hot water
DOE	Department of Energy
EF	Energy Factor
EPRI	Electric Power Research Institute
ER	electric resistance
erCOP	electric resistance mode theoretical average annual coefficient of performance
ERWH	electric resistance water heater
ft	feet
GE	General Electric
HC_f	heating and cooling system interaction factor
HDD	heating degree days
hpCOP	heat pump system only average annual coefficient of performance
HPWH	heat pump water heater
HVAC	heating, ventilation, and air conditioning
HZ1	Pacific Northwest Heating Zone 1
HZ2	Pacific Northwest Heating Zone 2
HZ3	Pacific Northwest Heating Zone 3
ID	identification
kW	kilowatt
kWh	kilowatt hours
kWh/day	kilowatt hours per day
kWh/yr	kilowatt hours per year
n	number of observations
NEEA	Northwest Energy Efficiency Alliance
OAT	outdoor air temperature
QC	quality control
RBSA	Residential Building Stock Assessment
sq.ft.	square feet
TMY3	Typical Meteorological Year (based mostly on 1976-2005)
VBDD	variable base degree day
W	Watts

Executive Summary

A substantial fraction of the electricity consumed in the residential sector is used to heat water. Indeed, it is the second-largest single end use of electricity next to space heating. Due to the sheer size of the load there is substantial opportunity for energy use reductions. Recognizing the importance, heat pump water heaters (HPWHs) have been an integral component in Northwest power plans for nearly thirty years (NPCC 2010). The introduction of a new generation of integrated HPWHs, in conjunction with the Northern Climate Heat Pump Water Heater Specification, over the past five years, has enabled those savings to be realized (NEEA 2012a).

In 2012, the Northwest Energy Efficiency Alliance contracted with Evergreen Economics and Ecotope to measure water heater performance and validate estimates of the energy use and savings. The HPWH Savings Validation study was designed to integrate all previous work in the Northwest on HPWHs with the purpose of establishing a proven unit energy savings (UES) estimate for the Regional Technical Forum (RTF). This project comprehensively draws on laboratory studies and, importantly, two previously conducted field studies. It augments the field studies by measuring water heater performance in climate zones and installation locations not previously observed. 70 sites had been previously studied in the field and this project added 50 more.

Study Design

The HPWH Model Validation Study was conceived in a way to gather determinants of performance and then employ that information in numerical simulations to estimate energy use. As such, the field work conducted in the project was designed to aid in the development of a field calibrated engineering model. The field data provided observations of the independent variables of water heater energy use including, but not limited to, ambient space temperature, inlet water temperature, tank set point, and hot water draw pattern. The simulation-based approach offers several distinct advantages: it shortens the overall metering period and the number of sites required. There is significant diversity in installation characteristics, so to achieve field monitoring with statistical significance in each category would require an enormous and cost-prohibitive field study. Instead, the simulation-based approach allows us to collect the independent performance variables across the region, measure dependent variables (energy use) at selected sites, and then predict energy use for the myriad equipment types and installations. Therefore, the project approach is specifically designed to provide information for simulation purposes.

Analytic Methodology

Data on the two previous field studies were obtained from the Bonneville Power Administration (BPA) and NEEA. Together, they were integrated with the field data collected in the current study to create one, comprehensive dataset. The data were then inspected, curated, and cleaned. Anomalous data, reflecting events that we determined did not actually occur, were discarded.

To account for the non-annual basis of the Ecotope monitoring and the sporadic missing data inherent in all long-term field projects, we developed an annualization algorithm to adjust average values for seasonal effects. Water heating itself and many of its determinants vary seasonally, so annualization was necessary to make fair comparisons between sites that did not contain a contiguous year's worth of data.

Five sites were set aside for "flip-flop" testing to assess the space heating impact. In the flip-flop test, the HPWH was manually changed between heat pump mode and resistance mode. The idea was that the contrast between space heat with and without HPWH operation could be quantified through degree day regression.

Findings

Daily average flow was calculated as 23 gallons for a single occupant home, with an additional 11 gallons per day for each additional occupant. Mean energy use, normalized by flow, varied between 8 kWh/100 gallons and 13 kWh/100 gallons, depending on make and install location. Annual energy consumption of HPWHs typically ranged from 1,000 to 2,000 kWh per year.

We defined the quantity "aCOP", the average annual coefficient of performance, as the ratio of useful energy delivered to input energy, which assesses efficiency including degradations due to standby losses and resistance heat.[1] Across different combinations of HPWH make and install location, aCOP varied between 1.6 and 2.4, which represents a two- to three-fold increase in efficiency over a resistance tank. That improvement is for water heating energy alone and excludes any additional load placed on the space heating system from certain installation scenarios.

Space heating impacts were investigated through "flip flop" tests at five sites, where the HPWH was manually switched between heat pump mode and resistance mode. These tests proved inconclusive. In addition we explored the ambient space temperature depression during water heater operation. Those results show that the space heating impacts (and penalty) are less than 100%. That is, not every unit of energy removed by the HPWH from the inside air is replaced by the heating system. Exploratory analysis suggests the interaction should be no greater than 0.9 and a reasonable lower bound is likely 0.5. Further, the data suggest there is no noticeable interaction for garage and unheated basement installations.

Average delivered water temperature was roughly 124 °F. In tanks larger than 50 gallons, only about 1% of water was delivered below 105 °F. In 50 gallon tanks, 2.5% of water was delivered below 105 °F. In other words, HPWHs essentially always met the load demanded of them.

The detailed investigation of data quality – and the quantitative relationships within the dataset – revealed slow-developing performance anomalies with Air Generate ATI water heaters monitored in the previous NEEA study. The anomalous time periods were excluded from the dataset and the analysis proceeded only on those units that were clearly operating as designed.

---

[1] "aCOP" is an analogous quantity to the Energy Factor (EF) but we opt not to use the term "EF" because it is specifically defined under a set of lab test conditions. We use aCOP to maintain the distinction of results observed in the field.

Extended Findings

Inlet water temperature was modeled to vary linearly with a 7 week moving average of outdoor air temperature, with the elasticity of the change determined by water source (city surface, city ground/community, city mixed, and well).

Average setpoint across all units was measured at 128 °F. This value was not found to change with obvious factors like water heater make or number of occupants. For modeling purposes we assert that the mean setpoint should be set to 128 °F.

Intake air temperature profiles were modeled for each of the four installation scenarios. Garages and unheated basements are calculated based on fits to various outdoor temperature lags. The interior, non-ducted case, was modeled as an exponential temperature decline as the HPWH runs. For the interior, exhaust ducted cases, we determined that the house space temperature from a simulation can be used without modification.

Draw schedules for 1, 2, 3, 4, and 5+ occupants were crafted from the observed data. There are both typical day and typical week schedules. The typical week is most appropriate for simulation as it captures more of the variability inherent in hot water use. Each schedule is tuned to the observed average daily water draw per occupancy category. Within the schedule, the time, size, and duration of draws is informed by the field data themselves.

The generalized inputs, in combination with a validated simulation, succeeded in translating the findings in the engineered field sample to the population of houses at large. The simulation output echoed the findings of the field study but in a way usable for general estimates. The simulation output produced unit energy savings estimates adopted by the RTF.

Conclusions

Overall, the study provided the necessary field observations of the independent determinants of HPWH energy use in order to predict their behavior with confidence across the general population of houses in the Northwest. The energy and performance measurements show the integrated HPWHs can deliver energy use one-half to one-third below the base case resistance tanks. The findings also confirm there are differences in energy use between Tier 1 and Tier 2 Northern Climate Specification tanks. Certainly, the Tier 1 tanks, due to more resistance heat use, require more energy. The increased resistance heat use comes from control strategy differences and from the inability of Tier 1 tanks to operate their refrigeration cycle below 45 °F. The low temperature cut out is largely only of concern for garage installations but it significantly tempers energy savings there. In the end, the project data and simulation were used to update the unit energy savings estimate at the RTF and produced a validated simulation to be used in future estimates of HPWH energy savings.

1. Introduction

In the Pacific Northwest, 55% of the 4 million households use electricity to heat domestic hot water (DHW) (Baylon 2012). The dominant technology consists of electric resistance elements in an insulated tank. For the average household with an electric tank water heater of 2.57 occupants, that technology uses 3,227 kWh/yr to heat water (Ecotope 2014). This end use alone represents a significant load (approximately 850 aMW) on the electric power system. The traditional method to reduce energy use has been to require more tank insulation. Such improvements have been federally mandated since 1990 (DOE 2014).

Reductions in energy use through insulation alone are limited, however, because they only reduce standby losses and do not change the way in which the water is heated. As such, substantial gains in electric water heater efficiency are only available with different heating methods. Indeed, heat pump water heaters (HPWH) have been a significant, potential, conservation measure in nearly every Power Plan since 1986 (NPCC 2010). Heat pump water heaters are able to heat the water 2-4 times more efficiently than traditional systems. Assuming HPWHs are twice as efficient as the traditional systems, they can save at least 425 aMW. Nevertheless, the actual technology and availability has not, until recently, risen to the level of promise the energy savings would indicate.

Beginning with efforts in the 1980s, various manufacturers have introduced heat pump technologies to meet DHW demand (Hanford 1985). In several of those efforts, the technical and/or market challenges proved insurmountable. In 2008, the national EnergyStar program announced a labeling specification for electric water heaters which helped prompt the development of a new generation of heat pump water heaters (NEEA 2012b). Several manufactures, including AO Smith, Rheem, General Electric, and AirGenerate introduced EnergyStar qualified water heaters beginning in 2009. Concurrently, NEEA launched the Northern Climate Specification (NC Spec) for residential heat pump water heaters setting out important technical criteria for the successful application of HPWHs in cold climates.

The combined availability of promising product and the NC Spec ushered in a round of evaluation and testing beginning with a lab investigation of three HPWHs in 2009 (Larson 2011). Subsequently, the Bonneville Power Administration (BPA) engaged the Electric Power Research Institute (EPRI) to measure the energy and water use of those same three HPWH models in approximately 40 residences across the Northwest (Bedney 2012, BPA 2012). Over a similar time period, and continuing to the present, NEEA conducted laboratory investigations of HPWH to understand their performance (Larson 2012a, 2012b, 2012c, 2013a, 2013b, 2013c). Likewise, using the lab and field data available at that time, the Regional Technical Forum (RTF) accepted a provisional unit energy savings estimate for HPWH installations replacing traditional electric resistance water heaters (ERWH) (RTF 2011). Simultaneously, NEEA launched a 30 site field study of Tier 2 Northern Climate units (Fluid 2013).

In 2012, the Northwest utilities embarked on a project to move the provisional savings estimate of HPWHs to a more solid, proven number. To do so, NEEA contracted with Evergreen Economics and Ecotope to conduct this project, the HPWH Model Validation Study. Per the research plan, developed in conjunction with the RTF, the project was designed to integrate all the HPWH projects in the Northwest drawing from both lab and field sources. The result would be a comprehensive study of water heater behavior and energy use.

Under the first stage of the HPWH Model Validation Study, Evergreen Economics conducted a market test assessment which assessed the supply- and demand-side market acceptance of HPWHs in the Northwest and assessed the implementation strategy of NEEA's Heat Pump Water Heater Market Test (Evergreen 2013).[2] In the second stage, Ecotope researched all the necessary components to provide a reliable estimate of HPWH energy use and savings in the Northwest.

1.1. Goals and Objectives

The main goal of the second phase of the Validation Study, and this report, is to establish a proven unit energy savings (UES) estimate for heat pump water heaters in the Northwest. The project does so in a comprehensive way by drawing on laboratory studies, multiple field studies, and a calibrated simulation of water heater performance. Accordingly, this report includes an accounting of the field metering methodology, results of the metering including characterization and performance data, integrated analysis on the full Northwest HPWH field dataset, a complete discussion of the engineering model validation, and a final estimate of energy savings. The report is a project reference document and serves as the basis for proposing a unit energy savings number to the RTF.

To accurately estimate heat pump water heater energy use across houses in the Northwest, it is necessary to understand their behavior well enough to predict performance under a wide variety of operating conditions and installation scenarios. The operating conditions span a range of ambient air temperatures, inlet water temperatures, and occupant hot water use patterns. Installation scenarios span the range from conceptually simple garage locations to complex configurations involving placement inside a conditioned house, heated by a heat pump, with the HPWH exhaust air ducted outside. Consequently, to assess all possible operating conditions and installation configurations, we turn to software simulations, supported by field measurements. Restated, the goal of field metering heat pump water heaters is to quantify all of the independent performance variables in enough detail to predict energy use (the main dependent variable) for all installation types in the Northwest.

---

[2] Refer to NEEA website for the report: http://neea.org/docs/default-source/reports/northwest-heat-pump-water-heater-market-test-assessment.pdf?sfvrsn=6

2. Methodology

The Validation study was designed to support the main project goal of establishing a unit energy savings estimate for HPWHs. Broadly, the project needed to gather information on the independent variables that determine the performance of the HPWH in a way that could be employed to inform numerical simulations of energy use. In addition, the field work and field data collected in this project, and the laboratory work leveraged by the earlier phases of the project, was designed to serve the needs and purposes of the engineering calculations and simulations. To that end, the field measurements quantify the independent variables which determine HPWH performance, including:

- Environmental Variables – those dependent on installation type, location, and climate
  - Ambient conditions of the space where the water heater is installed
  - Cold water inlet temperature
- User Variables – those dependent on the house occupant over time
  - Hot water draw pattern
  - Tank setpoint temperature
  - Operating mode selection
- Equipment Variables – those dependent on qualities of the particular tank installed
  - Tank storage volume
  - Heating component output capacity
  - Control strategies
  - Tank heat loss rate
  - Heat pump efficiency (over a range of given environmental conditions)

To support the objective of developing a credible software simulation, the project collected enough measured performance data to validate and tune any such simulation so its predictive accuracy can be improved and quantified. The simulation approach offers several distinct advantages: it shortens the overall metering period and the number of sites required. There is significant diversity in installation characteristics, so to achieve field monitoring results with statistical significance in each category would require an enormous and cost-prohibitive field study. Instead, the simulation-based approach allows us to collect the independent performance variables across the region, measure dependent variables (energy use) at selected sites, and then predict energy use for the myriad equipment types and installations. Therefore, the project approach is specifically designed to provide information for simulation purposes.

2.1. Study Design

HPWH performance depends on a number of environmental, user, and equipment variables. The environmental variables are, in most ways, pre-determined by the climate zone where the house is located and the location within the house where the HPWH is installed. The user variables present a greater source of variability, as draw patterns between similarly sized households typically show much more variation than accompanying environmental variables such as inlet water or ambient temperature. Many of the equipment variables have also been measured in the lab in previous studies. Consequently, the objective of site selection is to construct a population across the broadest possible range of environmental variables (to observe climate variation) while simultaneously spanning a long enough time and diverse enough user base to capture variation in draw patterns.

Previous projects, funded by BPA and NEEA measured field performance at approximately 40 and 30 sites respectively (Bedney 2012, BPA 2012, Fluid 2013). Those sites comprised both Tier 1 and Tier 2 HPWHs. To leverage the existing funding and datasets, Ecotope planned to use the existing projects and expand on the climate and installation scenarios to complete our understanding of HPWH performance in diverse applications.

The site selection process is an engineered selection designed to cover all of the independent environmental and equipment variables, as opposed to a random sample of the population. That is, our site selection focused on collecting metered data across heating climate zones, installation locations, and tank types, regardless of natural HPWH geographical and installation configuration distribution. This allows us to build models of HPWH performance, which can then be overlaid on maps of the distribution of housing types and potential installation options to estimate expected region wide energy savings. The existing BPA and NEEA field studies already cover many of the important cells in this sample design (based on geography, HPWH, and installation characteristics). As a result, the units and geographies in this study could be limited to the coverage needed to meet the study design goals of the RTF.

2.1.1. Engineered Site Selection and Recruiting

Ecotope presented a preliminary, engineered sample design at the RTF in August, 2011 consisting of 165 sites (Larson 2011a). In collaboration with the RTF HPWH Evaluation Subcommittee[3], NEEA, and other stakeholders, Ecotope refined this sampling plan to optimize and, ultimately, reduce the number of sites required. The keys to refining the sampling plan included achieving enough diversity across climates, installation location, HPWH performance tier, and tank size with the minimum number of field sites. Ecotope leveraged the existing field studies conducted by BPA and NEEA, which encompass approximately 70 installations. Those sites were then compared against the sampling plan. The remaining sites then formed the basis of our metering group.

Important influences on the sample design include recruiting enough sites in colder climates to measure how HPWHs perform in these challenging conditions. Further, the sample was stratified in such a way as to directly feed into the performance validation model. The recruiting plan was driven by the need to fill in the areas that were not covered by the previous two studies. This led to a matrix covering geographies and installation types totaling 50 sites.

Table 1 shows the matrix considering installation configuration, NC Spec Tier, heating zone, and specific equipment type. HZ1, HZ2, HZ3 correspond to the Pacific Northwest heating zones

---

[3] http://rtf.nwcouncil.org/subcommittees/hpwh/

defined by the Northwest Power & Conservation Council.[4] The climates range from mild heating requirements in HZ1 to substantial heating requirements in HZ3. The installation configurations are defined as follows:

- Garage – an unheated space attached to the house. As such, essentially no water heaters are installed in those spaces in colder HZ3.
- Unheated Basement – does not have a positive heat supply but it may have heat gains contributed from furnaces, appliances, ductwork, and conduction from the first floor of the house.
- Interior – inside a house's temperature controlled space where the HPWH exhausts air directly to that space. Locations typically included utility and laundry rooms.
- Interior ducted – one inside a house's conditioned space but the exhaust air is ducted outside the conditioned envelope.

The remaining strata are on NC Spec tier and HPWH equipment type. The GeoSpring is a 50 gallon tank from General Electric (GE). The Voltex 60 and 80 gallon tanks are from AO Smith. All three qualify for Tier 1 under the Northern Climate Specification. The ATI 66 gallon is a Tier 2 qualified product from AirGenerate. When the project was conceived there were no other makes and models available that qualified under the NC Spec.

Table 1. HPWHs Installed under Ecotope Portion of Validation Project

Installation Location	Tier 1						Tier 2		
	GeoSpring 50 gallon			Voltex 60 or 80 gallon			ATI 66 gallon		
	HZ1	HZ2	HZ3	HZ1	HZ2	HZ3	HZ1	HZ2	HZ3
Garage	0	2	0	5	2	0	0	1	0
Unheated Basement	3	2	1	4	3	1	0	0	2
Interior	0	2	1	1	6	1	0	0	0
Interior Ducted	0	0	0	0	0	0	3	1	9

Neither household size nor other demographics were considered in the sampling strata. However, household size did have some impact in that it drove the tank size selection – larger occupancy counts needed larger tanks.

Initial project plans called for recruiting participants who already had a HPWH installed. However, in quarter three of 2012, when the recruiting began, few houses had water heaters installed in the configurations required by the sample plan. Consequently, the project opted to purchase and install the necessary HPWH for each house.

In order to reach these specific targets, Ecotope called about 250 sites and inspected 109 potential participant houses. Some participants were drawn from the Residential Building Stock Assessment survey; others were drawn from the pool of people who had installed ductless heat pumps in their residences. Many sites failed the initial inspection for reasons that were not identified in the initial phone interview. These reasons included the size of the space around the

---

[4] See "Weighting Factors" worksheet in this workbook:
http://rtf.nwcouncil.org/measures/support/files/RTFStandardInformationWorkbook_v2_0.xlsx

water heater, an unworkable location of the electrical panel, and insufficient space in the electrical panel. These initial inspections were extremely effective in filtering out unworkable sites and led to significant cost savings – the cost of scheduling a site manager, plumber and electrician to show up at a site only to have to walk away is considerable.

2.1.2. Integrated Field Site Disposition

An essential step in this project was to collect and integrate all the field data on HPWHs that has been measured in the Northwest. Ecotope worked closely with the funders and contractors of the other studies (BPA and EPRI, NEEA and Fluid/CLEAResult) to obtain and understand the existing measurements. After collecting the data, Ecotope subjected all of it to the same rigorous quality controls described in section 2.3.1. Ultimately, Ecotope assembled a cleaned, quality-controlled working dataset of 107 sites with characteristics shown in Table 2.

Table 2. Comprehensive Northwest HPWH Field Site Locations

Equipment	Climate Zone	Installation Location				
		Basement	Garage	Interior	Interior Ducted	Total
Voltex 60 & 80 Gallon	**HZ1**	4	9	1	0	14
	HZ2	2	4	6	1	13
	HZ3	0	0	1	0	1
	All	6	13	8	1	28
ATI 66 gallon	**HZ1**	3	8	0	12	23
	HZ2	0	5	0	11	16
	HZ3	0	0	0	7	7
	All	3	13	0	30	46
GeoSpring 50 gallon	**HZ1**	4	16	2	0	22
	HZ2	3	2	2	0	7
	HZ3	2	0	2	0	4
	All	9	18	6	0	33
Total	**All**	18	44	14	31	107

Included in other studies, but excluded from this dataset were two sites with Daikin Altherma, split-system, heat pump water heaters, and eight sites with the first generation of a Rheem 50 gallon HPWH. The Rheem HPWH turned out not to qualify for either Tier 1 or Tier 2 status. The more recent model Rheem, however, does (Larson 2013b). This project focuses on integrated heat pump water heaters meeting either Tier 1 or 2 of the Northern Climate Specification. All other types of heat pump water heaters were excluded from the analysis.

2.2. Field Measurement

Ecotope's field measurements of HPWHs are discussed below. The BPA and other NEEA projects generally followed similar plans and protocols making it possible to eventually meld all three datasets. For specifics on the BPA and other NEEA projects, refer to their reports (Bedney 2012, BPA 2012, Fluid 2013).

2.2.1. HPWH and Datalogger Installation

The HPWH was installed onsite along with the datalogging package. The datalogging package consisted of a datalogger, water temperature probes, a water flow meter, current transformers and temperature sensors. The installation protocol is described in detail in Appendix A: Metering Protocol. The onsite team consisted of the site manager, a plumber and an electrician. The team installed the HPWH, performed a full house audit, installed the datalogger and temperature probes close to the HPWH, and installed the current transformers in the electrical panel and connected them to a power transducer which also measured line voltage. The output from the power transducer was connected to the datalogger.

Ecotope deployed meters that recorded the data needed for this project at 1-minute intervals. In previous field projects, Ecotope has found that 5-minute intervals are sufficient for logging most household energy uses but the need to observe water flow events, often lasting only seconds, required 1-minute logging (Baylon 2012b, Ecotope 2014). Temperature data were 1-minute averages while water flow and electricity usage were accumulated every minute. Electricity use channels recorded true power, true energy, and power factor. Ecotope also collected data on outdoor temperature, hot water inlet and outlet temperature and intake and exhaust air temperature.

Ecotope deployed a metering network that reliably and accurately aggregated metered data for transmission to Ecotope servers. This system was self-contained and did not depend on the homeowner’s internet connection or assistance. Dataloggers uploaded accumulated data on a regular schedule (approximately every 6 hours) to the manufacturer’s internet servers and cached data locally for remote retrieval if needed. Ecotope retrieved this data on a daily basis. Dataloggers had the capacity to store at least three weeks’ data between transmissions. This meant that no data were lost due to network outages.

2.2.2. Data Management

The data collection for the metered sites started after the first sites were installed. Major data collection activities included: automated daily download for power, water flow, outdoor temperature and HPWH temperature readings; and ongoing monitoring and troubleshooting of remotely downloaded data

Data were uploaded securely to Ecotope’s servers. Log files were also stored on the dataloggers onsite to protect against data loss. The dataloggers were monitored remotely to quickly detect any problematic equipment. Data were also processed through a daily automated check to ensure that meters were reporting realistic numbers. Site visits were performed as necessary to effect repairs.

2.2.2.1. Data Quality Monitoring

As each site was added to the study, Ecotope evaluated the data streams coming from the site. Ecotope developed an automated script to perform data quality checks to ensure and verify the accuracy and completeness of data. This script ensured that data were flowing and that all readings were within reasonable ranges. This script also generated a daily report that allowed

Ecotope to rapidly identify and address malfunctioning equipment (both dataloggers and the HPWH itself).

2.3. Analytic Methods

2.3.1. Data Quality Control

Data quality is of primary concern. Elaborate methodology and well-laid plans are irrelevant if the data are mishandled or incorrectly collected. In addition to the ongoing data quality monitoring procedures for the data collected by Ecotope, we applied a number of post-processing filters to the data across all three field studies. These techniques were deployed sequentially to trace down and provide fixes to observed data anomalies:

1. Summary graphics, showing the entirety of data for a given site.
2. Comparison graphics, in which energy is plotted against flow for a similar class of sites, for example all 80 gallon Voltex water heaters, or all garage installs.
3. Output from a diagnostic regression model.

Figure 1 shows an example of the first method – a summary graphic showing the entirety of the data for a given site. This graphic shows data for a site outside Prineville, Oregon in the prior NEEA study. It is obvious from the wild oscillations of the lavender outlet water temperature line and the slate green inlet water temperature line that something unusual occurred during the summer. As this house is located rurally in the desert, we believe that the measured water temperature oscillations were caused by pressure changes in the water system during lawn irrigation, and the lack of a proper check valve in-line with the water temperature measurements. Water flow measurements of this kind were particularly vexing across all three studies. After diagnosing the problem early on in our field data collection, Ecotope installed check valves to guarantee on-way flow at the meter so as to not over-count water use.

Figure 1. Example of Data Quality Method Number One.

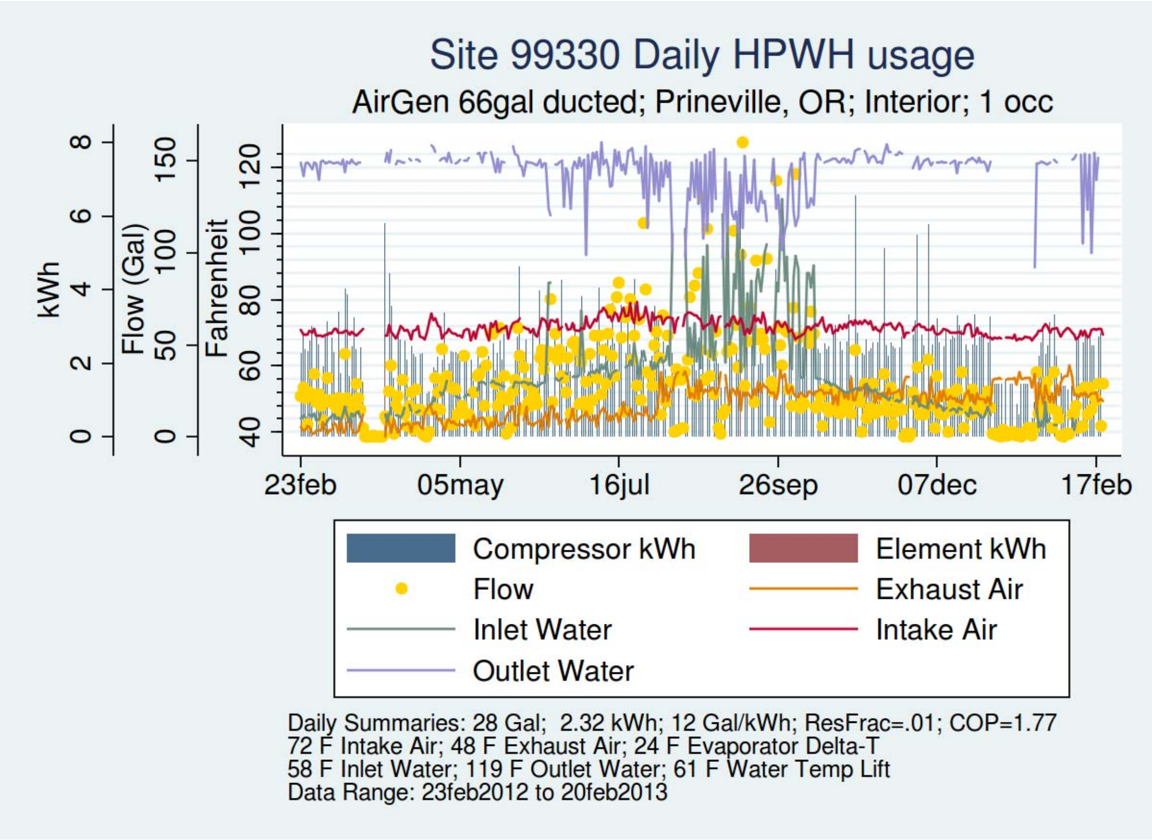
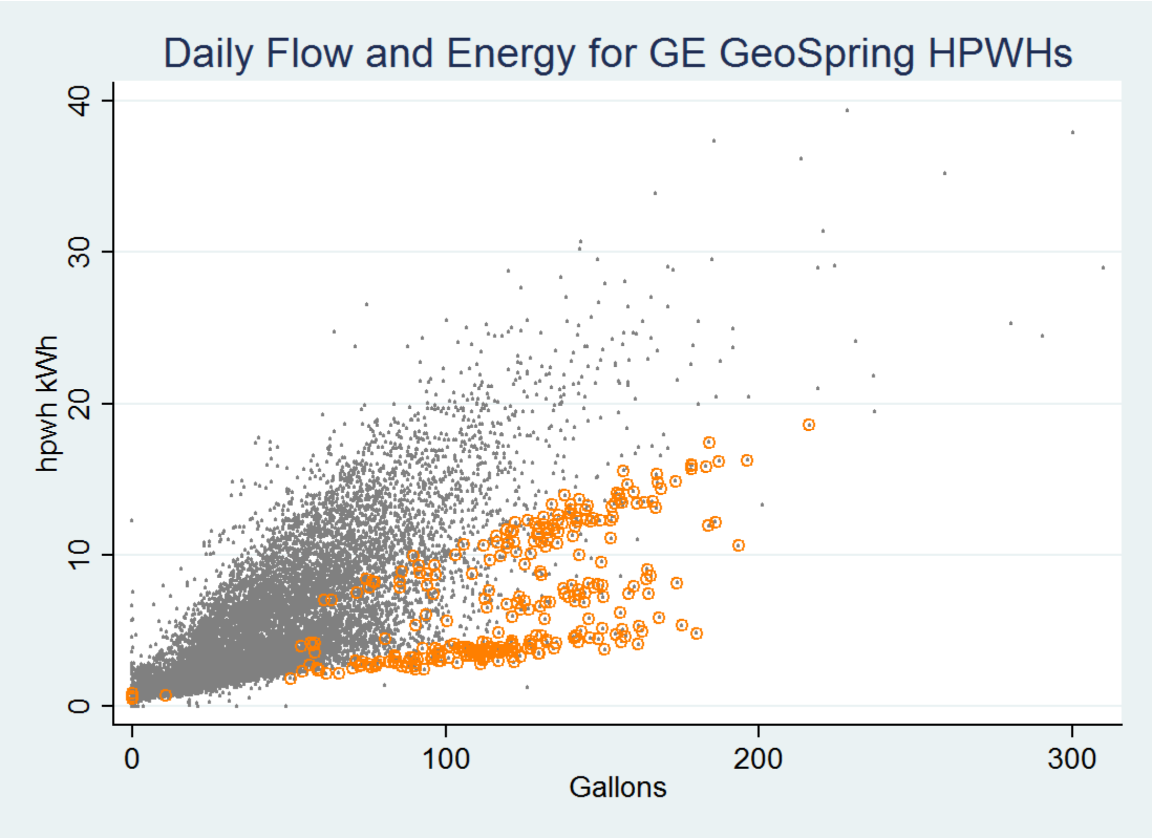


Figure 2 shows an example of data quality method number two. The plot shows daily flow and energy for GeoSpring water heaters, with site 99114, a GeoSpring from the BPA study, highlighted in orange. It is evident that, for a subset of the observed days, site 99114 falls outside the bounds of credible GeoSpring water heater operation. This problem was traced back to an incorrect flow measurement.

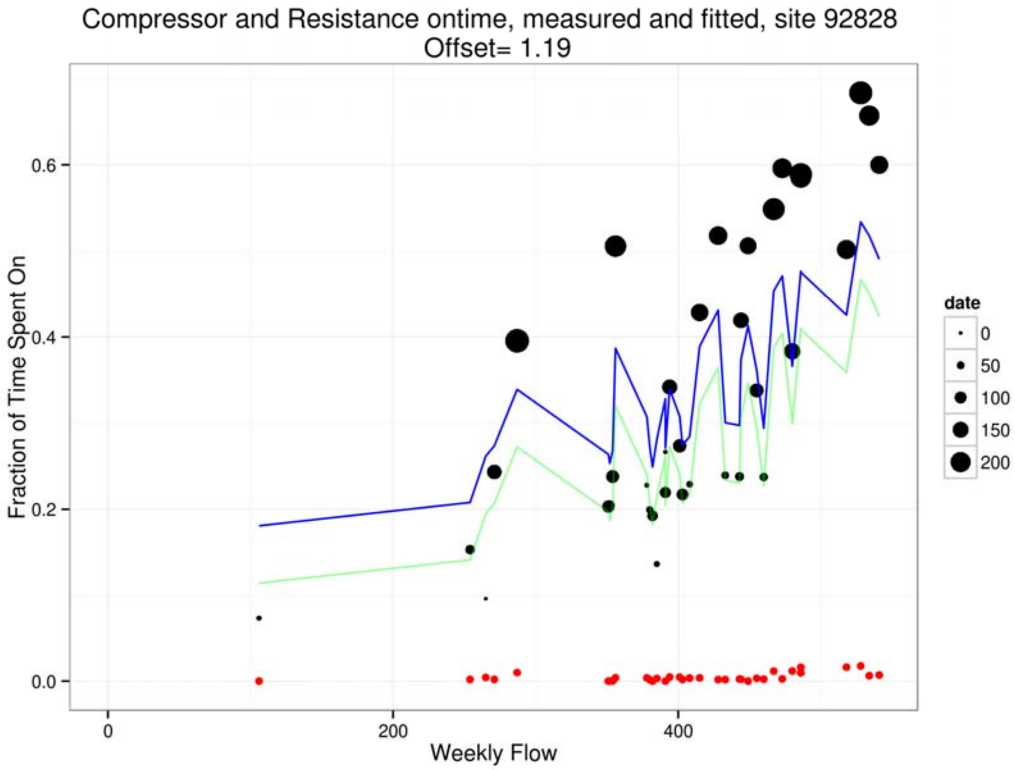
Figure 2. Example of Data Quality Method Number Two.



Most data problems were detected through a combination of those two procedures: looking at the entirety of the data for a given site, or viewing a quick comparison of the efficiency for a given unit compared to other units of the same make. However, some lingering anomalies prompted the development of a diagnostic regression model.

The basic idea behind the diagnostic regression model was that, conditional on all other aspects of heat pump operation, the amount of on-time should be predictable. Given the hot water draw, inlet water temperature, intake air temperature, tank setpoint, amount of resistance heat, make, and install parameters, one should be able to “predict" with great accuracy the amount of heat pump runtime needed to meet that demand. Note that this model is necessarily diagnostic for data previously observed and cannot be used for out-of-sample predictions – or UES estimates – because it requires conditioning on all data points save one: heat pump on-time. It is, however, a powerful tool to identify incorrect data or malfunctioning heat pumps. More details about the diagnostic regression model can be found in Appendix C: Diagnostic Regression Model.

Figure 3. Example of Data Quality Method Number Three – Diagnostic Regression Model.



The unusual looking graphic of Figure 3 shows output from the diagnostic regression model, data quality method number three. Of all the plots used in quality control, these are probably the most difficult to initially interpret, but reveal the most subtle indications of malfunctioning data loggers or water heaters. The x-axis shows weekly water draw in gallons, and the y-axis denotes fraction of on-time: black circles for heat pump on-time and red circles for resistance element on-time. The green line represents the average prediction across all water heaters of that make, and the blue line represents the predictions specific to the current unit, in this case an ATI at Ecotope site 92828. The size of the black bubbles is proportional to elapsed time since the start of the study. In the plot, the large black bubbles exceed the prediction line, which indicates a loss of performance later in the monitoring period. This unit was eventually diagnosed as having gradually lost refrigerant charge.

2.3.2. Annualizing Measurements

Not all sites from all studies yielded contiguous, one-year regions of usable data. Since water heating is a seasonal load, a fair comparison between units, sites, and conditions must be done under some sort of annualizing algorithm. Annualizing summaries of the data is especially important with a heat pump water heater, as, in addition to seasonally changing inlet water temperatures, seasonally changing intake air temperatures can alter heat pump efficiency, or necessitate the use of supplementary resistance heat.

Note that annualized data is only a concern in that we would like to descriptively view tables of the data. The annualized data procedure is not directly related to the calibrated engineering approach for developing the proven UES. It is only relevant to view and discuss the measured performance characteristics; even though we will ultimately use the data to develop and calibrate the simulation, it is still worthwhile to view and discuss what actually occurred in the three studies.

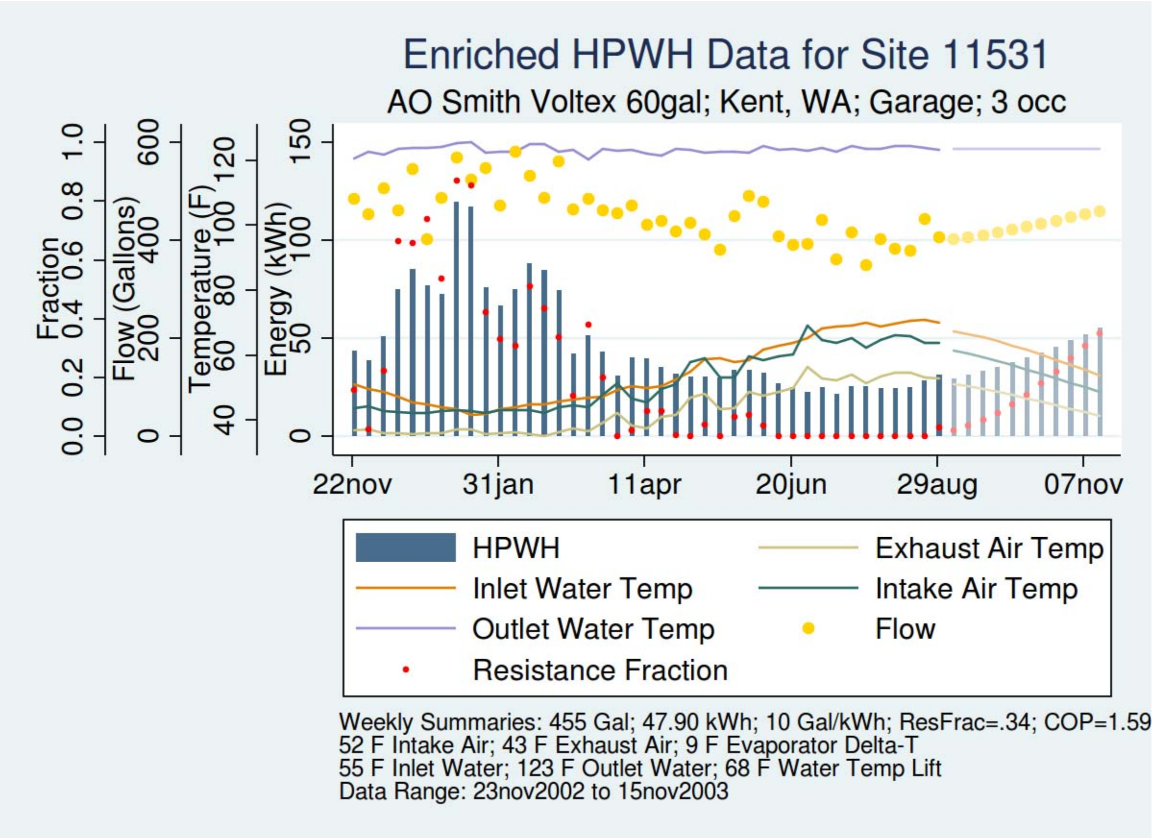
The BPA study observed sites for over a year with an average span of 470 days. The prior NEEA study monitored locations for exactly one year. The Ecotope study observed the water heaters for an average of 240 days each – less than a year for all units.

The following measurements were annualized for the purposes of the tables in this section: flow, intake air temperature, inlet water temperature, heat pump water heater total input energy, and fraction of heat delivered through resistance elements.

The basic idea behind the annualizing method was, for any given site, to "borrow" from the entirety of the dataset to produce a smooth and reasonable prediction of what would have happened had we observed the entire year. Note that this is a poor substitute for actually observing the data in a classic statistical sample study, but for the purpose of tabulating summary statistics, it is adequate. Refer to Appendix B: Details of Data Annualization for a more details.

Figure 4 shows observed and enriched (annualized) data for Ecotope site 11531. The actual monitoring ended in early September 2013, but to make fair comparisons of summary statistics we need annual averages, and so the data were supplemented with predictions from the regression model. In the figure, solid colors indicate real data, and faint colors indicate predicted data used to annualize the HPWH performance.

Figure 4. Observed and Enriched Data for Ecotope Site 11531.



2.3.3. Heating System Interaction

By virtue of their design, HPWHs extract heat from the ambient air surrounding them. Depending on the configuration, that cooler air is either recirculated to the surrounding space or exhausted outside the house. In the former case, there is a local cooling effect. In the latter, there is a change to the house infiltration. In either case, the result is an interaction with the HVAC system. The interaction is expected to be most pronounced for interior installations but it can theoretically still exist for garage and basement configurations. For example, the HPWH has the potential to lower the space temperature in the garage which, in turn, increases the rate of heat conduction through the house-garage walls. In the heating season, the interaction manifests as an heating energy penalty. While in cooling, the HPWH provides an energy (or comfort) bonus. In mild weather, with the heating or cooling system off, it has no impact on the house space conditioning load.

To explore the interaction, we define the quantity, HC_f, on a scale of 0 to 1, to be the heating and cooling interaction factor. A value of 1 means that every unit of energy extracted from the ambient air is replaced by a unit of energy from the heating system. Likewise, in cooling, a value of 1 means that the cooling system benefits 100% from the energy removed. Due to the climates under study in the Pacific Northwest, we often refer to the heating aspects only but there is little reason to suspect the framework for cooling would differ.

We explored several methods to ascertain the heating system interaction. They included "flip-flop" testing and monitoring the ambient space temperature depression. With both these methods we did not conclusively measure a heating interaction. Ultimately, we resort to basic principles of energy balance and heat transfer to guide engineering experience and judgment.

2.3.4. Hot Water Draw Profile Development

The main objective in collecting hot water draw data was to create representative profiles to be used in water heater simulations. At its most basic, the draw profile is sometimes thought of as an average daily volume assumed to be constant over the year. Field and lab studies have shown, however, that the control strategies employed by the HPWHs respond differently to different draw patterns. In other words, given the same daily total amount, the exact time and amount of water use across multiple days can dramatically change equipment efficiency. Consequently, if we expect a simulation to produce reasonable performance estimates, it is important to develop typical draw profiles based in the data we observed.

Ecotope's extensive literature review showed that no existing work could be directly used in developing representative draw profiles. Numerous studies have attempted to quantify draw profiles for various purposes (Perlman 1985, Becker 1990, Fairey 2004, Lutz 2006, Hendron 2008, Lutz 2012). The goals of some were to determine an average residential draw shape while others quantified the water volume, number of draws, and time between draws on a daily basis. The work of Lutz is the most comprehensive, producing summaries from datasets across the country of daily volume, daily draws, flow rates, time of recovery intervals and draw duration (2012). The report also created three summary groups based on clustering median daily hot water use but did not report the number of occupants for those clusters. For the purposes of simulating water use across a housing population it is necessary to group characteristics based on the number of occupants. The work of Hendron (2008) produced the most readily useable results for simulation purposes. Those draw schedules, however, were designed for use over an entire year and differed day-to-day. Ideally, for simplifying the simulation, the same draw profile would repeat daily or, at most, weekly. Thus, we devised our own method to identify and build typical draw patterns.

This method employs a descriptive characteristics approach to describe both daily and weekly draw patterns. Recognizing that hot water use is driven by occupants, we set out to characterize draw patterns in terms of 1, 2, 3, 4, and 5+ occupants. There were not enough households with 5 or more occupants to categorize larger occupancies separately. Within each occupancy count, we examined the days and weeks for clusters of draws, total draw volume, the number of small, medium, and large draws, and the average size of the small, medium, and large draws. Small hot water draws were defined as 1-2 gallons, medium as 3-9 gallons, and large as 10+ gallons. See Appendix E: Draw Profiles for a detailed explanation.

We analyzed the data to determine a typical day and a typical week. Due to the variation in draws, we realized that creating only a 24-hour long draw pattern was not enough. In order to capture the variation from small to large daily usage, a full seven-day time span was needed. Put another way, a single day of only small to medium sized draws would never cause the HPWH's resistance elements to engage. Likewise, a single day of only large draws would trigger the resistance elements an undue amount. The two scenarios neither represent what we observed in the field nor accurately estimate annual energy use. Consequently, we turn to a full seven days instead of a single 24-hour period. Because daily draws are integral to current water heater ratings (EF) and testing we have presented both typical days and weeks in the report.

3. Findings

The findings based on the observed field measurements form the basis for the rest of the report. They come from the entire, engineered sample, across all three studies (section 2.1.2), so do not immediately translate to generalized results. Section 4 seeks to generalize results to the Pacific Northwest housing stock. This first findings section features tables, graphs, and discussion of the equipment as measured and annualized where necessary. These initial results are not intended to generalize to the population at large, but rather serve as a reference for later engineering calculations and simulations.

3.1. Occupancy and Installation Characteristics

In contrast to space conditioning, there are few characteristics determining hot water consumption. The main driver is the number of occupants per household. Additional drivers are the location in the house where the water heater is installed and the climate.

3.1.1. Occupancy

Overall, there were 2.7 people per household in the study which matches the average observed in the RBSA (Baylon 2012). Table 3 and Table 4 show the occupant distribution across the metered sites by climate, installation location, and equipment type.

Table 3. Occupant Count by Installation Type and Heating Zone

Number of Occupants	Heating Zone				
	HZ1	HZ2	HZ3	All Zones	n
Basement	2.5	2.4	2.5	2.4	18
Garage	3.0	2.8	0.0	2.9	44
Interior	3.0	2.8	2.3	2.7	14
Interior Ducted	2.5	2.3	3.1	2.6	31
All Installations	2.8	2.6	2.8	2.7	107

Table 4. Occupant Count by Equipment Type and Installation Location

Installation Equipment	Number of Occupants					
	Basement	Garage	Interior	Interior Ducted	All Installations	n
Voltex 60 & 80 Gallon	2.83	3.38	2.75	5.00	3.14	28
ATI 66 gallon	2.67	2.54	2.50	2.50	2.52	46
GeoSpring 50 gallon	2.11	2.89	3.00	0.00	2.65	33
All Equipment	2.44	2.93	2.58	2.58	2.72	107

Table 5 and Table 6 display the age distribution and the spread of household size across the field sites.

Table 5. Occupant Age Distribution

Installation Equipment	Count of Occupants by Age Group					
	Preteen	Teen	Adult	Retired	Total Occupants	n
Voltex 60 & 80 Gallon	0.75	0.29	1.61	0.46	3.14	28
ATI 66 gallon	0.69	0.13	1.38	0.69	2.52	46
GeoSpring 50 gallon	0.44	0.24	1.74	0.21	2.65	33
All Equipment	0.60	0.23	1.62	0.40	2.72	107

Table 6. Occupant Count Distribution

Count of Occupants	Installation Location			
	Basement	Garage	Interior	All Locations
1	1	5	6	12
2	11	15	22	48
3	4	9	7	20
4	1	10	5	16
5+	1	5	5	11
All Counts	18	44	45	107

3.1.2. Water Heaters

Table 7 shows the final distribution of water heater installation locations across climate zones. The breakdown based on equipment type was given previously in Table 2.

Table 7. Water Heater Installation Locations Across Climates

Installation Location	Heating Zone			
	HZ1	HZ2	HZ3	All Zones
Basement	11	5	2	18
Garage	33	11	0	44
Interior	3	8	3	14
Interior Ducted	12	12	7	31
Total	59	36	12	107

3.2. Annualized Measurements

This section contains tables of annualized measurements for each water heater make, one row per monitored unit. The annualization was performed as mentioned in section 2.3.2, and described more fully in Appendix B: Details of Data Annualization. A detailed discussion of the calculations for COP and what we denote "aCOP" can be found in section 3.7.1. Note that these are not generalized findings to extend to the population at large, but rather merely a descriptive summary of the data from the various HPWH field projects, standardized to annual estimates for ease of comparison and interpretation.

The following naming conventions exist in all tables:

- Flow – daily average gallons/day of hot water draw
- hpwh – daily average kWh/day of total HPWH energy use
- resfrac – fraction of input energy provided by resistance elements
- Tin – average inlet water temperature (°F)
- Tout – average delivered outlet water temperature (°F)
- Tintake –average intake air temperature to the evaporator, while the HPWH runs (°F)
- aCOP – average annual coefficient of performance, the ratio of useful energy delivered to input energy, which includes the penalty due to resistance heat and standby losses.
- hpCOP – heat pump only annual average coefficient of performance – excludes resistance heat and standby loss effects

Table 8. GE GeoSpring Water Heater Annualized Measurements

siteid	Location	City	Flow	hpwh	Resfrac	Tin	Tout	Tintake	aCOP	hpCOP
99086	Garage	Arlington, WA	47	7.1	69%	50	116	56	1.24	2.16
99122	Garage	Hood River, OR	51	8.4	71%	50	126	59	1.30	2.48
99085	Garage	Stanwood, WA	50	7.7	52%	50	123	58	1.36	2.42
90093	Basement	Seattle, WA	28	4.3	42%	53	121	57	1.41	1.73
99107	Garage	Eugene, OR	51	6.5	62%	53	117	60	1.44	2.83
99102	Garage	Vancouver, WA	92	10.6	72%	53	117	63	1.46	3.09
99140	Garage	Milton Freewater, OR	23	3.0	40%	56	118	61	1.55	2.40
99108	Garage	Springfield, OR	30	4.4	40%	50	125	58	1.57	2.15
99124	Basement	Hood River, OR	49	6.4	59%	47	126	73	1.62	3.55
99155	Garage	McMinnville, OR	82	8.8	44%	56	121	59	1.64	2.52
99088	Garage	Everett, WA	34	4.0	38%	52	117	59	1.66	2.38
99149	Garage	Richland, WA	34	4.4	39%	56	128	67	1.67	2.54
90253	Garage	Sisters, OR	46	7.3	57%	48	142	57	1.68	2.94
99065	Garage	Vancouver, WA	49	5.1	51%	55	117	59	1.71	2.66
99118	Interior	Frenchtown, MT	22	3.7	40%	45	138	70	1.73	2.36
99119	Basement	Frenchtown, MT	40	4.6	34%	47	121	67	1.74	2.19
99098	Garage	Marysville, WA	60	7.4	47%	51	131	58	1.78	2.86
99105	Garage	Cathlamet, WA	63	6.0	39%	54	116	64	1.80	2.77
99087	Garage	Everett, WA	28	2.9	26%	54	117	60	1.89	2.36
90129	Basement	Seattle, WA	53	5.6	28%	56	131	59	2.04	2.67
99123	Interior	Parkdale, OR	38	3.6	25%	49	118	60	2.09	2.67
90012	Basement	Bend, OR	31	3.5	13%	44	126	56	2.19	2.41
90153	Interior	Kalispell, MT	60	6.1	34%	50	130	60	2.19	2.92
99142	Basement	Bonners Ferry, ID	56	6.0	29%	45	131	68	2.21	2.90
23544	Basement	Idaho Falls, ID	39	3.7	17%	52	127	58	2.30	2.65
99150	Garage	Kennewick, WA	19	1.8	12%	57	124	63	2.36	2.66
21723	Interior	Ephrata, WA	19	1.8	13%	53	122	68	2.40	2.67
90105	Interior	Seattle, WA	46	3.7	25%	53	126	66	2.57	3.25
99104	Basement	Naselle, WA	24	1.8	4%	51	117	62	2.68	2.77

Table 9. AO Smith Voltex Water Heater Annualized Measurements

siteid	Location	Type	City	Flow	hpwh	resfrac	Tin	Tout	Tintake	aCOP	hpCOP
90135	Garage	60gal	Bend, OR	8	2.2	26%	47	120	52	1.33	1.52
20814	Garage	80gal	Vancouver, WA	122	17.3	43%	55	148	56	1.72	2.83
90050	Garage	60gal	Spokane, WA	29	4.6	23%	49	138	53	1.78	2.24
90069	Interior	80gal ducted	Cheney, WA	111	9.6	37%	65	124	61	1.80	2.35
11531	Garage	60gal	Kent, WA	64	6.6	19%	56	123	53	1.83	2.20
23744	Interior	60gal	Spokane, WA	12	2.0	0%	50	129	55	1.90	1.91
90034	Garage	80gal	Bend, OR	64	6.9	43%	47	123	51	1.92	3.01
99094	Garage	80gal	Vancouver, WA	63	5.4	17%	55	117	56	2.04	2.46
22897	Garage	80gal	Vancouver, WA	23	2.8	1%	56	135	58	2.21	2.24
13265	Garage	60gal	Everett, WA	24	2.4	6%	52	121	58	2.23	2.34
90028	Interior	60gal	Spokane, WA	17	1.9	0%	53	123	58	2.25	2.26
90030	Basement	60gal ducted	Spokane, WA	36	3.2	4%	50	120	64	2.30	2.37
90131	Garage	80gal	Redmond, OR	76	5.7	22%	50	113	54	2.30	2.95
21578	Garage	80gal	Vancouver, WA	84	6.4	12%	54	122	56	2.39	2.71
99106	Garage	80gal	Springfield, OR	26	2.4	1%	56	127	59	2.46	2.49
11289	Basement	60gal	Seattle, WA	25	2.1	0%	56	120	54	2.50	2.50
22096	Interior	60gal	Rigby, ID	74	6.0	13%	47	125	67	2.56	2.81
90130	Interior	80gal	Bend, OR	69	5.2	8%	48	122	66	2.57	2.73
90015	Basement	60gal	Eugene, OR	47	3.6	9%	52	123	57	2.67	2.87
90159	Interior	60gal	Corvallis, MT	40	2.9	0%	51	120	58	2.73	2.73

Table 10. AirGenerate ATI Water Heater Annualized Measurements

siteid	Location	Type	City	Flow	hpwh	resfrac	Tin	Tout	Tintake	aCOP	hpCOP
99307	Basement	66gal unducted	Oregon City, OR	31	5.4	6%	55	128	55	1.30	1.32
99328	Interior	66gal ducted	Redmond, OR	59	8.3	7%	50	123	71	1.40	1.44
99324	Garage	66gal unducted	Bend, OR	34	5.1	2%	52	123	54	1.42	1.43
99302	Garage	66gal unducted	Tigard, OR	30	4.5	4%	53	123	57	1.45	1.47
90166	Basement	66gal ducted	Idaho Falls, ID	37	4.9	7%	53	127	59	1.61	1.66
99323	Interior	50gal ducted	Redmond, OR	30	3.7	8%	52	121	70	1.63	1.68
92828	Basement	66gal ducted	Idaho Falls, ID	58	7.2	6%	52	128	60	1.70	1.75
99312	Garage	66gal unducted	Beaverton, OR	45	5.3	11%	60	127	54	1.70	1.80
99319	Basement	66gal ducted	Madras, OR	29	3.6	5%	57	125	62	1.70	1.74
99306	Garage	66gal unducted	Beaverton, OR	64	6.8	7%	56	123	58	1.73	1.79
99310	Basement	66gal unducted	Spanaway, WA	34	4.4	9%	55	127	51	1.74	1.83
90162	Basement	66gal ducted	Idaho Falls, ID	27	3.8	2%	52	133	65	1.78	1.79

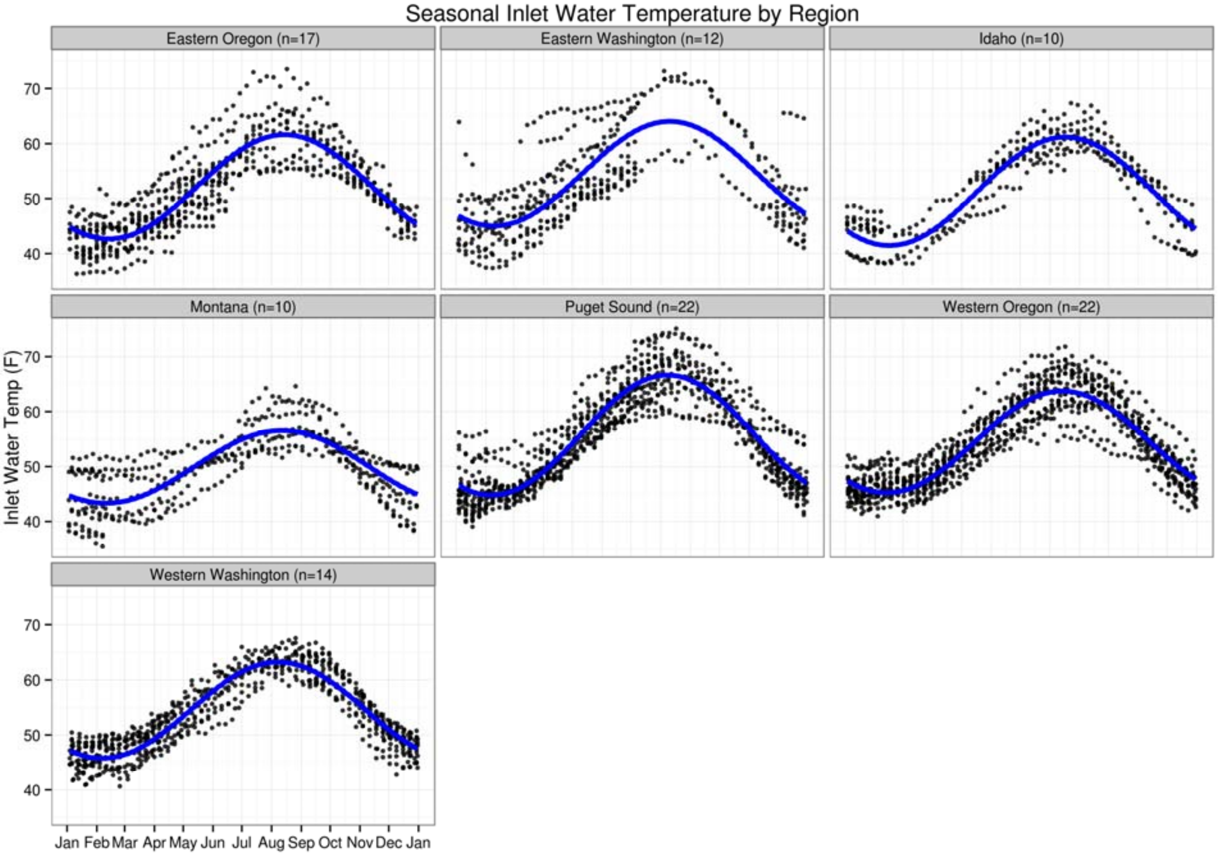
99316	Interior	66gal ducted	Portland, OR	16	2.1	8%	52	118	60	1.81	1.90
99320	Garage	66gal unducted	Bend, OR	17	2.4	1%	52	124	56	1.81	1.82
99309	Basement	66gal unducted	Issaquah, WA	58	6.6	0%	57	132	53	1.82	1.82
99317	Garage	66gal unducted	Redmond, WA	28	3.0	3%	56	115	51	1.82	1.85
99313	Interior	66gal ducted	Bend, OR	28	3.2	9%	51	123	74	1.85	1.93
99311	Interior	66gal ducted	Renton, WA	12	1.8	0%	60	130	67	1.86	1.86
99330	Interior	66gal ducted	Prineville, OR	20	2.3	1%	52	123	71	1.88	1.89
99301	Garage	50gal unducted	Olympia, WA	11	1.6	4%	55	117	51	1.93	1.97
99329	Interior	66gal ducted	Bend, OR	13	1.5	1%	51	122	70	1.98	2.00
99318	Interior	66gal ducted	Portland, OR	16	1.7	0%	53	118	71	2.01	2.01
92887	Basement	66gal ducted	Polson, MT	58	5.9	11%	50	126	66	2.06	2.20
99303	Interior	66gal ducted	Portland, OR	23	2.4	0%	55	122	65	2.06	2.06
92690	Basement	66gal ducted	Idaho Falls, ID	96	9.2	27%	51	129	67	2.11	2.58
99325	Basement	66gal ducted	Portland, OR	29	3.0	5%	55	126	63	2.11	2.18
99305	Interior	66gal ducted	Portland, OR	49	4.4	3%	55	125	71	2.12	2.15
99315	Garage	66gal unducted	Hillsboro, OR	25	2.6	3%	55	123	59	2.13	2.18
99308	Garage	66gal unducted	Sammamish, WA	51	4.8	9%	63	132	55	2.18	2.31
99322	Garage	66gal unducted	Bend, OR	38	3.4	6%	53	119	59	2.19	2.28
24339	Interior	66gal ducted	Mead, WA	22	2.2	0%	48	120	71	2.20	2.21
11219	Garage	66gal ducted	Tacoma, WA	46	4.3	6%	52	125	63	2.22	2.30
99326	Interior	66gal ducted	Portland, OR	23	2.1	1%	54	121	67	2.24	2.25
99327	Garage	66gal unducted	Oregon City, OR	23	2.1	1%	54	121	67	2.24	2.25
99304	Interior	66gal ducted	Portland, OR	36	3.1	3%	56	124	73	2.25	2.30
92661	Garage	66gal ducted	St Ignatius, MT	48	4.2	10%	50	121	59	2.26	2.42
92719	Basement	66gal ducted	Polson, MT	22	2.1	0%	48	117	62	2.30	2.30
10292	Basement	66gal ducted	Renton, WA	54	4.3	5%	56	123	64	2.35	2.43
90003	Basement	66gal ducted	Seattle, WA	62	4.6	0%	58	124	65	2.42	2.43
93144	Interior	66gal ducted	Idaho Falls, ID	85	6.4	13%	54	123	65	2.44	2.69
90168	Basement	66gal ducted	Idaho Falls, ID	37	2.9	6%	52	121	67	2.49	2.60
90169	Basement	66gal ducted	Idaho Falls, ID	49	4.1	6%	54	128	65	2.49	2.61
92602	Basement	66gal ducted	Polson, MT	31	2.6	3%	48	118	64	2.49	2.55

3.3. Inlet Water Temperature

3.3.1. Seasonal and Climatic Variation

Measured inlet water temperatures are displayed by day of year and geographic region in Figure 5. These inlet water temperatures were reported only during flow events, when water moving across the sensor enables accurate measurements. The measured temperature can “drift” during periods of no flow for several reasons. The presence of hot water in the tank or exposed piping to ambient temperatures can affect the reading, so measurements are only valid during flow events. The temperatures during flow events were tabulated and averaged daily for the plot of Figure 5. As expected, the annual profiles revealed sinusoidal shapes, with the greatest amplitudes observed in milder, coastal regions more reliant on surface water, and the smallest amplitudes observed in inland regions more reliant on ground water. Section 4.1 describes how these data were used to create generic inlet water temperature simulation inputs.

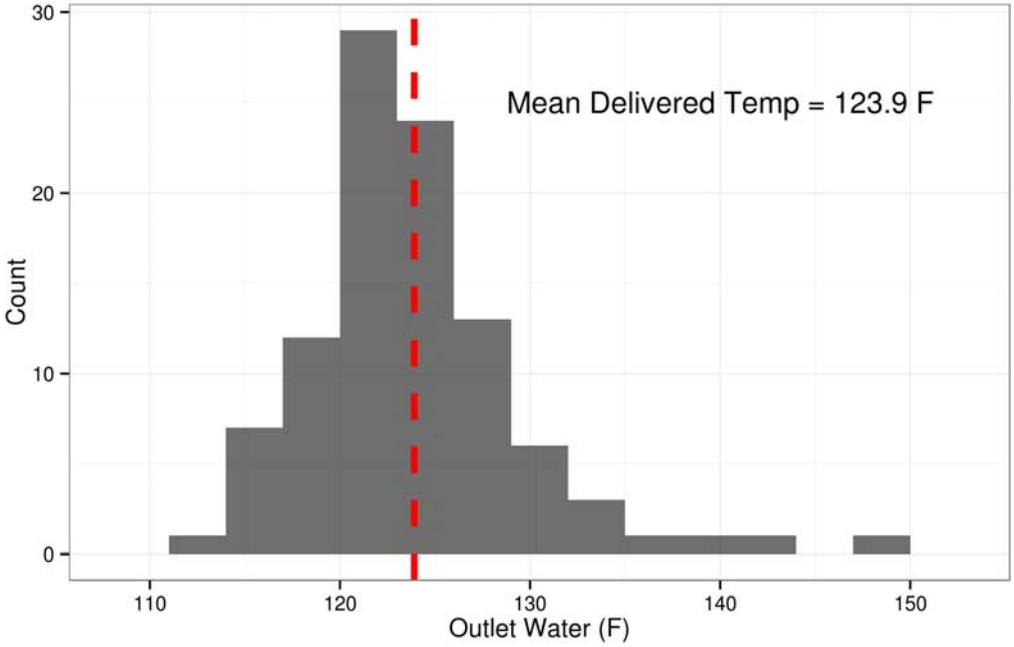
Figure 5. Measured Inlet Water Temperature by Region



3.4. Outlet Water Temperature

Figure 6 shows the distribution of average delivered water temperature, as measured just downstream of the tank outlet, across all sites. Similar to the inlet water temperature, outlet water temperature measurements are only valid during flow events, due to temperature drift of water in the piping close to the tank. Most sites delivered hot water at an average temperature between 115° F and 130° F. Note that this is not the tank setpoint, but necessarily some value below the actual setpoint. Water heaters are designed to heat water to a specified setpoint and include a "deadband" through which the tank temperature drifts down before the tank is reheated. Water withdrawn before a reheat cycle will therefore be somewhat lower than the target setpoint. A few sites showed average delivered water temperature above 140° F. This seemed unusual and suspicious, but further investigation revealed the data to be valid.

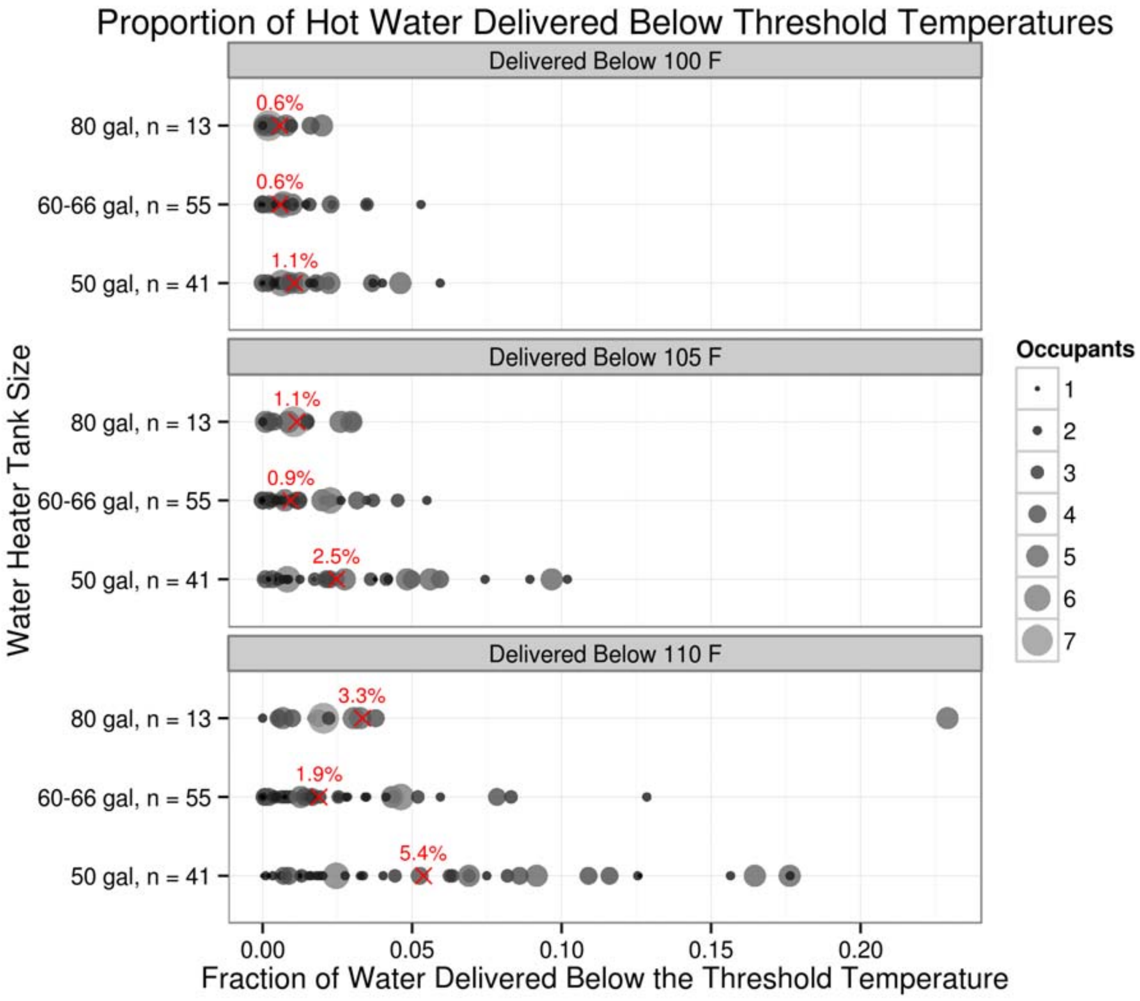
Figure 6. Distribution of Average Delivered Water Temperature



The outlet temperature data also allowed an investigation into whether the heat pump water heaters successfully maintained high enough delivered water temperatures. In other words, maintaining a high outlet temperature means that the occupants' demands for hot water are being met. Figure 7 shows the proportion of hot water delivered below three threshold temperatures – 100° F, 105° F, and 110° F – tabulated by tank size. The data were aggregated at 5 minute intervals. The red "x" marks the average fraction in each category.

Across all tank sizes, approximately one percent of all hot water was delivered cooler than 100° F. Therefore, at a basic level, heat pump water heaters were able to meet the demand or, alternatively, occupants curtailed their use when the temperature dropped below a useful level (whether that demand was met efficiently, or through costly calls to resistance heat, is not summarized here). At 105° F the proportions were largely similar, although the 50 gallon tanks delivered 2.5% of water below 105° F compared to 1.1% below 100° F. The amount of water delivered below 110° F was a bit higher: on average 5.4% for the 50 gallon tanks, 1.9% for the 60-66 gallon tanks, and 3.3% for the 80 gallon tanks. The higher mean fraction for the 80 gallon tanks was mostly driven by a home in the Ecotope study with six occupants and a setpoint of approximately 120° F. The slightly higher fraction of the 50 gallon tanks appears mostly driven by homes with three or more occupants. This is possibly a case of under sizing, as a 50 gallon heat pump water heater is likely inappropriate for larger households.

Figure 7. Proportion of Hot Water Delivered Below Threshold Temperatures



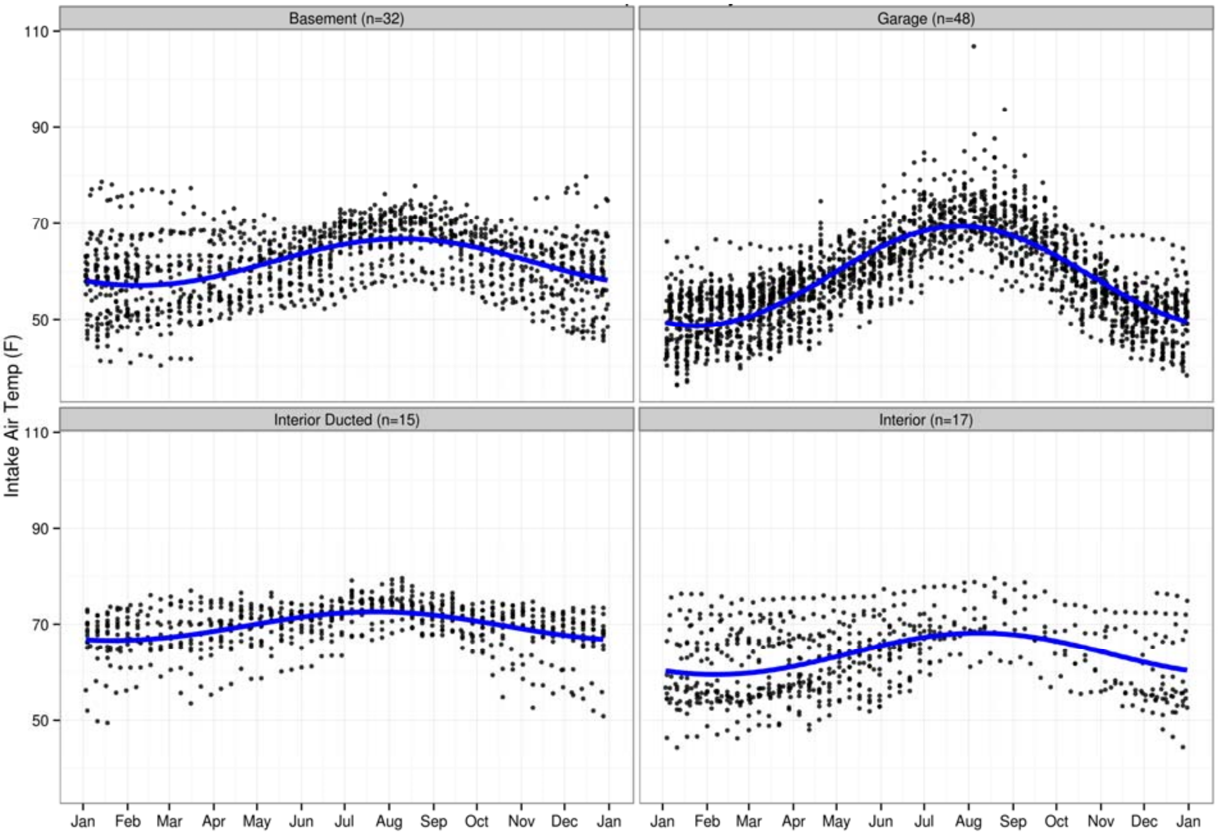
3.5. Ambient Space Temperature

The ambient air conditions surrounding the water heater are a critical determinant of energy consumption. With all storage tank water heaters, the ambient air influences the amount of heat lost through the tank insulation. With HPWHs, the ambient air provides the intake air, and energy source, for the refrigeration cycle. Higher intake air temperatures equate to higher efficiencies. In an interesting feedback loop, the integrated HPWHs, installed in an enclosed space have the ability to cool off the same ambient air they use as the heat source. Knowing the ambient air conditions is crucial to understanding HPWH performance.

Figure 8 displays the temperature across all sites according to the four, primary installation configurations: basements, garages, interior with exhaust ducting, and interior with recirculating air. In the figure, each point is the weekly average temperature for a single site. Much like the water measurements, we only use air temperature measurements when the HPWH is running.

Garages, which aren't conditioned and are the least thermally coupled to the house, show the largest temperature swing across the year. Basements, with more ground contact and a better thermal connection to the house, are second in the amplitude of the swing. Comparing the interior installations is illuminating. The installs with exhaust ducting send all the colder air outside the envelope while drawing in conditioned air from other parts of the house. They are, in effect, supplied with air from a regulated temperature source. In contrast, the interior installs, without any ducting, exhibit a cooler temperature and more annual change. This can be attributed, in large part, to the water heaters cooling down the space in which they are installed. Unlike the ducted cases, there is not a continuous, regulated supply of conditioned air. As the entire house gradually warms in the summer, so does the intake air for the water heater.

Figure 8. Weekly Average Intake Air Temperature by Install Location



The fraction of time in a year spent at a given ambient temperature is shown in Figure 9, Figure 10, and Figure 11 for heating zones 1, 2, and 3 respectively. The data are reported using the 1-minute logging interval only when the HPWH is running. Like the other data in the findings section, the results are annualized where there isn't a complete year on record. Each temperature bin is 5 degrees wide centered around the value given. For the HPWHs in this study, the critical temperature bins are 47° F and below. Some of the water heaters turn off their heat pump and switch to resistance heat when the temperatures fall below 45° F. Consequently, all water heating in 42° F bin and below is done at the low efficiency of the resistance element. The 47° F bin is important because it is the marginal case. Slight changes in weather or temperature measurements could push the water heater into resistance heat. For reference, the percent of time in each bin is also printed at the top of each bar. As expected, garages trend colder than basements, which, in turn are colder than interior spaces. No garage installations are present in zone 3. It is extremely rare to install a water heater in any unconditioned place in these colder climates.

Figure 9. Ambient Temperature Profiles, Heating Zone 1

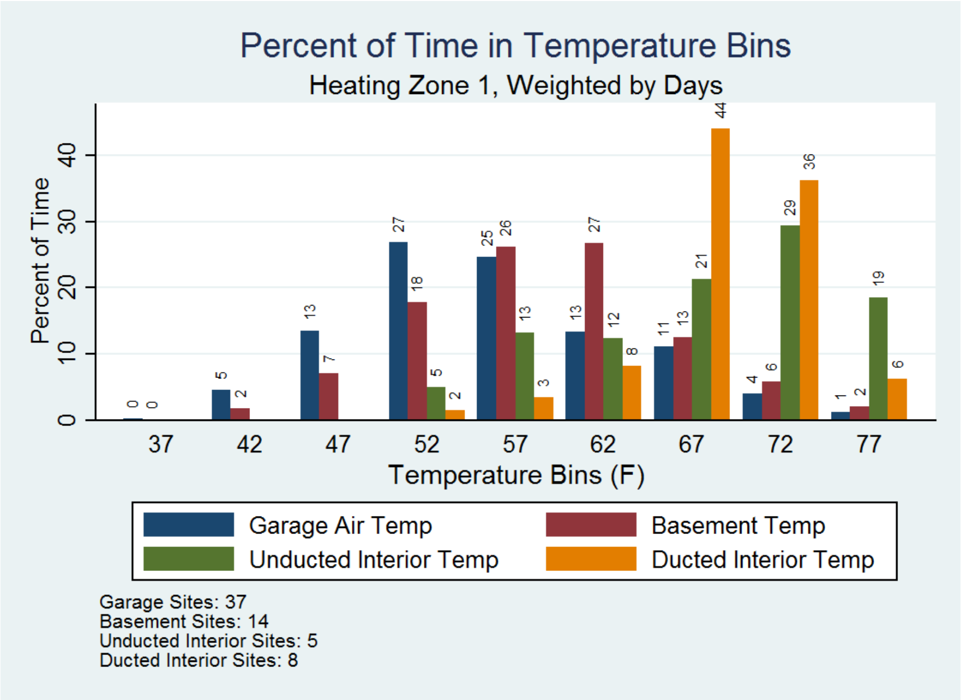


Figure 10. Ambient Temperature Profiles, Heating Zone 2

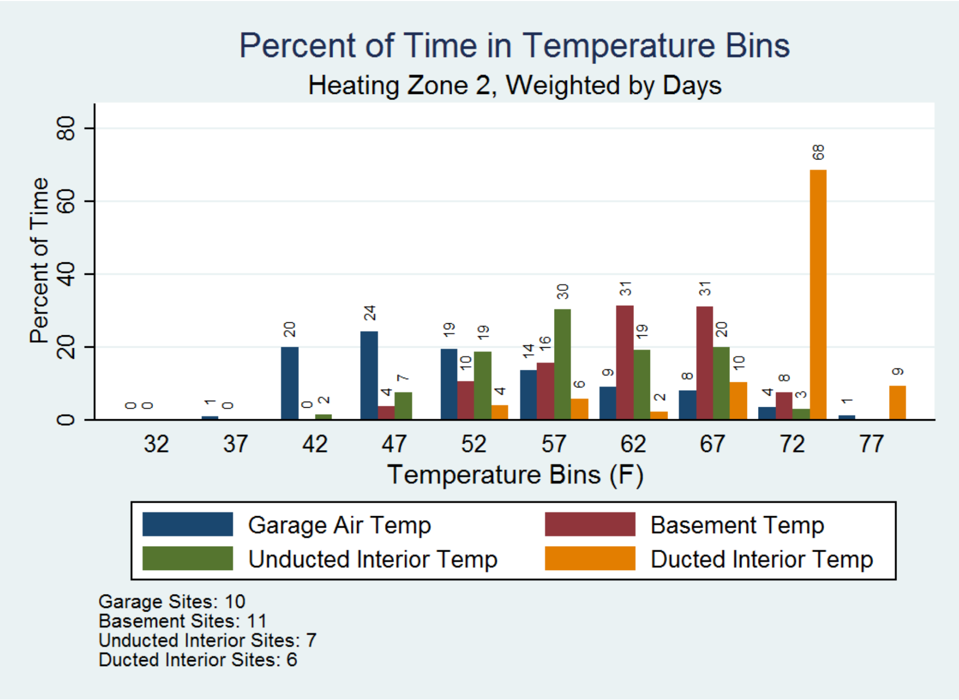
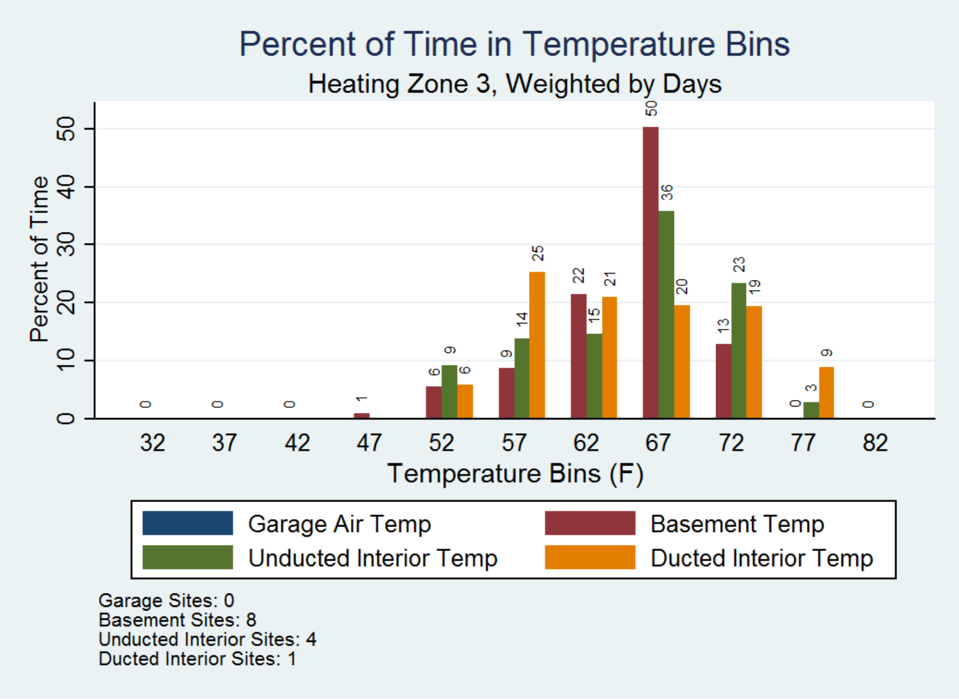


Figure 11. Ambient Temperature Profiles, Heating Zone 3

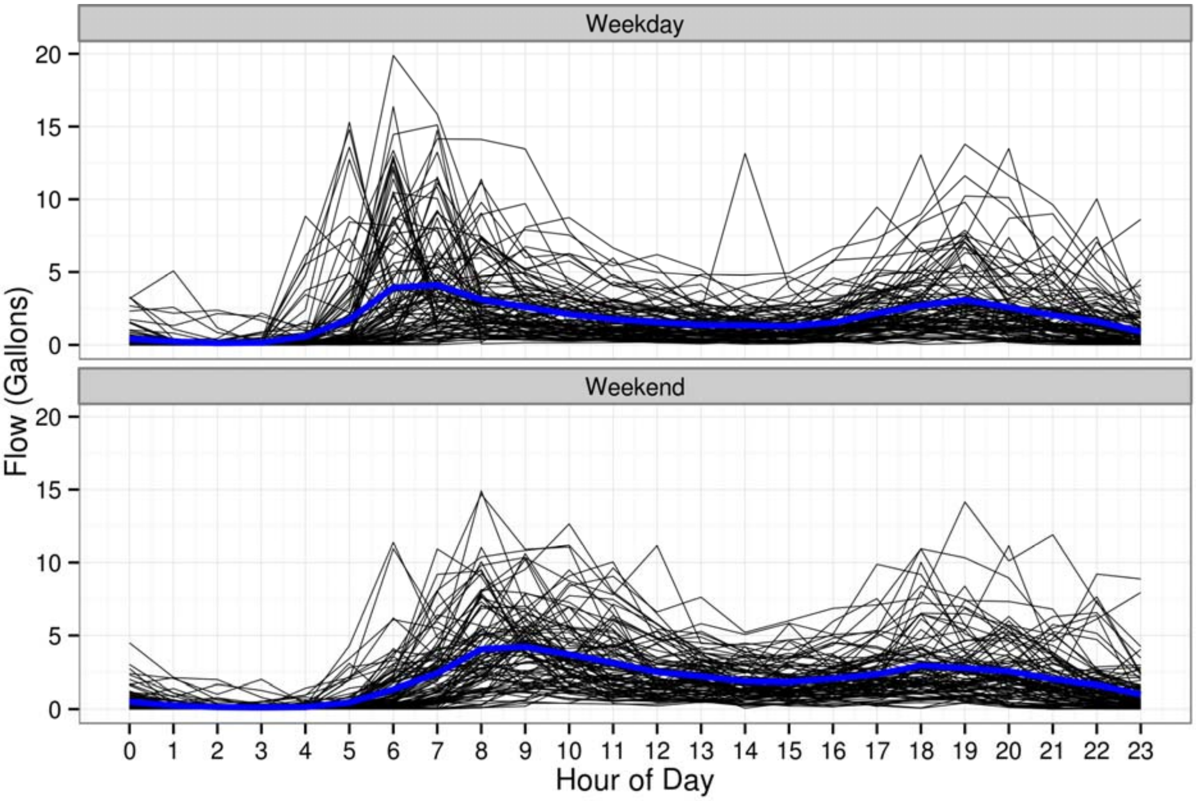


3.6. Hot Water Draw Patterns

3.6.1. Weekday and Weekend Shapes

The average hot water draw for weekdays and weekends is shown in Figure 12. Each site's hourly average flow across the entire monitoring period is plotted as a thin, black line while the average across all sites is plotted in blue. Unsurprisingly, this daily load shape for hot water flow closely resembles the energy load shape of other storage water heaters as measured in the RBSA Metering project (Ecotope 2014). The shape shows the expected peak use in the morning and a secondary peak in the evening. Further, the weekend morning peak is delayed compared to the weekday.

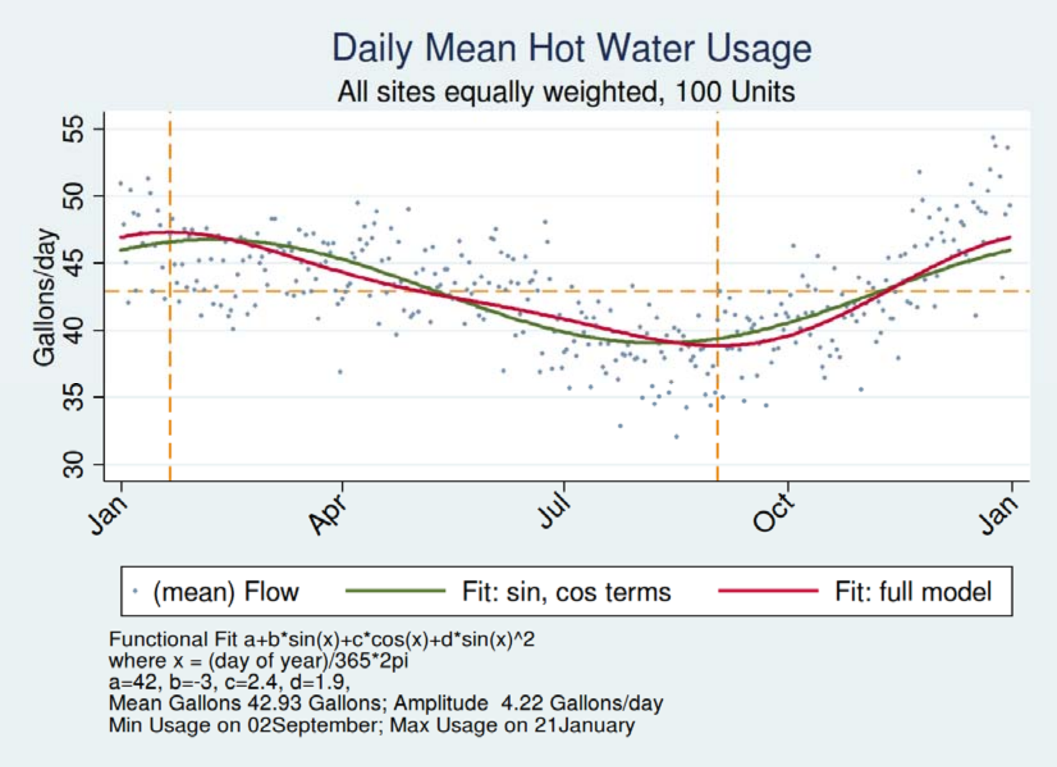
Figure 12. Average Hot Water Draw by Hour of Day, 99 Sites



3.6.2. Annual Shapes

Due almost entirely to changing inlet water temperatures, the daily average water use changes over the year. While we expect that occupant-driven water usage patterns are consistent across seasons, in colder months, more hot water must be mixed with the colder inlet water to provide a comfortable temperature at the faucet or showerhead. Figure 13 plots the average water used each day across all sites on an annual basis. The green and red lines are fits to the data. Two fits were conducted on the annual hot water usage data, one with four terms and one with three. The four term fit had constant, sine, cosine, and sine-squared terms while the three term fit omitted the sine-squared term. The sine-squared fit predicts the water draws decidedly better, which is seen by how well it follows the data. The minimum hot water usage occurred on September 2nd, and the maximum on January 21st, times which match the fit for electricity usage quite well (Figure 14). The seasonal variation of ±4 gallons per day can be explained by the changing temperature of the water supply.

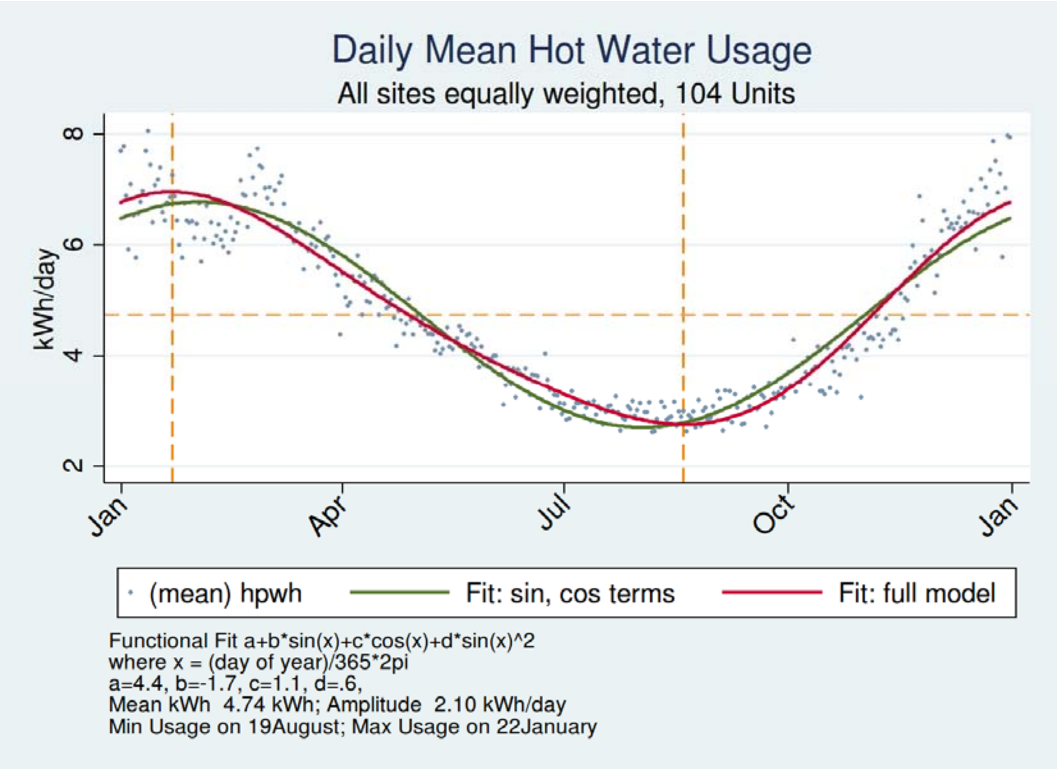
Figure 13. Seasonal Daily Average Hot Water Usage



Functional Fit a+b*sin(x)+c*cos(x)+d*sin(x)^2
where x = (day of year)/365*2pi
a=42, b=-3, c=2.4, d=1.9,
Mean Gallons 42.93 Gallons; Amplitude 4.22 Gallons/day
Min Usage on 02September; Max Usage on 21January

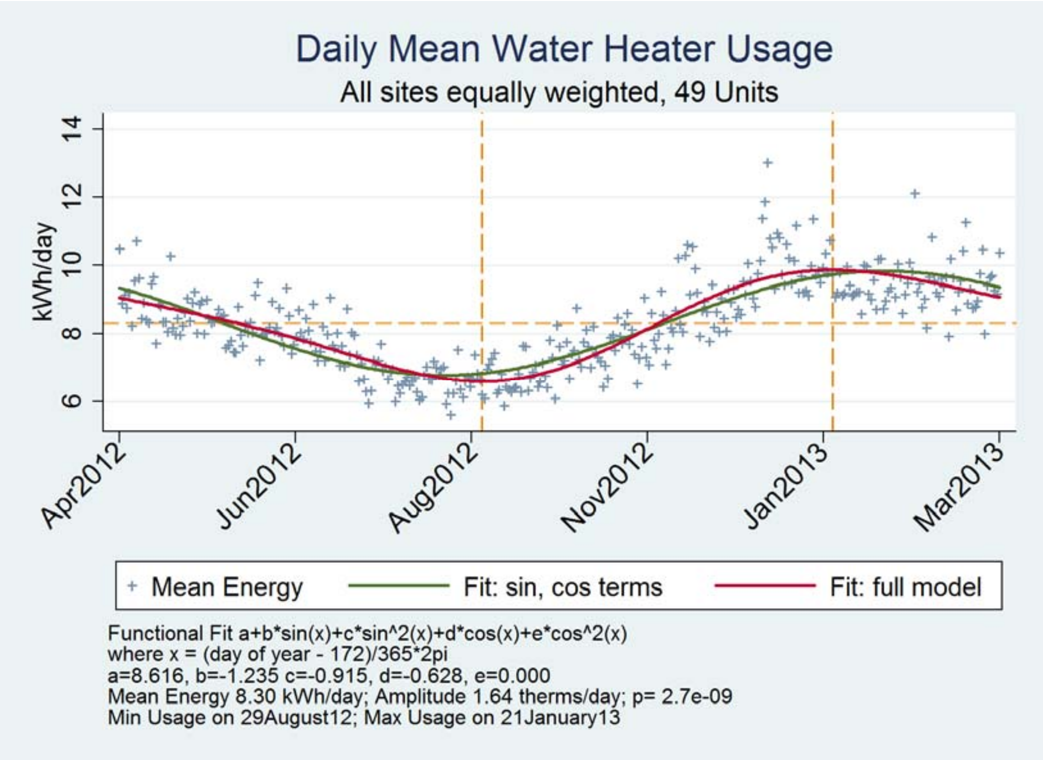
Excluding tank standby losses, the amount of electricity used depends on the underlying hot water draw pattern; as a result, the electricity use has the same annual shape as the hot water use. Fits similar to those discussed earlier were performed with results in Figure 14. The average daily electricity usage is 4.7 kWh/day, and the fit indicates the seasonal variation of the electric energy draw is ±2.1 kWh/day with a high on January 22th and a low on August 19th. Both the energy extrema match the water draw extrema well.

Figure 14. Seasonal HPWH Electricity Usage



Compare the data for heat pump water heaters to the data for electric resistance water heater (ERWH) usage in the RBSAM study: a model with the same terms showed an average electricity usage of 8.3 kWh/day, and a seasonal variation of ±1.64 kWh/day with a high on January 21st and a low on August 29th. In this RBSA Metering study these homes had an average occupancy of 2.2 people per house. Using equation 3 from the RBSA Metering report to predict the energy use of ERWHs for 2.7 people – the same occupancy as in this study – gives 9.2 kWh/day. The daily electricity usage is nearly double, as expected, due to the relative inefficiency of resistance water heaters compared to heat pump water heaters. The extrema match up fairly well; the 10 day difference in August is small compared to the amount of variation present in the data.

Figure 15. Seasonal ERWH Usage (RBSA Metering)



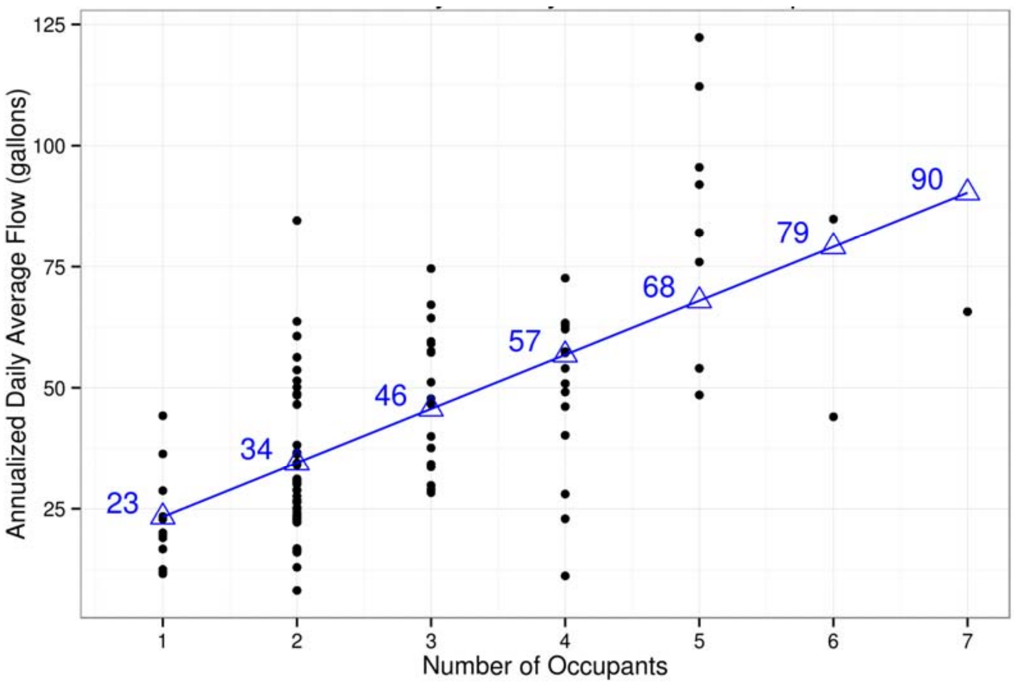
3.6.3. Daily Draw Volume and Events

Table 11 tabulates annualized average daily DHW flow in gallons by number of occupants. Figure 16 shows the data from which the table was derived along with a regression line to estimate mean daily flow by occupancy. The average flow for single occupant homes was 22 gallons per day. Notice in the table that the average usage of 3 occupant homes was identical to the average usage of 4 occupant homes. This is likely due to sampling variability. The regression model and data displayed in Figure 16 suggest an average usage of 23 gallons per day for a single occupant home, with each additional occupant contributing an additional 11 gallons per day. Consequently, the average house with 2.7 occupants is expected to use 42 gallons per day.

Table 11. Average Daily Flow by Occupancy

Occupants	Annualized Daily Flow (Gal)		n
	Mean	SD	
1	22	10.0	12
2	34	15.3	43
3	48	13.8	19
4	48	17.4	14
5	85	25.8	8
6	64	28.9	2
7	66	-	1

Figure 16. Average Daily Draw Sizes by Number of Occupants



Examining the field data for draw characteristics as outlined in section 2.3.4, results in Table 12. For a given household occupant count, the table describes the typical draws per day. In the table, “Clusters” refers to the number of event clusters, typically 60 minutes in duration but also ranging 30-90 minutes, over the course of the day. Total flow is the total hot water drawn over the course of the day. The small, mid, and large flow columns delineate how many gallons of water were drawn by the small, mid, and large draws. Similarly, total draws refers to the number of draws per day while small, mid, and large draws list the average number of that size draw per day. For example, the single occupant household had three event clusters using 23 gallons in all. Small flows account for 6/23 or 26% of the daily draw. There are typically 4.5 small flow events per day. Likewise, there are only 0.7 large draws per day. Last, the “Sites” column lists how many different sites were used to create the data summaries and the “Days Metered” column tells how many different days’ worth of draws were observed. Refer to Appendix E: Draw Profiles to see tables of the weekly characteristics and for information on how draws are distributed within each event cluster.

Table 12. Daily Draw Characteristics

Occupant Count	Clusters per Day	Gallons per Day				Draw Count per Day				Sites	Days Metered
		Total Flow	Small Flow	Mid Flow	Large Flow	Total Draws	Small Draw	Mid Draw	Large Draw		
1	3	23	6	5.5	11.5	6.4	4.5	1.1	.7	7	2160
2	5	34	10.4	7.7	16	12.6	9.7	1.7	1.1	32	10602
3	5	46	13.8	10.7	21.5	15.3	11.9	2.2	1.3	14	5193
4	5	57	13.8	12.5	30.9	14.6	10.7	2.3	1.7	13	4440
5+	5	72	14	14.7	43.3	18.5	12.6	3.2	2.7	10	2448

3.7. Overall Water Heater Energy Use

Table 13 shows annualized estimates of water heater annual kWh for the observed units. These are useful to develop a summary of typical usage, as the averages in each cell include sites of varying draw sizes and operating conditions. Table 14 shows a metric for total water drawn normalizing kWh used per 100 gallons delivered by make and location.

Table 13. Annualized Annual Energy (kWh/yr) of Monitored Water Heaters

Annualized Water Heater Energy Use (kWh/yr)						
Equipment	Basement		Garage		Interior	
	Mean	n	Mean	n	Mean	n
ATI	1,678	16	1,380	13	1,201	15
GeoSpring	1,600	9	2,185	17	1,549	6
Voltex	1,696	6	2,208	13	1,785	8

Table 14. Annualized Annual kWh per 100 gallons delivered

Annualized kWh per 100 Gallons Delivered						
Equipment	Basement		Garage		Interior	
	Mean	n	Mean	n	Mean	n
ATI	10.3	16	10.7	13	10.7	13
GeoSpring	10.8	8	13.2	17	10.7	5
Voltex	8.2	3	12.0	13	9.9	6

3.7.1. System Efficiency

System efficiency is defined as useful energy output divided by energy input. We calculated the system efficiency over the entire year of operation and denoted it “aCOP” for average (or annual) coefficient of performance as defined in equations 1-3:

$$\boldsymbol{aCOP = Q_{delivered}/Q_{input}}$$ **Equation 1.**

$$\boldsymbol{Q_{delivered} = mc_p(T_{outlet} - T_{inlet})}$$ **Equation 2.**

$$\boldsymbol{Q_{input} = Q_{heat\,pump} + Q_{resistance}}$$ **Equation 3.**

Where, m is the mass of water passing the flow meter, c_p is the heat capacity of water, T_{outlet} is the outlet water temperature, T_{inlet} is the inlet water temperature, $Q_{heat\,pump}$ is the energy input to the heat pump system (includes compressor and fan), and $Q_{resistance}$ is the energy input to the resistance elements. “aCOP” is an analogous quantity to the Energy Factor (EF) but we opt not to use the term “EF” because it is defined by a specific set of test conditions not the actual operating conditions observed here. Both quantities, however, are concerned with the useful energy output. That is, energy lost through the tank insulation in standby operation reduces the efficiency. In other words, some input energy is used to offset the stand-by losses but it is never realized as “useful” energy in water leaving the storage tank. For instance, with a resistance tank, this is the distinction between an EF of 0.9, and the fact that the heating element converts electrical energy to heat with 100% efficiency.

Examining the relationships in the aCOP equation and those found in the field data, we find that the yearly average efficiency depends on three items. They are, in order of most to least important: (1) how often the resistance elements run, (2) the efficiency of the heat pump cycle, and (3) the insulation level of the tank. Table 15 shows the observed annual, average, system efficiency at all the sites sorted by installation location, heating zone, and equipment.

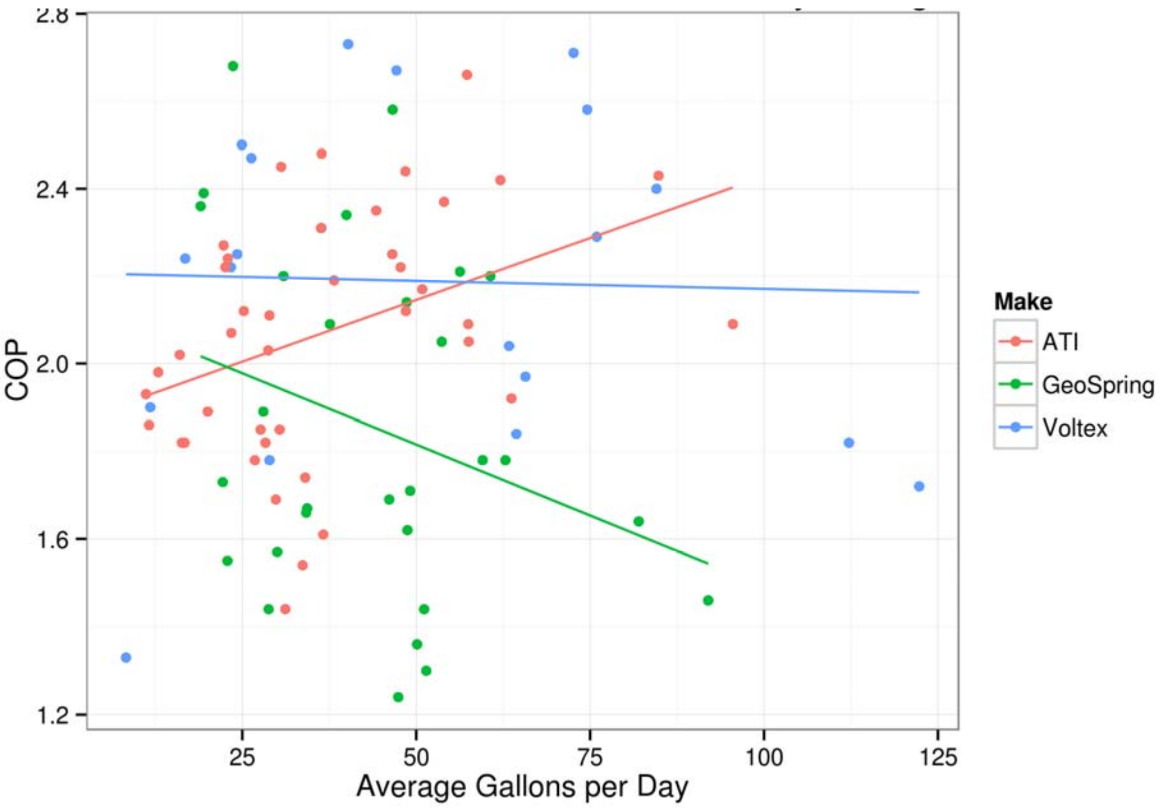
Table 15. Annualized average aCOP by Equipment, Location, and Heating Zone

	Basement		Garage		Interior	
	Heating Zone 1					
Make	**Mean**	**n**	**Mean**	**n**	**Mean**	**n**
ATI	2.03	6	2.05	9	2.06	7
GeoSpring	1.95	4	1.63	15	2.34	2
Voltex	2.59	2	2.13	7	-	0
	Heating Zone 2					
ATI	2.20	4	1.94	4	1.96	5
GeoSpring	2.20	2	1.69	1	2.39	1
Voltex	2.31	1	1.84	4	2.28	5
	Heating Zone 3					
ATI	2.18	6	-	0	2.43	1
GeoSpring	2.24	2	-	0	1.97	2
Voltex	-	0	-	0	2.58	1
	Overall					
ATI	2.13	16	2.02	13	2.05	13
GeoSpring	2.09	8	1.63	16	2.20	5
Voltex	2.49	3	2.03	11	2.33	6

Even though the grouping of the sites in to similar installation locations aligns ambient air conditions, occupancy counts and draw patterns varied among the installations which obscured some of the performance trends and limits the conclusions that can be drawn from Table 15. Nevertheless, some limited generalizations are possible. The ATI units were most consistent in aCOP across operating conditions, with overall ratios of useful heat to input energy of roughly two in all install conditions. The GeoSpring was more efficient in interior installs, but much more heavily penalized by the harsher conditions of a garage install. The Voltex saw the highest overall average COPs in the observed sample.

Figure 17 shows the relationship (or lack thereof) between aCOP and the average daily draw volume. There is clearly a huge amount of variability in the aCOP with much of it due to the installation location ambient air conditions. What little trends are available from the graph show that the GeoSpring tank aCOP decreases as the average daily draw increases. As shown in section 3.7.4, this is due to the control strategy’s use of resistance heat.

Figure 17. aCOP as a Function of Daily Draw



3.7.2. Heat Pump Cycle Efficiency

Extending the analysis to better understand how the equipment functions, we explore the efficiency of the heat pump cycle only. We define a new quantity "hpCOP":

$$hpCOP = \frac{Q_{delivered} + Q_{standby} - Q_{resistance}}{Q_{heat\ pump}}$$ **Equation 4.**

Where $Q_{standby}$ is the energy lost through heat conduction during standby periods. The hpCOP is simply a measure of how efficiently the heat pump transfers heat from the ambient air to the water in the tank. Since we don't have a measure of the internal tank temperature, like in a laboratory setting, the calculation of hpCOP is necessarily an estimate. It is calculated by excluding energy input from the resistance elements and accounting for standby losses. The resistance energy is directly identified in the field data while standby losses must be inferred.

A full estimate of standby loss requires knowledge of the tank heat loss rate, the temperature of the water inside the tank, and the temperature of the environment surrounding the tank. The laboratory testing provided measurements of the heat loss rate, which, lacking conclusive evidence to the contrary, we have assumed to broadly reflect the heat loss rate occurring in the field (see Appendix F: Measuring Tank Heat Loss). The field studies recorded intake air temperature to use as proxies for space temperature. Those sensors were affected by the water heater and only provided reliable readings during heat pump operation, when air actively moved across the sensor. When off, the intake air temperatures indicated unreliable readings suggestive of stagnant, stratified air so we opted not to use them in this mode. Adding to the uncertainty, we don't actually know the average water temperature inside the water heater at any given time – a luxury afforded in detailed lab testing, with a thermocouple tree, but unavailable in field studies. In the end, we used the laboratory measurements of heat loss rate, took the intake air measurements during water heater operation to represent the ambient space temperature surrounding the water heater, and assumed average tank temperature of ten degrees below average delivered temperature. Ultimately, these assumptions, while crude, should offer decent estimates of standby loss, which is itself mainly a second order effect when calculating efficiency.

Table 16 shows the estimated hpCOP averaged over an entire year of operation. The hpCOP estimates are the theoretical maximum efficiency one could expect if there were no standby losses and no element use. The ambient temperature profile of a given location will change the annual hpCOP but the table shows no consistent trends. For example, garages, on average, are always colder than interior locations, so should always show lower hpCOP values. The fact they don’t suggests other, independent variables are influencing the performance.

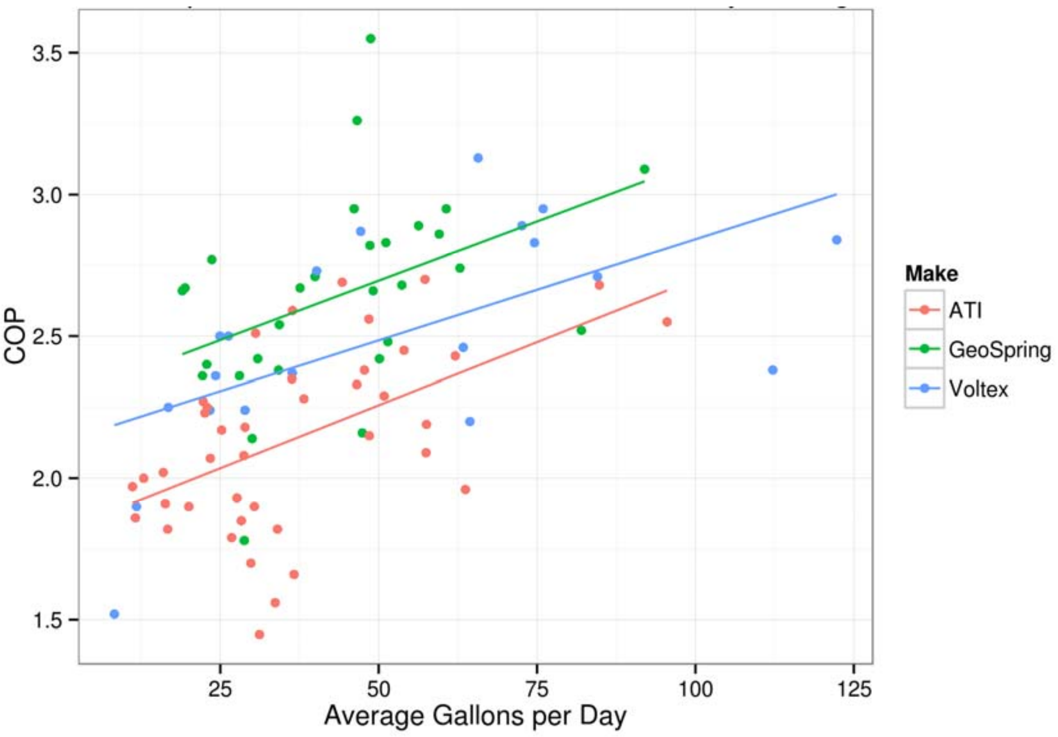
Table 16. Annualized average hpCOP by Make, Location, and Heating Zone

	Basement		**Garage**		**Interior**	
Heating Zone 1						
Make	**Mean**	**n**	**Mean**	**n**	**Mean**	**n**
ATI	2.07	6	2.13	9	2.09	7
GeoSpring	2.69	4	2.55	15	2.97	2
Voltex	2.68	2	2.47	7	-	0
Heating Zone 2						
ATI	2.26	4	2.01	4	1.99	5
GeoSpring	2.66	2	2.95	1	2.67	1
Voltex	2.37	1	2.46	4	2.43	5
Heating Zone 3						
ATI	2.31	6	-	0	2.68	1
GeoSpring	2.76	2	-	0	2.65	2
Voltex	-	0	-	0	2.83	1
Overall						
ATI	2.21	16	2.10	13	2.10	13
GeoSpring	2.70	8	2.57	16	2.78	5
Voltex	2.58	3	2.47	11	2.50	6

Selecting the calibrated engineering approach has proved especially fortuitous, as it has become obvious that there is a significant dependence between heat pump operating efficiency and draw size. See Figure 18 which shows the hpCOP, by site and average daily flow as used to make Table 16. Houses with large daily draws tend to see the heat pump condenser working against much cooler water on average, as the bottom of the tank is repeatedly flooded with cold tap water under numerous water draws. In contrast, houses with small draws see most compressor operation working against warmer water: a home with no draws at all, just standby recoveries, would see the HPWH trying to add heat to 110-120° F water exclusively, which would be relatively inefficient. Further, to the first approximation, most tanks have similar standby losses, so a house with a smaller daily draw has relatively more lost heat through standby than a house with a larger daily draw.

Differences in aCOP between water heaters are highly confounded by differences in draw profiles, as well as other differences in operating conditions, which makes an average COP for an individual unit difficult to interpret. Compared to the Figure 17 graph of system aCOP, Figure 18 is much more orderly because it has used the definition of hpCOP which excludes resistance element use. Consequently, we can conclude that resistance heat, used for whatever reason, is a large influence in the variability and difference in performance between sites. Moreover, it is possible conclude to that the GeoSpring had the most efficient heat pump, followed by the Voltex, and then by the ATI. Likewise, lab tests showed the same finding.

Figure 18. Estimated hpCOP as a Function of Daily Draw



3.7.3. Baseline Water Heating Efficiency

An additional variation on yearly performance is useful to calculate: the "all resistance" or base case efficiency. It is the efficiency of what would have happened in the case that the heat pump were disabled and all heat was provided with the resistance elements. The all resistance COP is defined as "erCOP":

$$\boldsymbol{erCOP} = \boldsymbol{Q_{delivered}} / (\boldsymbol{Q_{delivered}} + \boldsymbol{Q_{standby}}) \qquad \textbf{Equation 5.}$$

The equation states all heat is provided at efficiency of 1. Total energy input, the denominator, is equal to energy delivered plus energy lost. The erCOP is useful because the base case performance of a specific water heater is not 1 or even the rated EF of, say, 0.9. Actual base case efficiency is determined by the relative amount of standby losses encountered which change based on where the water heater is installed and how much it is used. Consequently, the difference in erCOP and aCOP is the improvement offered by the HPWH.

Table 17 shows the erCOPs and suggests a reasonable rule of thumb for electric resistance tank performance on an annual basis is 0.82. The equipment names listed are used in Table 17 only to correspond to those in the previous tables. This is a theoretical table of what performance might be if all the water heaters, in all the installed configurations, used only resistance heat. By comparing the erCOP to the aCOP in Table 15, one can estimate the reduction in energy use achieved by the HPWH. For example, an aCOP of 1.6 uses roughly half the baseline energy while an aCOP of 2.0 uses roughly 40%.

Table 17. Annualized "erCOP" by Make, Location, and Heating Zone

	Basement		Garage		Interior	
Heating Zone 1						
Make	**Mean**	**n**	**Mean**	**n**	**Mean**	**n**
ATI	0.83	6	0.79	9	0.76	7
GeoSpring	0.83	4	0.85	15	0.86	2
Voltex	0.80	2	0.84	7	-	0
Heating Zone 2						
ATI	0.82	4	0.80	4	0.79	5
GeoSpring	0.86	2	0.86	1	0.76	1
Voltex	0.85	1	0.78	4	0.80	5
Heating Zone 3						
ATI	0.86	6	-	0	0.93	1
GeoSpring	0.87	2	-	0	0.83	2
Voltex	-	0	-	0	0.92	1
Overall						
ATI	0.84	16	0.79	13	0.79	13
GeoSpring	0.85	8	0.85	16	0.83	5
Voltex	0.82	3	0.82	11	0.82	6

3.7.4. Control Strategy Characteristics

The primary quantity of interest with respect to control strategy is resistance heat, as the key to achieving efficient water heating is avoiding costly invocations of the heating elements. Figure 19 demonstrates the range of energy use due to the control strategy for Voltex 60-gallon water heaters. The data are color-coded by fraction of resistance heat. Bright red indicates sole use of resistance elements, and blue the sole use of heat pump heat. It is clear from the graphic that control strategy – and the relative split between heat pump heat and resistance heat – plays a large role in determining the efficiency of this technology.

Figure 19. Daily Flow and Energy for 60 gallon Voltex

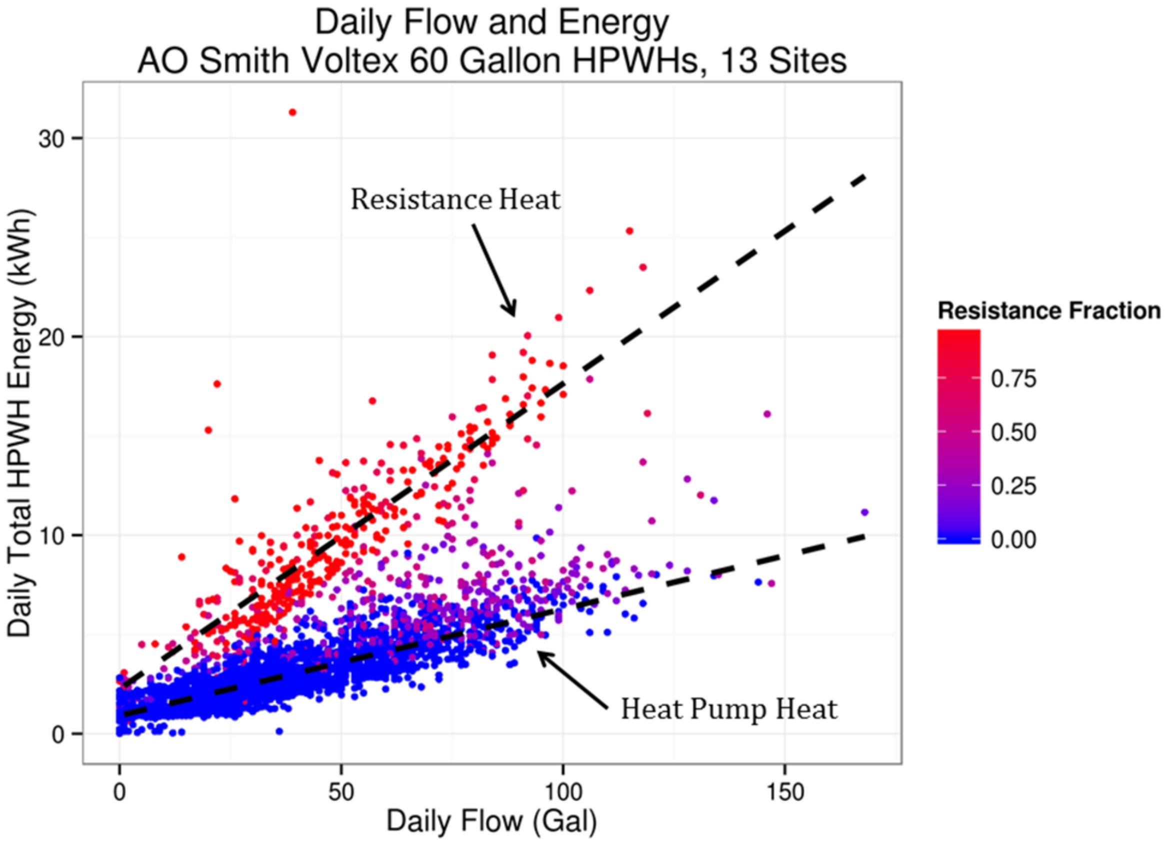


Table 18 shows annualized resistance heat fractions by equipment and install location. The resistance heat fraction is defined as the proportion of total input energy provided by resistance elements. As suggested by the lab research, the ATI showed the least amount of resistance heat, due to its ability to operate at cool space temperatures and hesitancy to invoke resistance heating (Larson and Logsdon 2012a). The Voltex units showed similarly small amounts of resistance heat in basement or interior installs, but the inability of the Voltex heat pump to operate below 45° F penalized its performance in garage installs. The GeoSpring water heaters showed the highest proportion of resistance heat, likely caused by the smaller capacity of a 50 gallon tank (most Voltex units in the study held 80 gallons, the ATIs 66 gallons), and a similar compressor lockout to the Voltex, where the heat pump does not operate with air temperature below 45° F (Larson and Logsdon 2012b).

Table 18. Annualized Resistance Heat Fractions

	Basement		Garage		Interior	
Make	**Mean**	**n**	**Mean**	**n**	**Mean**	**n**
ATI	6%	16	5%	13	4%	15
GeoSpring	25%	9	45%	17	28%	5
Voltex	4%	3	20%	13	10%	6

3.8. Space Heating Impacts

Space heating interaction is one of the primary questions regarding this generation of integrated, packaged heat pump water heaters, but is extraordinarily difficult to assess. A handful of sites were selected for co-heat tests (or flip-flop tests), in which the water heater was manually switched between heat pump mode and resistance heat mode in order to estimate the difference in space heating due to the heat pump water heater. These measurements proved inconclusive. The amount of daily variation in heating energy was large compared to the space heating effect of a heat pump water heater. In addition, we investigated the effect of HPWH runtime on evaporator entering air temperature. This provided an interesting display of the extent to which HPWHs cooled their immediate surroundings, but also did not offer usable estimates of space heating interaction.

3.8.1. Flip-Flop Sites

As described in section 2.3.3, five sites were set aside for a co-heat or flip-flop test, where the water heater was manually switched between heat pump and resistance modes, with the goal of assessing space heating impact. The idea was to estimate the annual impact by observing a heating signature through degree day regression in both operating modes, and then examining the difference when applied to a typical meteorological year (TMY). Refer to Appendix D: Space Conditioning Interaction for a detailed discussion of the method and the findings.

Overall, this test proved inconclusive. One of the five sites – a basement install in Woodland, WA – was ultimately unusable due to an error in the data-logging on the furnace. Of the remaining four, two lacked statistical significance, one showed unusual behavior that led to an unbelievable estimate, and only one showed a statistically significant, physically credible space heating impact.

The one site yielding credible results was #90051, an 80 gallon AO Smith Voltex water heater, installed in a Spokane, WA conditioned basement. Based on a weather-normalized analysis of the change in heating energy with outdoor temperature, the analysis shows heating energy was slightly higher with the water heater in heat pump mode as compared to resistance mode. We estimate the space heating impact of the HPWH at this site to be 1,500 kWh for a typical meteorological year, and that estimate is clearly distinguishable from noise in the regression.

Site 90051 uses an average of 55 gallons of hot water per day. Water use determines heat pump runtime which extracts heat from the air. The more water used, the more heat extracted. Therefore, we consider the finding at this one site in terms of interaction per gallon used: 27 kWh/yr per daily gallon. Previous energy modeling of an interior installation in the Spokane climate with an electric forced air furnace showed the interaction to be 37 kWh/yr/gal.[5] The previous work assumed that every unit of energy removed by the HPWH was made up again by the heating system, an HC_f of 1. At this particular site in Spokane, it appears that only 73% of the energy removed from the air was realized as a penalty at to the heating system energy. These are tentative findings and caution is warranted in applying them further.

---

[5] http://rtf.nwcouncil.org/measures/support/files/HPWH_interior_installs_94_v0_5.xlsm

3.8.2. Ambient Temperature Depression

A less direct means of assessing the space heating impact is to investigate the extent to which intake air temperature decreases during heat pump runtime. This did not lead to a directly usable estimate of space heating impact but it did show that HC_f should be less than 1. Further, it offered a view of the extent to which the HPWHs altered their surrounding temperatures.

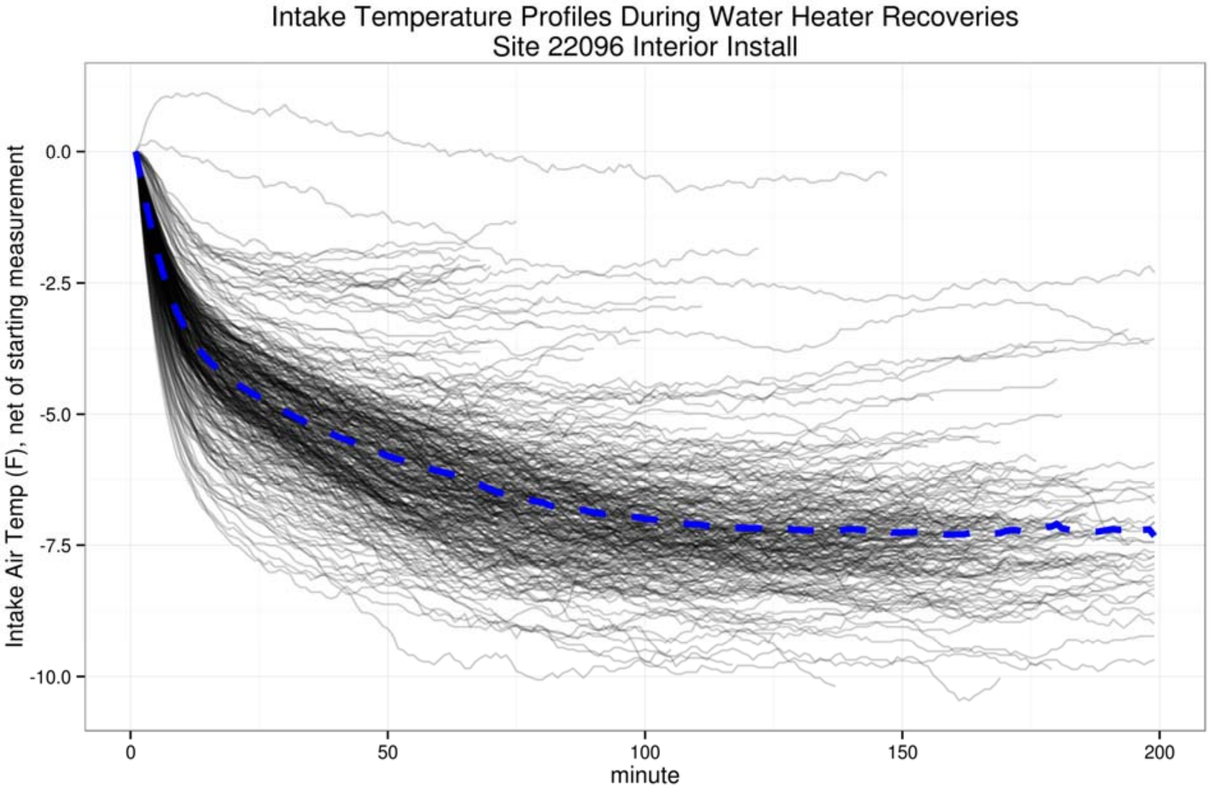
Figure 20 shows an example of this inquiry at a single site. The plot shows the measured intake air temperature during each heat pump recovery, relative to the starting intake air temperature. Each event starts at minute zero with “relative temperature” zero, and then the individual lines on the plot represent the intake air temperature profile along a single recovery event. The dashed, blue line shows the mean depression at each minute.[6] All temperatures are relative to the first valid temperature measurement of the recovery event. It is important to note that the beginning two minutes of intake air temperatures during the recovery events were discarded, as the air temperature measurements are only valid once air has begun flowing over the sensor. During HPWH inactivity, the presence of the nearby tank of hot water interferes with an accurate assessment of space temperature.

The pattern of Figure 20 was extremely common – an initial, rapid decrease, followed by a leveling off between one and two hours. The unit in question here was a 60 gallon unducted Voltex, installed within a vented, 180 cubic foot utility closet. The small volume of the install space likely explains the large depression of intake air temperature – ultimately reaching seven degrees on average. Most unducted units showed similar qualitative patterns, of a sharp initial decline followed by a more gradual decline. It is not clear the extent to which the abrupt decline during the first ten or so minutes represents the lag in air mixing within the space or heat extracted from the space.

---

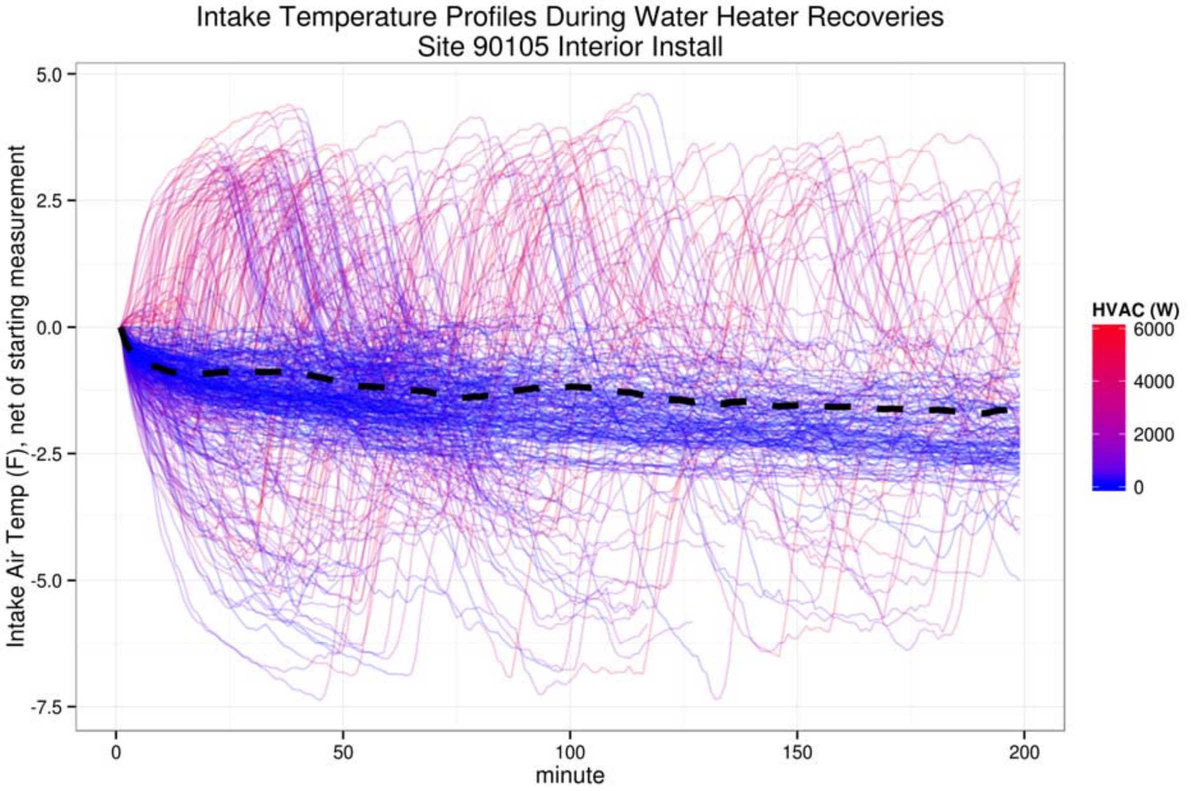
[6] This situation provides the unusual luxury of having so much data – having observed so many recovery events – that we don’t need to appeal to statistical modeling assumptions to derive a smooth curve of expected values. The dashed blue line is simply the mean observed value at each time point.

Figure 20. Temperature Depression During HPWH Operation, One Site



We frequently observed temperature oscillations, occurring due to the presence of a nearby heat source. See Figure 21 for an example. This graphic is similar to the previous, only with coloring according to heating system power draw. This site contained a GeoSpring water heater, installed in a 4,800 cubic foot laundry room, with zonal electric heating. Two distinct patterns become apparent from the graphic. One is the presence of red oscillations, which correspond to activations of the heating system warming the air entering the HPWH evaporator. There was no apparent correlation in time of the HPWH activation and heating system operation. The other pattern is a more orderly decline in intake air temperature with the space heating system off, colored in blue. Installed in a much larger space, the rate of decline – net of heating system interference – was much less abrupt in the unit of Figure 21 than that of Figure 20.

Figure 21. Temperature Depression During HPWH Operation, One Site with HVAC

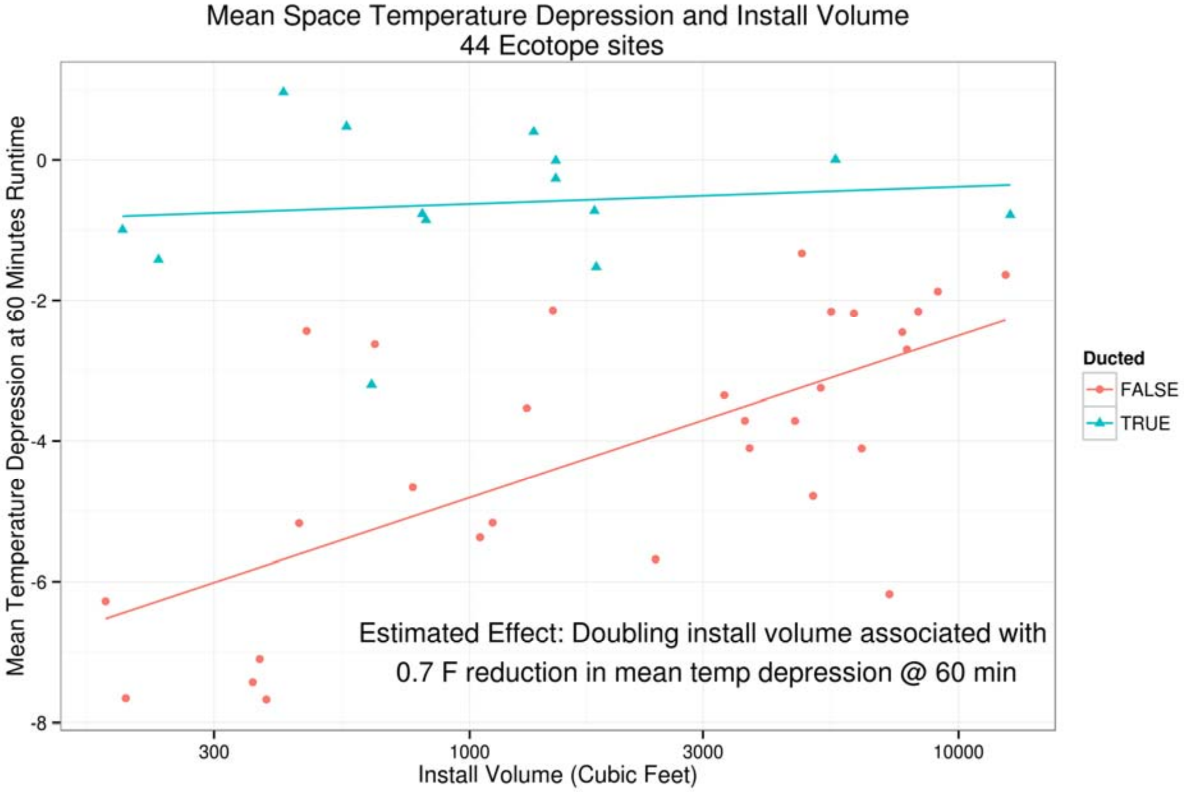


A logical investigation, given these graphics, is to pick a time interval and summarize the mean temperature depressions by install parameters, net of heating system intervention. For example, in the graphic of Figure 21, we would estimate the downward trajectory of the blue filaments, where the space heating system did not activate. Since much of the temperature depression occurred somewhere near one hour, we picked 60 minutes as the time period to consider. Typically, after one hour, or a little longer, the install location reached a steady-state, where further HPWH operation did not further reduce the evaporator entering air temperature.

Using this insight and barring heating system intervention, unducted units installed in extremely small spaces such as closets saw intake air temperature reductions around six or seven degrees Fahrenheit after one hour of heat pump water heater operation. Unducted units in large spaces saw temperature reductions between two and four degrees Fahrenheit after one hour of heat pump water heater operation. A regression suggested that doubling the volume of the install space resulted in a decrease of 0.7°F in the amount of temperature depression at 60 minutes of runtime (See Figure 22).

As expected, ducted units showed little relationship between temperature depression and install space volume. The mean temperature depression at 60 minutes runtime for the ducted units was not significantly different from zero. Those units penalize the heating system through additional infiltration load. Since we removed intervals containing central heating we would actually expect those units to see a slight intake air temperature decline over a one hour interval, as the whole house temperature floats downward due to the central heat being off (at least during the heating season).

Figure 22. Intake Air Temperature Depression During Water Heater Recoveries



To estimate space heating impact from the air temperature profiles, though, would require a detailed physical understanding of the heat transfer at work. Under some extremely general assumptions, it can be shown that interior installs – especially those in small spaces such as closets or utility rooms – theoretically cause exponentially decaying temperature reductions during HPWH runtime (as shown in Figure 20). The shape and severity of the ambient temperature decay profile in the water heater's install space are governed by the size of the room, the area of each exterior surface of the room, the thermal conductance of each exterior surface of the room, the permeability of those surfaces (or presence of vents or openings), the temperature of the main house (and also outdoor temperature or buffer space temperature, if appropriate), the rate of heat removal from the ambient surroundings, and the heat loss rate of the tank water to the ambient surroundings.

With an accurate accounting of the driving forces of heat transfer at a particular site, one could theoretically calculate expected temperature profiles under the assumptions of make-up heat totality – that every Joule removed from the space is made up by an extra Joule from the house heating system ($HC_f = 1$) – and under the assumption of no make-up heat – that the presence of the HPWH does not affect the house heating system ($HC_f = 0$). The actual measured air temperature profiles could be compared to the theoretical profiles under the two extremes to estimate a space heating interaction factor.

In the end, what we can assert from this investigation is that HPWHs installed in small spaces such as closets cool their surrounding air by roughly six or seven degrees after an hour of runtime, and HPWHs installed in much larger portions of the house cool their surrounding air by two to four degrees after an hour of runtime. The shape of the temperature depression profiles was found to approximate exponential decay (in the absence of space heat applied directly to the install space), as expected theoretically. Moreover, three crucial observations are that these temperature depressions (1) persist for long periods of time, (2) exist in the unducted installations but not the ducted ones (see Figure 8), and (3) are not directly correlated in time with heating system activation. Regarding the third item, the heat system does not directly respond to the water heater running. Therefore, the temperature of the install space is permanently lowered relative to the rest of the house. The lower temperature implies the heating system is not replacing all of the energy removed by the HPWH. Consequently, the heating and cooling interaction factor, HC_f, must be less than 1.

3.9. Noise Levels

A possible impediment to the acceptance of this technology is the increase in noise over a conventional electric resistance water heater. While an ERWH runs silent, the compressor and fan of a HPWH create a noticeable amount of noise. During installation visits, measurements were taken of the water heater noise in decibels (dBA). Five measurements were taken near the water heater according to the schematic of Figure 23. In addition, two measurements were taken in an adjacent room: one in the center of the adjacent room, and one near the wall shared with the HPWH room.

Figure 23. Sound measurement schematic

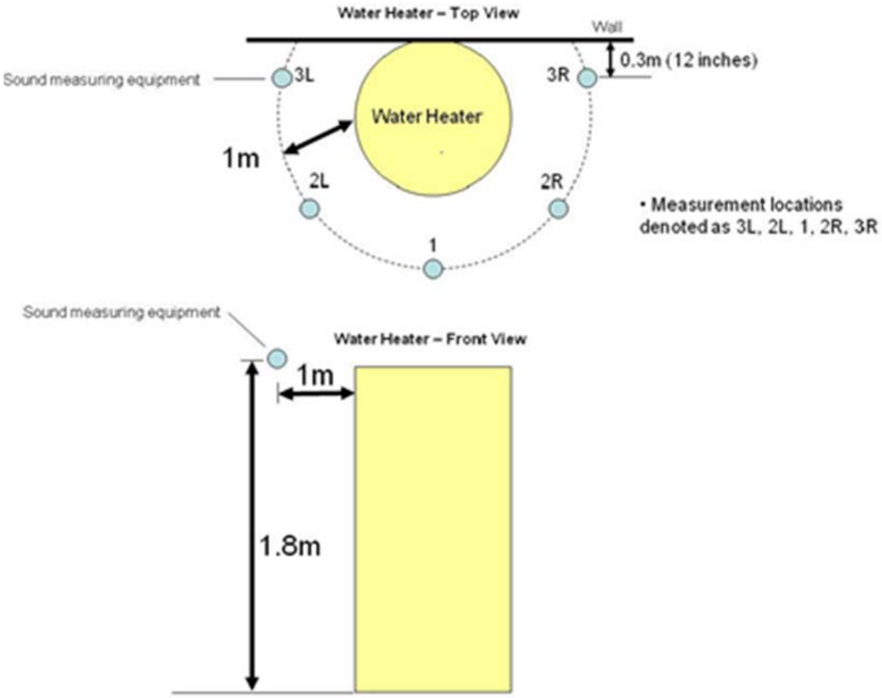


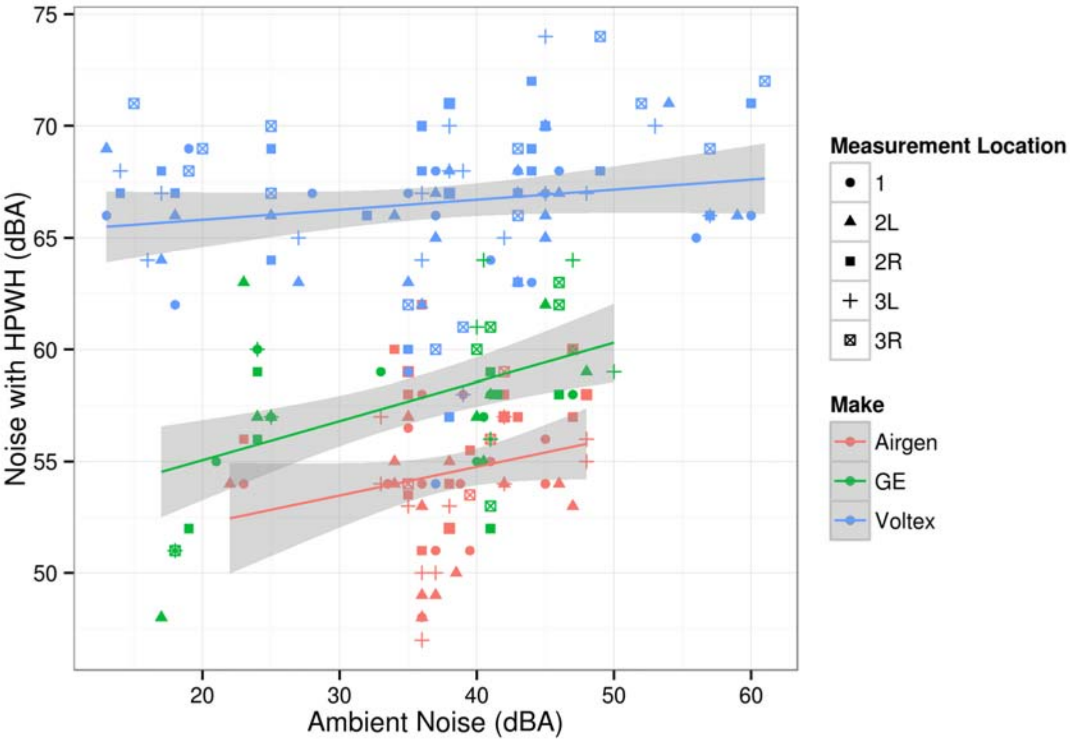
Table 19 summarizes the noise measurements of ambient background and HPWH decibels across water heater equipment for the adjacent room, install room, and from the lab testing(references). Across all cells the Voltex was measured as the noisiest water heater, both in terms of overall dBA and increase over ambient dBA. In the most relevant scenario of the adjacent room, the ATI water heater was measured as louder in an absolute sense than the GeoSpring, but, due to higher levels of background noise, created a smaller increase over ambient. In the field measurements from the install room the GeoSpring was both noisier in an absolute sense as well as causing a larger increase over background noise than the ATI, however this was reverse of the finding from the lab testing. Average noise in the same room as the HPWH ranged between roughly 55 and 65 dBA, while noise in an adjacent room ranged between roughly 35 and 45 dBA.

Table 19, Noise level in dBA

Noise in dBA	Ambient	HPWH	Difference	n
Adjacent Room				
Voltex	37.1	46.4	9.3	21
GeoSpring	30.7	37.5	6.8	10
ATI	37.8	41.2	3.4	15
Install Room				
Voltex	37.1	66.6	29.5	21
GeoSpring	35.3	57.7	22.4	10
ATI	38.4	54.6	16.2	15
Lab Testing				
Voltex	31.8	63.2	31.5	1
GeoSpring	37.9	54.6	16.7	1
ATI	33.6	58.5	24.9	1

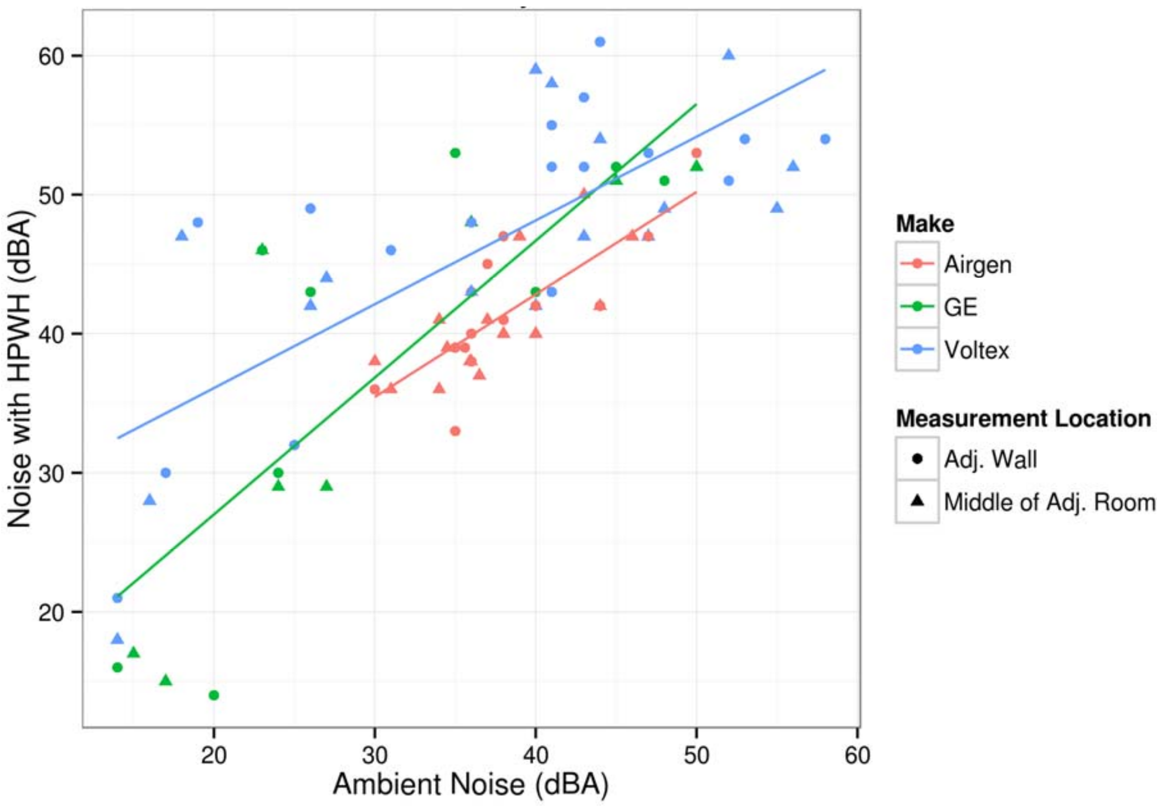
For a graphical view, Figure 24 shows measured noise levels in the install room, colored and shaped by water heater make and measurement location. The graphic shows that, in the install room, the water heater noise mostly swamps whatever ambient noise already existed. The Voltex was measured as the noisiest unit, with install room decibel readings typically around 65 dBA, and ranging mostly between 60 and 75 dBA. The ATI was the quietest with typical readings near 55 dBA, and a range from 50 to 60 dBA. The GeoSpring fell in between with typical readings between 55 and 60 dBA.

Figure 24. Ambient and HPWH Noise in Install Room



In most install scenarios, located in garages, basements, or utility closets, however, the noise levels in an adjacent room are more relevant in assessing the effects of noise on occupants. Figure 25 shows HPWH noise measured in an adjacent room, colored and shaped by make and measurement location. Whereas in the install room the water heater overwhelmed the ambient noise, in an adjacent room the decibels during HPWH operation depended strongly on the ambient decibels. The ATI and Voltex showed similar patterns of HPWH noise over ambient, with the Voltex again measured as the louder unit. The pattern of the GeoSpring noise measurements was influenced by two installs adjacent to nearly silent rooms (~ 20 dBA) that were barely affected by the HPWH.

Figure 25. Ambient and HPWH noise in adjacent room



3.9.1. Occupant Satisfaction with Noise Levels

During site decommissioning, homeowners were asked to rate their satisfaction with the noise level of the HPWH on a scale of 1-5, with 1 indicating "very dissatisfied" and 5 indicating "very satisfied." The mean satisfaction along with adjacent room noise averages are displayed in Table 20. Homeowners were most satisfied with the noise level of the GeoSpring, with mean satisfaction of 4.3, and least satisfied with the noise level of the Voltex, with mean satisfaction of 3.6.

Table 20. Noise and Occupant Satisfaction

Make	Noise in dBA (Adjacent Room)			Satisfaction
	Ambient	HPWH	Difference	(1-5 scale)
GeoSpring	30.7	37.5	6.8	4.3
ATI	38.1	43.4	5.3	4.0
Voltex	37.2	46.5	9.3	3.6

Average sound level satisfaction aligns well with the measured decibels in the adjacent room. Appendix G: Assessing Noise Satisfaction, investigates further to determine if the finding is significant or coincidental and examines other factors that may influence sound satisfaction. The assessment showed no significant correlation with HPWH make. It did show that houses with higher average draw volumes were less satisfied with the noise, although, oddly, there was not a strong relationship between runtime (which is depends on draw volume) and satisfaction.

4. Extended and Generalized Findings

This section extends all the measured findings to a more generalized case. For example, inlet water temperature profiles that are applicable to anywhere in the Northwest will be developed. This section is an analytic extension of the “raw” findings of section 3. Consequently, many of the same topics are covered albeit in a distinctly different way. The major goal of generalizing the findings is to develop results, grounded in field measurements, which can be used to build or run numerical simulations. Those simulations are the crucial step from translating the findings in the engineered field sample to the population of houses at large. This section concludes with simulation output showing the energy use and savings estimates of HPWHs as validated by this study.

4.1. Inlet Water Temperature

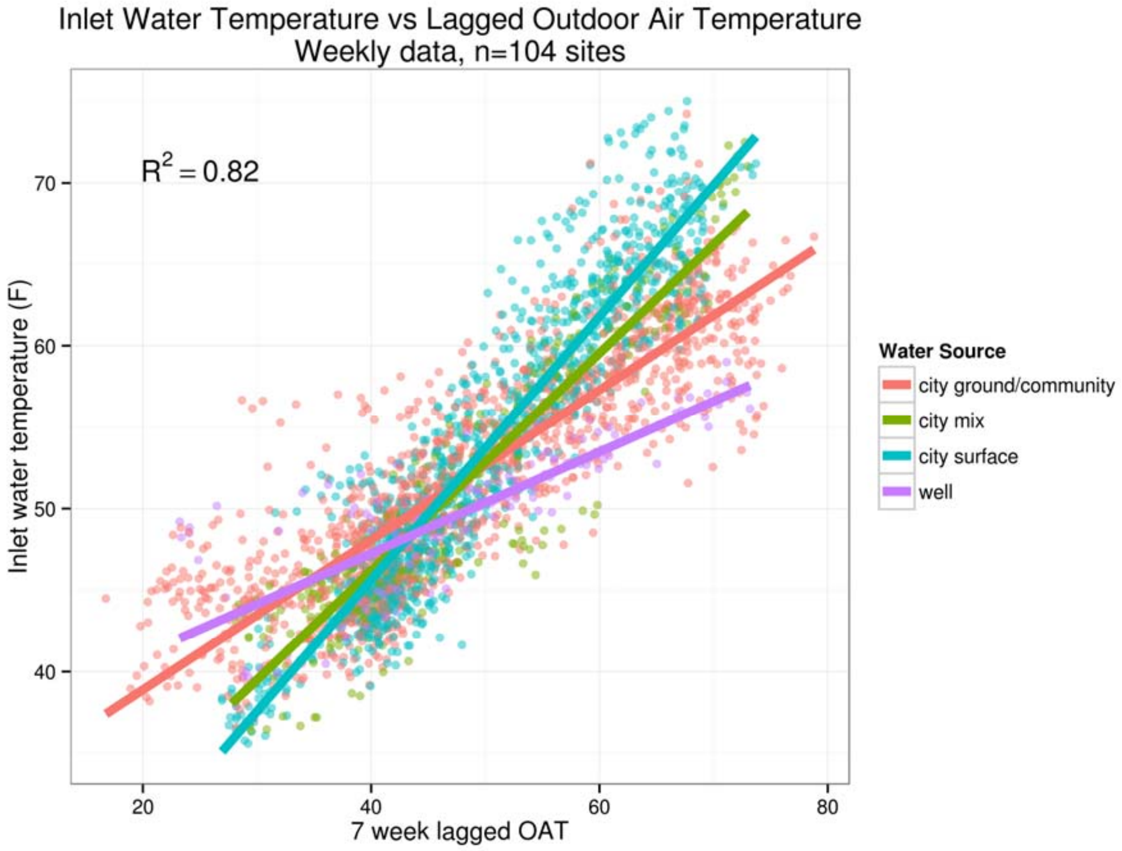
This section details the methodology to develop generic inlet water temperature curves for the Pacific Northwest, to be used in the simulation. The two main determinants of inlet water temperature are climate and water source. As such, the sites were recorded for nearest climate station, and classified into five water supply categories:

1. City Surface. A city water system with primarily surface water sources. Examples include Seattle and the Cedar River Watershed, or Portland and the Bull Run Watershed.
2. City Mix. A city water system that draws from both surface water and ground water sources. Examples include Bend, Oregon, drawing from the Deschutes Regional Aquifer and Bridge Creek.
3. City Ground. A city water system that draws primarily from ground water. Examples include Spokane, Washington and the Spokane Valley - Rathdrum Prairie Aquifer.
4. Small Community Water System. A water system, encountered in rural towns, drawing water from an underground source, with a small distribution network, to serve a small number of connections (on the order of tens to hundreds) in contrast to city water systems.
5. Well. A home which draws from a local underground source.

The climatic effect was assessed through a rolling average outdoor temperature. It was assumed that changes in inlet water temperature responded linearly to changes in moving average outdoor temperature, with the elasticity of the change determined by water source and the optimal window size for averaging selected from the data. Logically, surface water sources should display the most sensitivity to changes in climate, and well sources the least, with city groundwater, community systems, and cities that draw a mix of ground and surface water lying in between.

Figure 26 shows the data and regression lines for the inlet water temperature profiles. We considered moving windows to average outdoor temperature between one and three months, with the optimal window size estimated at 7 weeks, which is plotted in the graphic. Due to similarities between city ground and small community sources – both physically and realized in the data set – those two sources were collapsed into a single category.

Figure 26. Inlet Water Temperature Profile Regression



As evident from the graph of Figure 26, the data did not fit into these categories as tidily as one may like, but the overall means seem appropriate. City groundwater systems showed the greatest heterogeneity, as water may be drawn from aquifers of different depth, and the groundwater may be stored in above-ground tanks for differing periods in different cities. In short, the graphic shows plenty of variation, but the mean response of inlet temperature versus climate appears plausible and usable.

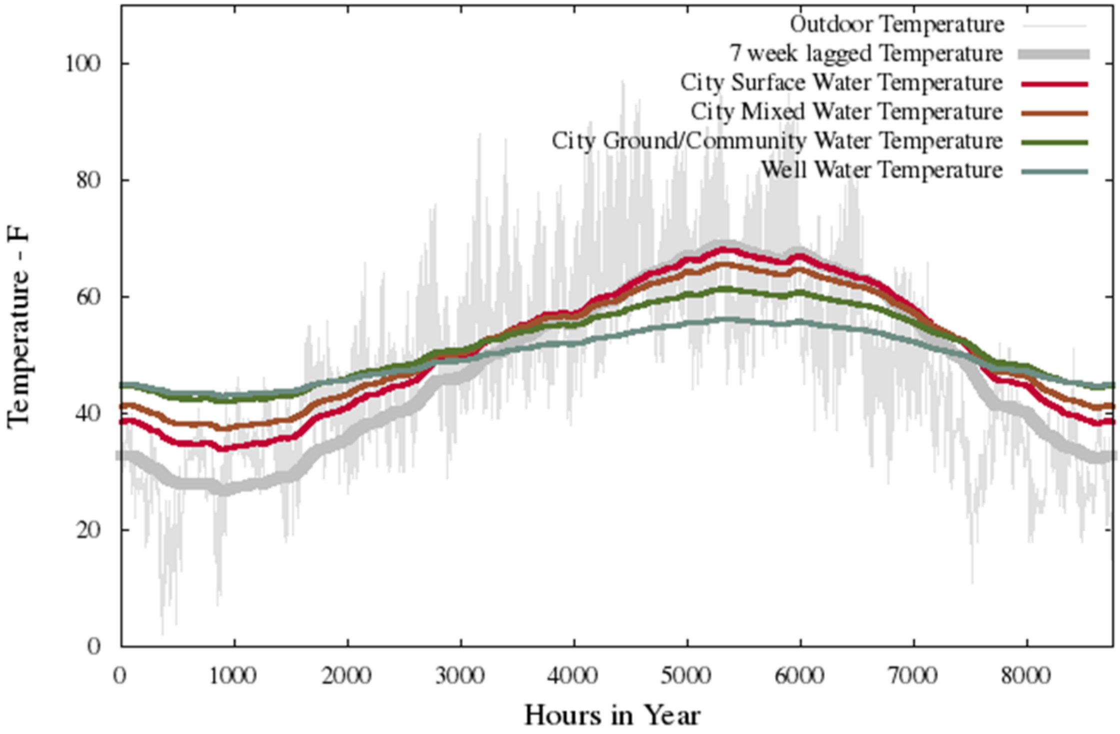
The coefficients to generate the lines of Figure 26 are shown in Table 21. As is evident on the graph, and as one would expect, surface water showed the highest responsiveness to changes in air temperature, with a slope of 0.81. This implies that an increase in one degree of rolling average outdoor air temperature corresponded with a 0.81 degree increase in water temperature. Wells showed the least responsiveness to changes in air temperature, also as expected, with a slope of 0.31: an increase of one degree in rolling average outdoor air temperature corresponded with a 0.31 degree increase in water temperature. While a necessary part of the regression to predict water temperature the intercepts alone, do not offer a similarly nice interpretation, as they represent the somewhat useless quantity of expected water temperature given a 7-week average air temperature of zero degrees Fahrenheit.

Table 21. Linear Coefficients of Inlet Water Temperature Against 7-Week Average Rolling Outdoor Temperature

	Intercept (°F)	Slope (°F/7 wk lag)
City Surface	13.3	0.81
City Mix	19.5	0.67
City Ground/Community	29.7	0.46
Well	34.8	0.31

As an example application of this model for developing inlet water profiles, consider Figure 27. This graphic displays Typical Meteorological Year (TMY3) weather data for Spokane, Washington, along with the lagged outdoor temperature and predicted inlet water temperature profiles for each water system. This output is developed as an 8760 profile – one predicted value for each hour in the year. Note that the hypothetical "city surface" system shows the greatest elasticity with respect to outdoor temperature[7], while the private well system shows the least.

Figure 27. Example Inlet Water Temperature Profiles, Spokane Climate



4.2. Outlet Water Temperature and Tank Set Point

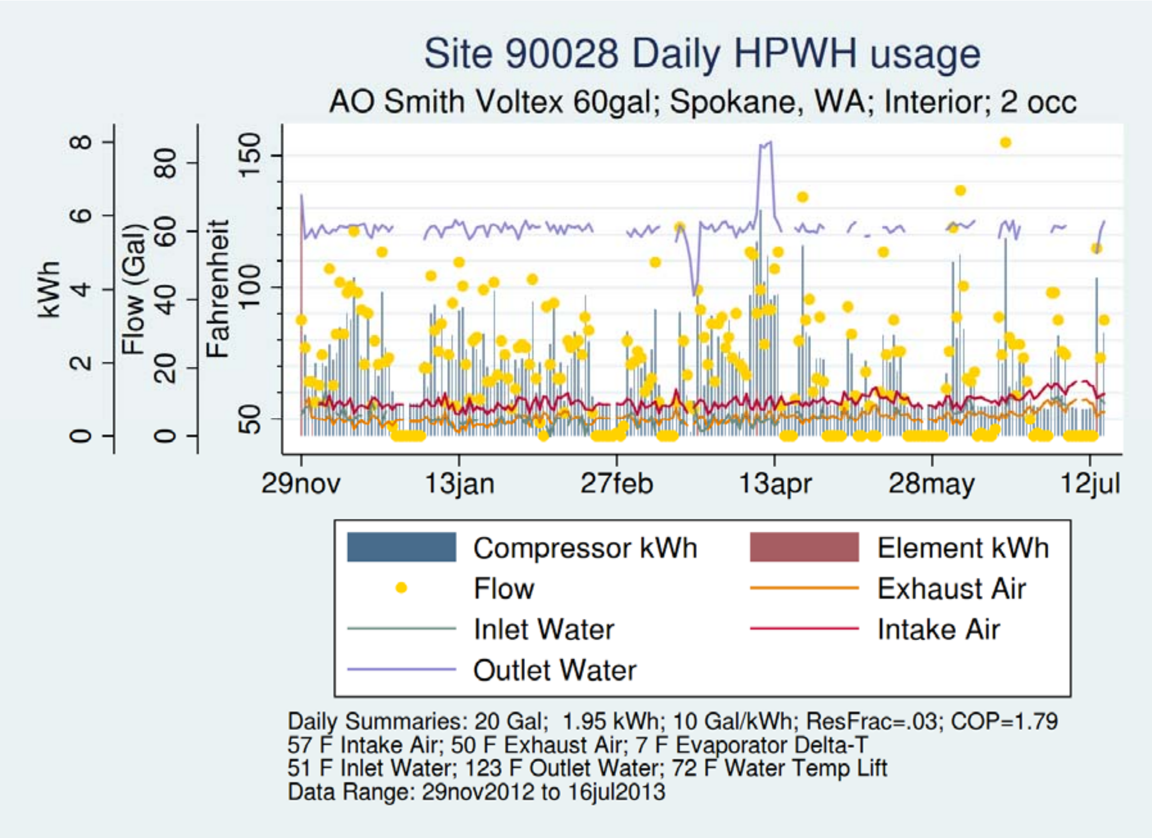
To develop a tank setpoint input for the simulation, we considered for each site the 95% quantile measurement of outlet water temperature. While it may seem preferable to use the maximum

---

[7] Spokane actually draws water from the extensively documented Spokane Valley-Rathdrum Prairie Aquifer. The different profiles presented here are merely illustrative of the simulation inputs.

observed outlet water temperature as our estimate of the setpoint at a site, the maximum is much more susceptible to vagaries in the data. The 95% quantile estimator is more robust to unusual occurrences. See Figure 28 for an example of this – the homeowner appeared to increase the setpoint from approximately 125° F to 150° F for a few days in April. Using the maximum observed outlet water temperature to estimate the setpoint would lead us to conclude on this site a setpoint of 150° F. That is clearly not a good estimate of the setpoint at this site, and the more robust 95% quantile estimator instead assigns a 127° F setpoint, which much more accurately describes the behavior.

Figure 28. An Example Where the Maximum Temperature Estimator Fails



Daily Summaries: 20 Gal; 1.95 kWh; 10 Gal/kWh; ResFrac=.03; COP=1.79
57 F Intake Air; 50 F Exhaust Air; 7 F Evaporator Delta-T
51 F Inlet Water; 123 F Outlet Water; 72 F Water Temp Lift
Data Range: 29nov2012 to 16jul2013

We investigated whether setpoint showed correlation with either water heater make, or number of occupants (see Figure 29 and Figure 30). No correlations were found, and so we use the mean setpoint of 128° F in the simulations. Note that the setpoint is necessarily different from the average delivered water temperature (section 3.4) primarily because the tank temperature floats in a "deadband" below the setpoint.

Figure 29. Setpoint as a Function of Water Heater Make

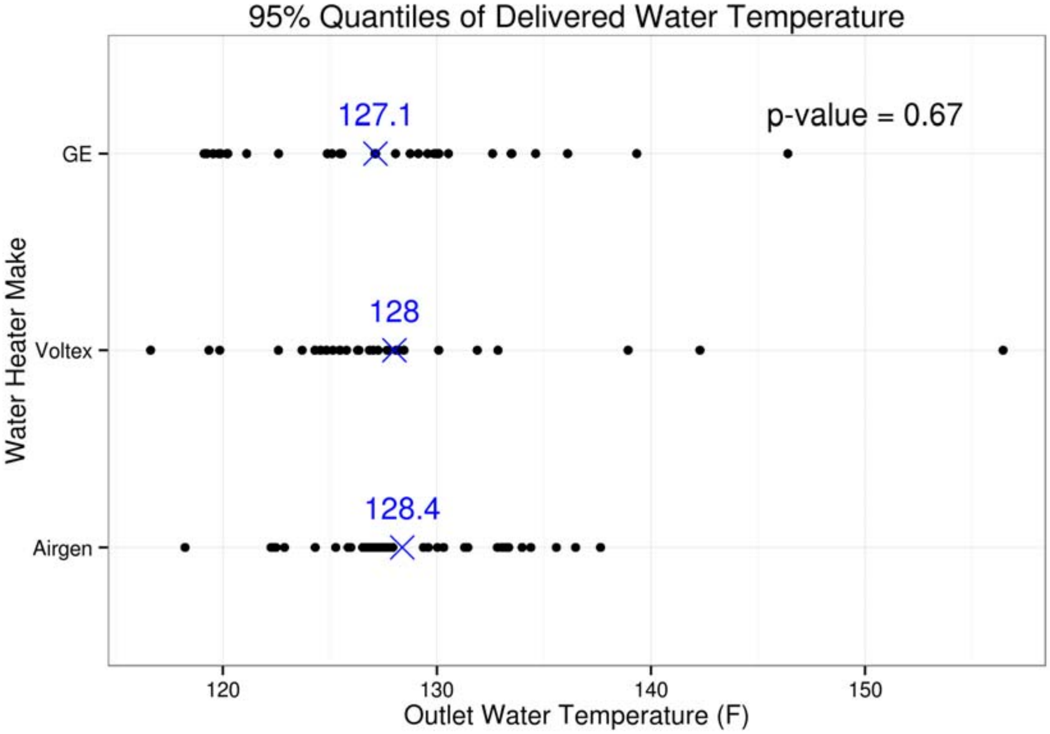
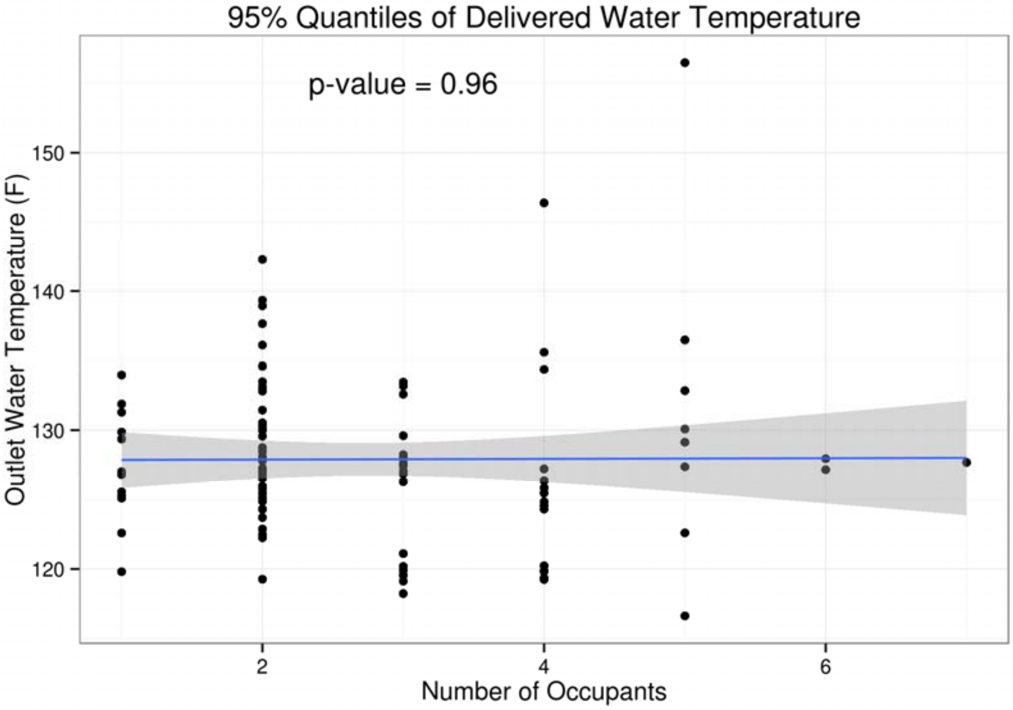


Figure 30. Setpoint as a Function of Occupancy

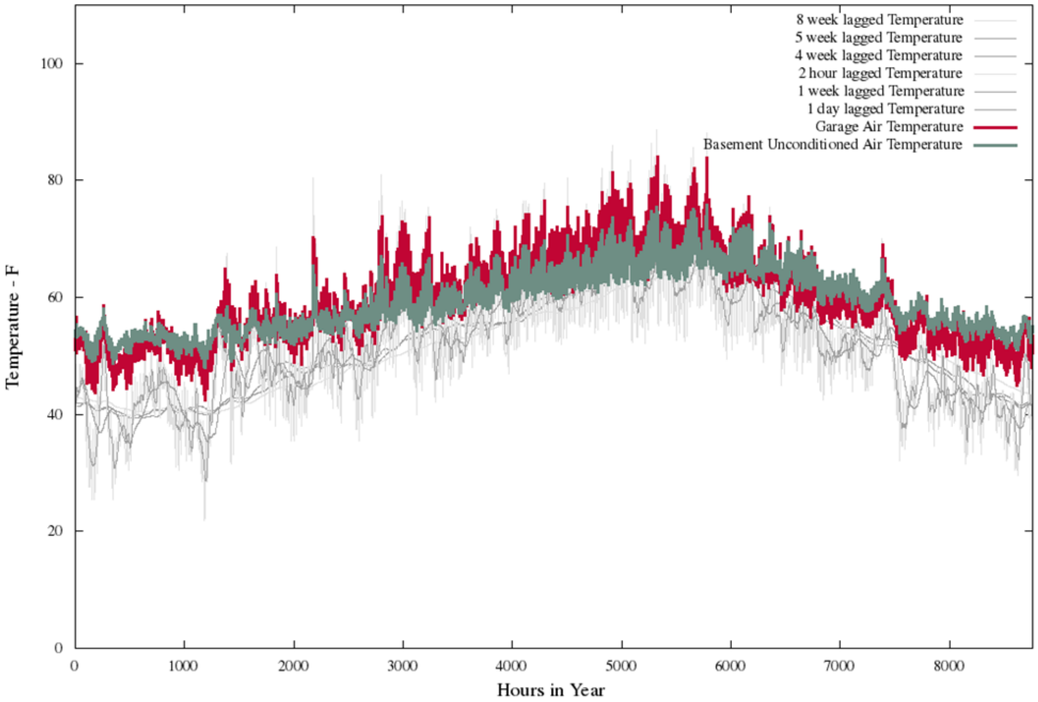


4.3. Ambient Space Temperatures

The generalized ambient space temperatures are derived from the measurements. There are four cases to consider: garages, basements, interior ducted, and interior non-ducted. Each installation scenario has a different, if only slightly, temperature regime. Previous work to estimate temperatures relied on simulations and assumptions about insulation levels and how well thermally connected to the house garages and basements are (Larson 2011a, RTF 2011). With measured data, we can create empirically derived relationships between garages, basements, and outside temperatures obviating the need for simulated assumptions.

Garage temperatures were observed to lag behind the outside air temperature and to also be significantly buffered by their contact with the house and the ground. Consequently, we developed a functional fit using hourly temperature data to predict the garage temperature. The garage temperature changes not only throughout the year but also has a diurnal variation, although largely damped, to account for daily temperature swings. A similar approach is taken for basements. Figure 31 displays the outdoor temperature from TMY3 data for Seattle and the associated garage and unheated basement temperatures as calculated.

Figure 31. Garage and Unheated Basement Temperature Predictions for Seattle



Equation 6 and Table 22 provide the method to calculate the garage and unheated basement temperatures for use in a simulation. To use the equations, calculate the trailing average of the outdoor temperature for the given length of time. For example, the four week lagged term is the average temperature over the previous four weeks.

$$\boldsymbol{T_{garage,basement} = c0 + c1 * lag_{8wk} + c2 * lag_{4wk} + c3 * lag_{1wk}}$$

$$\boldsymbol{+c4 * lag_{1day} + c5 * lag_{2hr} + c6 * lag_{0}} \qquad \textbf{Equation 6.}$$

Table 22. Garage and Unheated Basement Temperature Prediction Coefficients

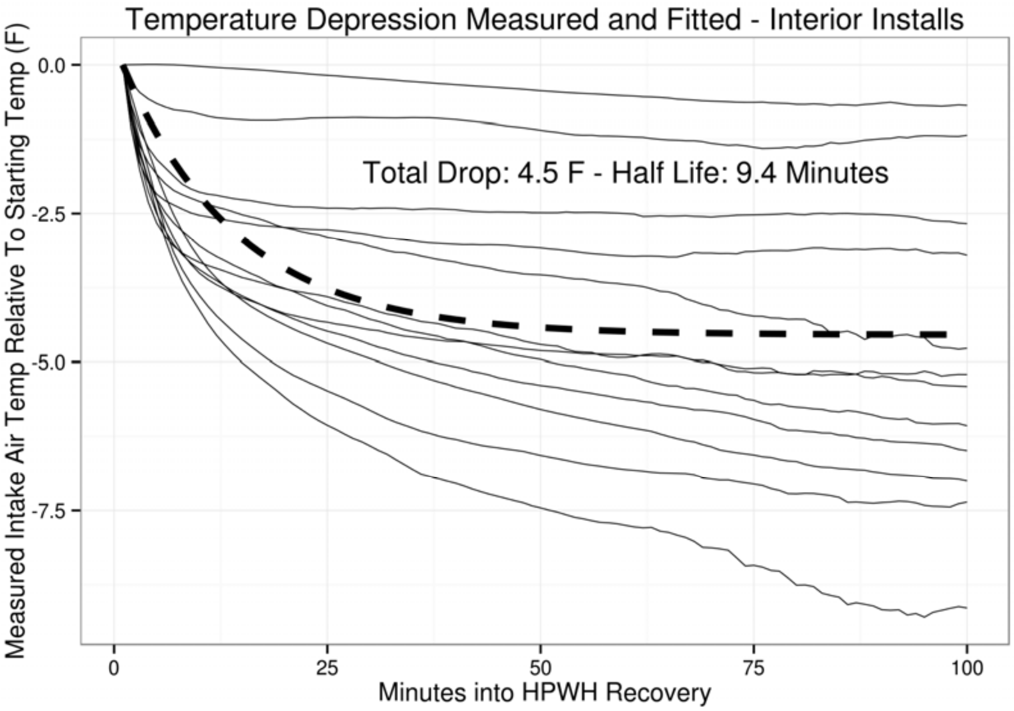
Coefficient	Garage	Basement	Length of Lag
c0	19.30326	40.32471	-
c1	0	0.244355	8 week
c2	0.20457	0	4 week
c3	0.089434	-0.08328	1 week
c4	0.166076	0.03047	1 day
c5	1.589119	0.354972	2 hour
c6	-1.29351	-0.19954	current

Generalizing interior space temperatures for use in a whole-house simulation presents a different situation from the buffer zones. Inside the house, the water heater is obviously more connected to the thermostat-regulated space temperature. For ducted cases, intake air temperatures measured in the upper 60s F in winter to lower 70s F in summer – exactly the temperatures one would expect. Consequently, it is possible to directly use the simulated space temperatures as the intake air temperature.

For non-ducted cases, section 3.8.2 showed the ambient air temperature was below what would be expected from the rest of the house. Therefore, we calculate an adjusted ambient air temperature which is used by the HPWH simulation. This bypasses the air temperature of the heating/cooling simulation but is necessary to accurately model the water heater. Figure 32 shows how the intake air changes after the HPWH turns on. Each solid line is the average observed temperature change for each site. The dashed line is the fit to all the sites. It is an exponential function decaying to a final temperature drop of 4.5F with a half-life of 9.4 minutes (see Equation 6 which calculates the temperature change as a function of the number of minutes, *t*, that the heat pump has been running). As is observed in field data, once the water heater turns off, the space temperature in the simulation rebounds in a mirrored way.

$$\boldsymbol{T_{dep}(t) = 4.5e^{-t*\frac{ln(2)}{9.4}} - 4.5} \qquad \textbf{Equation 7.}$$

Figure 32. Interior, Unducted Space Temperature Depression



4.4. Space Heating Interaction

The exploratory work on space heating impact – while perhaps scientifically interesting – did not illuminate the changes to space heating loads. Thus, we fall back on engineering experience and judgment to estimate the interaction. Ecotope worked with RTF staff and the RTF HPWH subcommittee (the “subcommittee”) to assess all available information on the interaction and decide upon a number to use until more research can be conducted. As throughout the report, the subcomittee considered the four scenarios of interest: garages, basements, interior recirculating, and interior exhaust ducted.

Of particular help in assessing the space heating interaction was a recent experiment using side-by-side manufactured homes (Widder 2014). The experiment measured the difference in heating load in houses identical albeit in their water heater configurations. The HPWH was always installed in the conditioned space and was exhaust ducted or not depending on the test. The experiment found HC_f = 0.49 for the interior, unducted installation and HC_f = 0.44 for the interior, exhaust ducted installation. These findings apply for this specific house configuration but it is unclear how to extend them to a general population of houses. The amount of “thermal coupling” or interaction between the HPWH and the heating system thermostat will depend on how close the thermostat is to the water heater and how temperature changes in one zone of the house ultimately propagate to other locations.

The outcome of the subcommittee work for each of the four installation scenarios is as follows:

- Garages. Although previous modeling work attributed some small, non-zero interaction to garage installations, the review of field measurements suggests that the interaction may not differ noticeably from zero. Following the rule that simple assumptions are the best in the face of inconclusive data, we use $HC_f = 0$ for garage installations.

- Basements. To clarify, installs in conditioned basements are expected to act like installs in any other conditioned part of the house. This category is for unconditioned basements. The unconditioned basement is more thermally coupled to the house than the garage. Nevertheless, changes in air temperature when the HPWH runs appear to be relatively minor. As with garages, in lieu of definitive data, we suggest the simplest assumption which is that $HC_f = 0$ for unheated basement installations.
- Interior with recirculating air (no exhaust ducting). The investigation of interior temperature change demonstrates that HC_f must be less than 1. Exploratory calculations of the change in house heating requirements for a reduced temperature in the HPWH install zone show that they could account for 10% of the interaction. That is HC_f could be no larger than 90%. The subcommittee ultimately chose $HC_f = 0.65$.
- Interior with exhaust ducting. The interaction is indirect through changes in house infiltration. When the HPWH runs, it acts like an exhaust fan to the rest of the house. As is the case without ducting though, the house has many zones and an increase in infiltration in one location may not be noticed by the thermostat in another. To keep with the simplest concepts, the subcommittee chose the same interaction factor as for unducted installs, $HC_f = 0\ .65$. Certainly, the heat transfers at play in the two interior installation scenarios are different but, for now, the results are estimated to be the same.

In addition to the providing judgment on the heating interaction, the subcommittee also recommended three possible avenues for future research to better measure HC_f:

1) A new experiment with the side-by-side lab homes to measure the range of thermal coupling by placing a known heat (or cool) source in four different locations throughout the houses and observing the heating system response. Possible locations to consider are the main living space near the thermostat, the master bathroom (a distant zone from the thermostat), a kitchen, and a bedroom.
2) A more detailed paper study to use existing data and models to help bound the interaction factor. Possible approaches include examining correlations with HPWH space temperature depression, hand calculations balancing heat flows to the HPWH zone from the interior with the heat flow to the exterior and heat extracted by the HPWH, and more informed modeling.
3) A large scale version of the flip-flop field study. Perhaps with enough sites (100-200), and data collection spanning an entire year, HC_f could be measured directly.

4.5. Hot Water Draw Patterns

We wish to create generalized draw patterns for use in a simulation to predict annual hot water energy use. Draw volume depends heavily on the number of occupants per house. Further, HPWHs have two heating systems (heat pump and resistance) with two drastically different efficiencies. The programed control strategy responds to the draw pattern to turn on one or another of the water heating systems. Large draws stress the water heater and have a tendency to trigger the use of resistance heat. Therefore, any generalized draw patterns should have periods of large and small draws in proportion to those observed in the field data. Further, we expect those profiles to occasionally trigger the resistance heat in similar proportion to the observed data.

In June 2014, the Department of Energy released the final version of the updated test procedure for residential water heaters (DOE 2014). The new procedure prescribes four, new draw patterns for the 24-hour simulated use test in contrast to the previous procedure's single pattern. The patterns are "point of use" at 10 gallons, low at 38 gallons, medium at 55 gallons, and high at 84 gallons. The "point of use" pattern is irrelevant for the purposes of this study. In the other three patterns, the draws are clustered in to three event groups throughout the day.

Overall, the DOE draw patterns don't appear to agree with the field data collected in the Northwest. First, the total daily draw volume for the DOE draws skew high. Second, the number of event clusters is lower than that observed in the field. Furthermore, in the context of creating credible energy use simulations we find it necessary to use more than a repeating, single day draw pattern to capture the variability seen in the field. Moreover, we wish to create draw profiles that are associated with an occupant count to inform future simulations as the housing stock changes. The DOE draw patterns aren't assigned a particular number of occupants. Consequently, we proceed in crafting draw profiles rooted in field measurements collect in the Northwest.

Due to the variability in draws, we created profiles for a range of household sizes (1, 2, 3, 4, and 5+). Thus, a future analyst can run simulations with all the different sizes and weight the output together in differing proportions relative to their saturation in the housing stock. Weighting results on output is a simpler way to produce an estimate of energy use for the average household size of 2.7 people without creating a separate profile for a fractional number of occupants.

Using the clustered draw analysis technique, we produced generalized draw patterns for 1, 2, 3, 4, and 5+ occupant households. Table 12 and Table 26 (in the appendix) provide the information for constructing the daily profiles. The characteristics of total flow, size, number, time, and duration of draws dictated how each profile was created. For instance, Table 12 shows three occupant households use 46 gallons per day separated in to 5 distinct event clusters. Analysis of all the sites shows when and for how long those events occur. Figure 33 visually presents the draw profile at a 3-person household. Specifically, the last event cluster is centered at 10pm, spans 43 minutes, and uses 6 gallons in 2 draw events. In the figure, the height of the line corresponds to how many gallons are drawn in the given minute.

Figure 33. Typical Daily Draw Profile, 3 Occupants

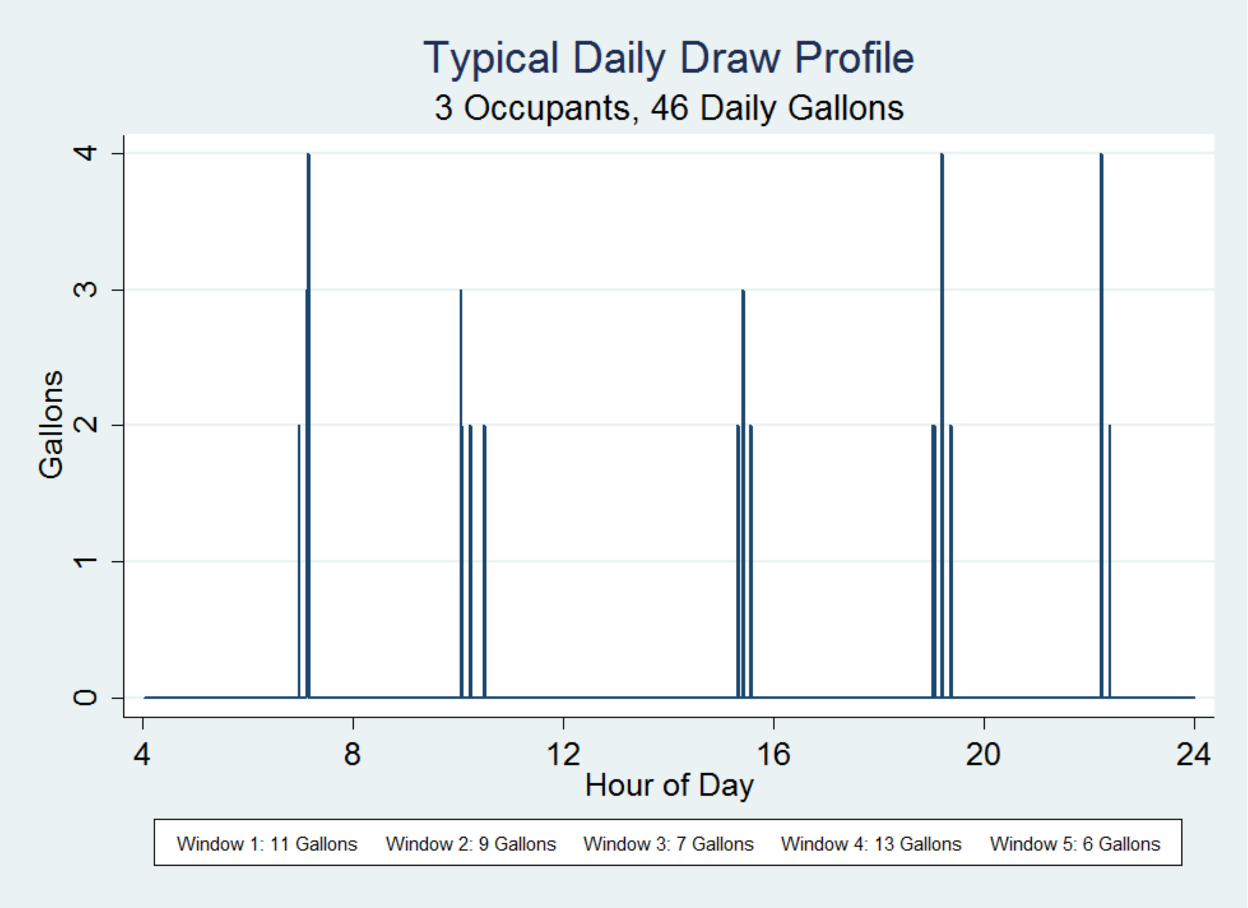
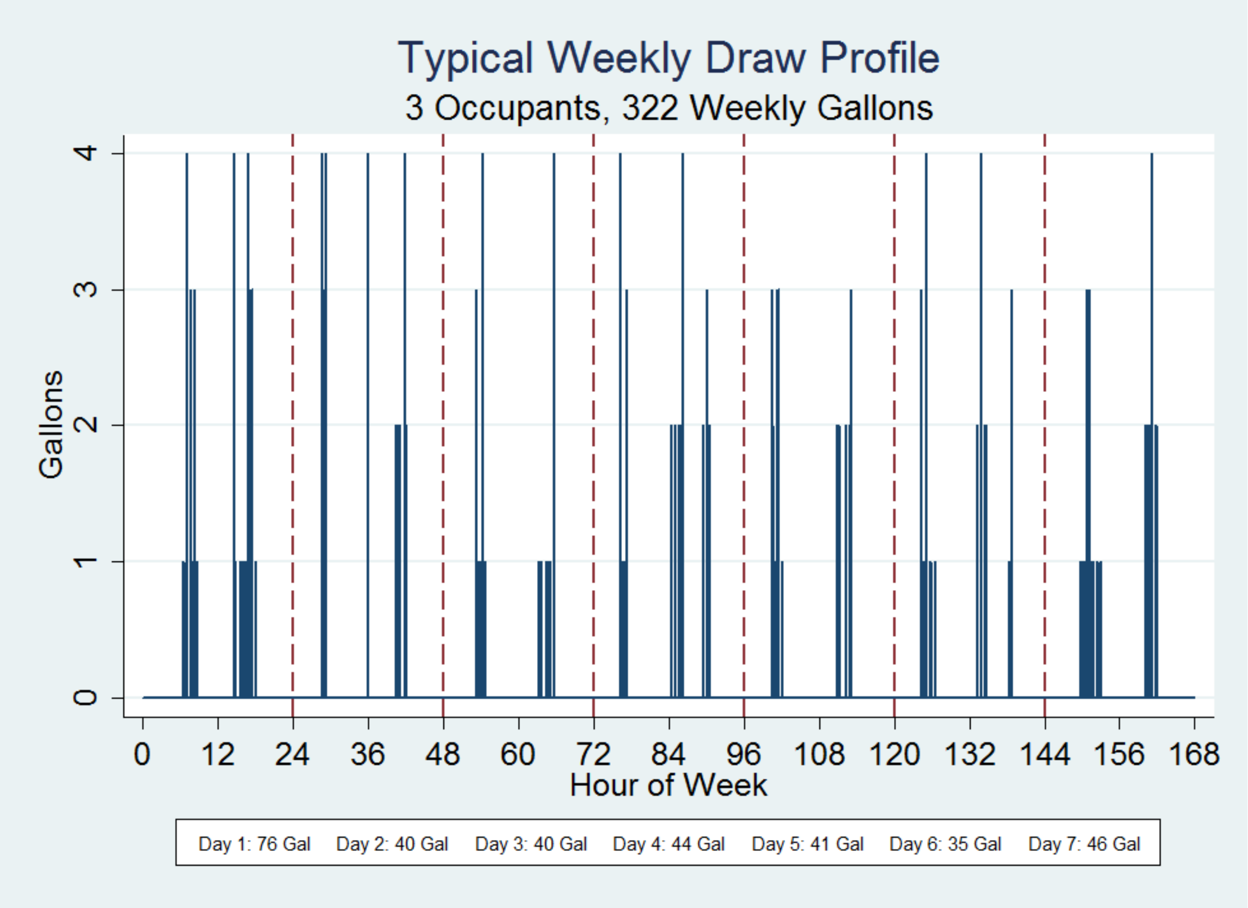


Figure 34 takes the draw profile development to the next level showing a typical pattern over seven days for three occupants. Investigations of the data show that the variability in draws can't be accurately captured (at least for simulation purposes) in one 24-hour period. Instead, it is necessary to use a week's worth of draws. As expected, the week in Figure 34 shows some days with heavier use and other days with lighter use. For instance, "Day 1" of the week uses 76 gallons while most of the other days range near the average from 40-46 gallons. Over the course of the week, the average daily flow remains at 46 gallons. For the specific draw schedule information refer to Appendix E: Draw Profiles.

Figure 34. Typical Weekly Draw Profile, 3 Occupants



All of the draw profiles were crafted directly from the characteristics observed in the field data (detailed in Appendix E: Draw Profiles). The exception is for the “Five+ Occupant” schedules which are additionally informed by the need to have at least one, large draw pattern for simulation purposes. The average daily draw volume of 5+ occupant households in the field sample of 72 gallons was for that specific distribution of household sizes. The distribution in the general population is different and can have more people and more water use. Further, a calibration exercise overseen by the RTF HPWH evaluation subcommittee concluded that having a larger draw available would more accurately simulate the amount of resistance heat use observed in the field data. Consequently, the 5+ occupant draw for both the daily and weekly basis uses 85 gallons per day on average which is increased from the observed 72 gallons per day.

4.6. Energy Use and Savings

The ultimate goal in generalizing the HPWH field study findings is to use them as inputs to the engineering model of water heater performance. The model, developed by Ecotope separately from this report, has been integrated with the SEEM residential energy simulation tool. Together, the HPWH model and SEEM simulation can be used to predict the energy use of electric water heaters across the region.

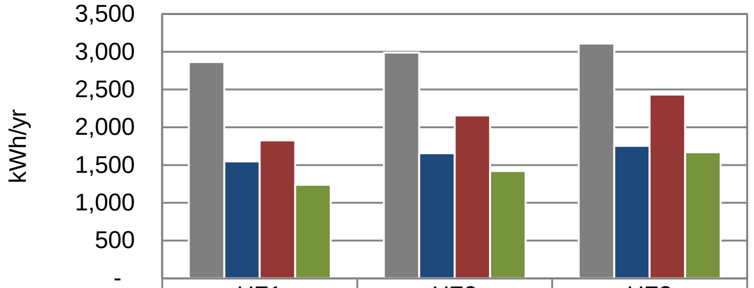
The final, generalized energy use and savings are comprised of simulations for a number of scenarios. Those include all combinations of make (GeoSpring, Voltex, ATI) to be aggregated in to the Northern Climate Specification Tiers, heating zone (HZ1, HZ2, HZ3), and installation location (garage, basement, interior, interior ducted). These scenarios describe how much energy the water heater alone uses. Over those runs, it is necessary to layer the heating system interaction for zonal resistance houses, gas furnaces, electric furnaces, and heat pumps. Taken together, those scenarios produce a set of conservation measures and savings estimates.

Broadly, the simulations are run using all the generalized inputs described in section 4. A number of simulations, using different combinations of input parameters are run, and then the output is averaged together using weights representative of a given parameter’s saturation in the housing population. For example, to simulate the energy use of the average household, we run all five water draw profiles for 1, 2, 3, 4, and 5+ occupants; and then weight the output so the average occupancy is 2.57 people – the current average occupancy for houses with electric tanks.

For comparison, and to calculate energy savings, the baseline energy use of an ERWH is also simulated. Water heater performance standards, as set by the Department of Energy, will change in April 2015. Consequently, in this analysis, we have assumed those standards to set the new baseline; tanks less than or equal to 55 gallons will roughly have an EF of 0.95 and those above 55 gallons will have an EF of 2.0 (essentially a heat pump water heater). The existing stock of electric tank water heaters has an EF near 0.9.

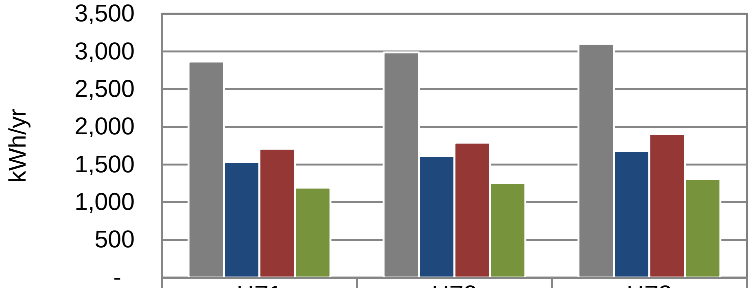
The graphs in Figure 35, Figure 36, and Figure 37 show the simulation output water heater energy use for garage, unheated basement, and interior installations across the three Northwest heating climates. All HPWHs show significant water heating energy reductions over the baseline ERWH. Next, all the water heaters show an increase in energy use as the climates become colder. This is due to both the average inlet water temperature and the average ambient air temperature decreasing. As observed in earlier sections, the GeoSpring’s use of resistance heat, especially at colder temperatures as seen in garages, results in more energy use. Like what was observed in the field, the ATI energy use is the least sensitive to changes in ambient temperature. When installed inside the conditioned space, the HPWHs tend to use more similar amounts of energy. Figure 37 plots the interior cases for both exhaust ducted and unducted installations. Only the ATI equipment was studied with ducting attached. Measurements showed that it was slightly less efficient when ducted due to reduced airflow. Consequently, overall energy use is slightly higher than the unducted case.

Figure 35. Garage Installation Water Heater Energy Use



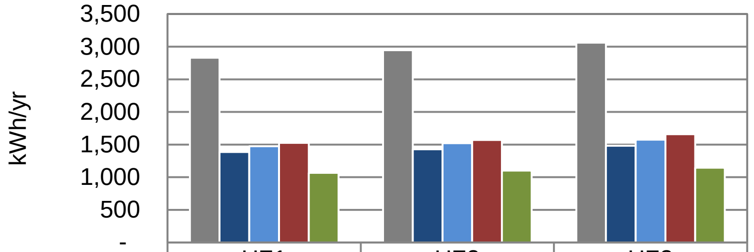
	HZ1	HZ2	HZ3
■ ERWH	2,864	2,989	3,109
■ ATI	1,547	1,655	1,754
■ GeoSpring	1,827	2,156	2,432
■ Voltex	1,238	1,421	1,668

Figure 36. Unheated Basement Installation Water Heater Energy Use



	HZ1	HZ2	HZ3
■ ERWH	2,866	2,986	3,102
■ ATI	1,535	1,609	1,676
■ GeoSpring	1,709	1,791	1,908
■ Voltex	1,193	1,254	1,309

Figure 37. Interior Installation Water Heater Energy Use

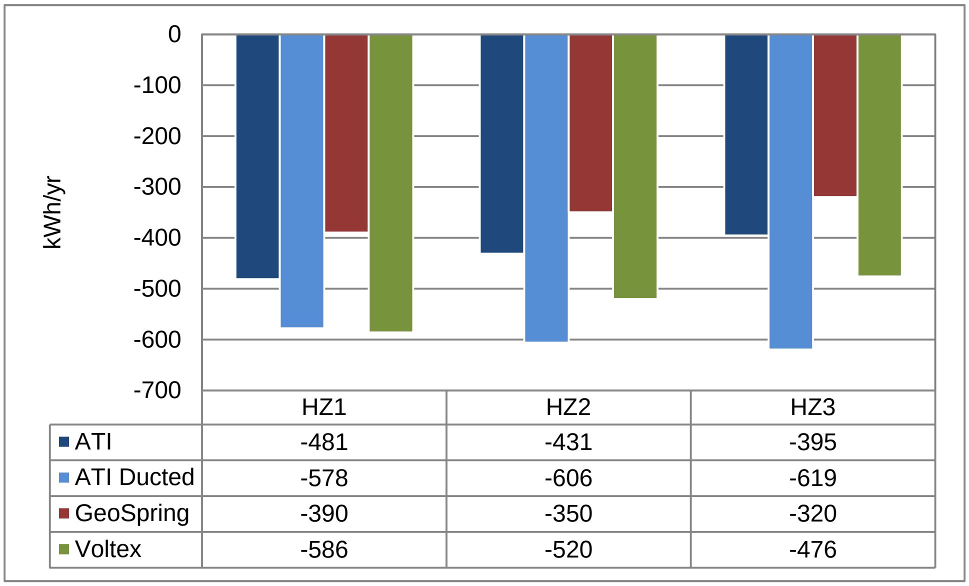


	HZ1	HZ2	HZ3
■ ERWH	2,830	2,946	3,061
■ ATI	1,386	1,429	1,482
■ ATI Ducted	1,476	1,521	1,577
■ GeoSpring	1,526	1,571	1,659
■ Voltex	1,068	1,102	1,145

For installations in garages and unheated basements, Figure 35 and Figure 36 tell the full savings story because there is no heating system interaction. For the interior locations, however, the water heaters either extract heat from the space directly or increase the infiltration load. Figure 38 depicts the heating interaction for houses with electric resistance zonal heat as modeled with SEEM and using $HC_f = 0.65$. The smallest heating penalty is associated with the HPWH which uses its compressor the least – the GeoSpring. The simulation output shows an ever increasing heating penalty for the exhaust ducted systems as the climate zones require more heating. Essentially, the added infiltration load on the house comes at a much colder air temperature than the exhaust from the HPWH.

For other heating system types, like furnaces or heat pumps, the heating system penalty will differ because those systems create and deliver heat to the space with differing efficiencies. For simplicity, the other systems are not shown here. Broadly, the overall impact on houses with electric furnaces is greater than baseboards due to distribution losses in the duct system. Next, it is smaller for houses with space heat provided by a heat pump because that system creates heat with efficiency greater than the resistance system. Finally, there is also an effect on gas furnaces, similar in energy content to electric furnaces, but the added heating load is made up with gas not electricity.

Figure 38. Heating System Interaction – Electric Resistance Zonal Heat



	HZ1	HZ2	HZ3
🟦 ATI	-481	-431	-395
🟦 ATI Ducted	-578	-606	-619
🟥 GeoSpring	-390	-350	-320
🟩 Voltex	-586	-520	-476

Table 23 presents the final electric savings estimates for each water heater make and installation type across climate zones. The savings include the added energy used by the heating system to replace the heat removed by the HPWH. Model output showed the additional cooling provided by the water heater offset only 10-40kWh/yr of cooling energy in most climates. Given the cooling savings is so small, that energy is excluded from the final savings calculations. Additionally, the increased therms used by gas furnaces although calculated in the analysis are not shown in the table. An exhaustive set of simulation output, in a spreadsheet format, is available on the RTF website.[8]

Table 23. Final Electric Savings Relative to ERWH with EF=0.95

Install Configuration	HZ1 Total Savings (kWh/yr)			
	ATI	ATI Ducted	GeoSpring	Voltex
Garage	1317	-	1037	1613
Unheated Basement	1331	-	1157	1661
Interior Gas	1424	1333	1288	1728
Interior EFAF	905	786	867	1093
Interior HP	1165	1107	1077	1409
Interior Zonal	963	776	914	1164
Install Configuration	**HZ2 Total Savings (kWh/yr)**			
	ATI	**ATI Ducted**	**GeoSpring**	**Voltex**
Garage	1334	-	832	1551
Unheated Basement	1378	-	1196	1718
Interior Gas	1498	1403	1360	1810
Interior EFAF	1032	856	982	1243
Interior HP	1186	1104	1103	1424
Interior Zonal	1086	819	1025	1309
Install Configuration	**HZ3 Total Savings (kWh/yr)**			
	ATI	**ATI Ducted**	**GeoSpring**	**Voltex**
Garage	1356	-	677	1422
Unheated Basement	1426	-	1194	1777
Interior Gas	1561	1462	1388	1883
Interior EFAF	1133	921	1042	1363
Interior HP	1235	1113	1128	1476
Interior Zonal	1184	865	1083	1425

---

[8] http://rtf.nwcouncil.org/measures/measure.asp?id=176

5. Conclusions

The HPWH Model Validation Study successfully integrated three datasets of field measurements of in-house performance of water heaters across the Northwest. The studies covered three makes of HPWHs, three climate zones, and four installation configurations. The project was specifically designed to measure the independent variables governing HPWH energy use. The analysis then generalized those variables for use in numerical simulations to predict HPWH energy use in the wider population of houses across the Northwest. Further, the field studies measured the energy of the water heaters and observed their control strategies – both of which are critical to constraining and validating any simulation. The following paragraphs summarize the field findings on draw volume and energy use:

- Measurements of HPWH intake air temperature showed that garages spent over 5% and 21% of their time below 45°F in Heating Zones 1 and 2 respectively. There were no garage installs in Heating Zone 3. Moreover, 13% and 24% of the time (in HZ1 and HZ2) was spent in the next warmest temperature bin of 45°-50°F, were. The 45°F is critical for the Voltex and GeoSpring because below that temperature, they operate in resistance-only heating. Basement temperatures were generally warmer and spent only a small amount of time in the cold temperature range.
- Daily average flow was calculated as 23 gallons per day for a single occupant home, with an additional 11 gallons per day for each additional occupant on average.
- Mean energy use normalized by flow varied between 8 kWh/100 gallons and 13 kWh/100 gallons, depending on make and install location.
- Annual energy consumption of HPWHs typically ranged from 1,000 to 2,000 kWh per year depending on occupancy, make, and install location. The corresponding electric resistance water heater energy use is 3,200 kWh/yr for the average household.
- In terms of efficiency, avoiding resistance heat is, at a basic level, the most important characteristic of a successful heat pump water heater. Small, 50 gallon tanks and garage installs of units with compressor lockouts used the most resistance heat. Larger tanks and tanks with compressors that may operate below 45 °F showed the least resistance heat.
- In addition to examining the hpCOP, which we use in the report as the observed efficiency of the refrigeration cycle, we introduced the concepts of "average annual" aCOP and "electric resistance" erCOP. The aCOP assesses the overall efficiency of useful energy delivered, and includes degradations due to standby losses and resistance heat. The erCOP is a hypothetical comparison for a resistance tank experiencing the same proportions of useful delivered energy and standby losses which serves as the baseline energy use comparison.
- Purely in terms of heat pump operating efficiency hpCOP, the GeoSpring showed the highest efficiency, followed by the Voltex and the ATI. This says the compressor and heat exchanger combination on the GeoSpring was the best performer. However, the GeoSpring also showed the highest fraction of resistance heat, while the ATI showed the lowest. Overall, without accounting for the heating system interaction, the GeoSpring model performed well inside conditioned locations in low-occupant homes, but

experienced dramatically lower aCOP under departures from those conditions. In contrast, the ATI had the least efficient heat pump but was also the most consistent at using heat pump heat nearly exclusively, delivering an aCOP around two almost regardless of install parameters. The Voltex fell in between these two extremes.

- It seems reasonable to attain an aCOP around two (or somewhat better) with this technology, depending on tank size, model, and install location. The observed range in aCOP of 1.6-2.4 represents a possible two-fold to three-fold increase in efficiency over a resistance tank, although the gain in water heating efficiency will be partially offset by an additional burden on the heating system (depending on install location).
- Quantifying the additional space heating burden imposed by an integrated, packaged unit HPWH is of utmost importance to developing estimates of energy savings, but is also challenging. The magnitude of the impact is small relative to the normal daily variation in residential HVAC. This does not mean that the impact is negligible, but rather that it is difficult to accurately measure. Flip-flop tests proved inconclusive. Examining the intake air temperature depression provided an interesting view of how the HPWHs affected their surrounding environment, but also did not lead to useable measurements of space heating interaction. What was useable from the field measurements combined with engineering judgment and experience suggests the following:
    - There is no noticeable heating interaction for garage and unheated basement installations.
    - The interaction factors, HC_f, for both types of interior installations, unducted and exhaust ducted, should be the same. Houses are multi-zone buildings with the HPWH often in a zone which is not fully thermally coupled to the main part of the house. The monitoring of ambient air temperatures demonstrates that the interaction is likely no greater than 0.9. The current body of research suggests 0.65 is a reasonable interaction factor.

In addition to these descriptive findings of energy use, draw sizes, and efficiency, we sought to generalize the relevant parameters and inputs of water heating for use in a simulation. The data collection, acquisition, and meta-analysis of this report were not intended for use as a statistical sample study, but rather to provide the requisite measurements and inputs for a physics-based simulation. Generalized versions of the independent variables include:

- Inlet water temperature was modeled to vary linearly with a 7-week moving average of outdoor air temperature, with the elasticity of the change determined by water source (city surface, city ground/community, city mixed, and well).
- Average measured setpoint across all units was calculated at 128 °F. This value was not found to change with obvious factors like water heater make or number of occupants. For modeling purposes we assert that the mean setpoint should be set to 128 °F.
- Intake air temperature profiles were modeled for each of the four installation scenarios. Garages and unheated basements are calculated based on fits to various outdoor temperature lags. For the interior, recirculating case, we modeled an exponential temperature decline as the HPWH runs. For the interior, exhaust ducted cases, we

determined that the house space temperature from a simulation can be used without modification.

- The recently updated and expanded DOE 24-hour simulated use test draw patterns did not align with the field measured data in terms of daily draw volume, number of draw event clusters per day, and the overall variability in draws. Accordingly, draw schedules for 1, 2, 3, 4, and 5+ occupants were crafted from the observed data. There are both typical day and typical week schedules with the typical week being most appropriate for simulation as it captures enough of the variability inherent in hot water use. Each schedule is tuned to the observed average daily water draw per occupancy category. Within the schedule, the time, size, and duration of draws is informed by the field data themselves.

The project succeeded in reaching its major goal of quantifying all independent performance variables, in detail, to predict energy use for all installation types in the Northwest. The field measurements were generalized and used in a validated HPWH model to calculate energy use and savings. The simulation offered enough flexibility to assess all possible operating conditions and installation configurations. Indeed, the simulation output shows a broad range of savings estimates, corroborated by the field measurements, depending on whether, for example, the HPWH is in a garage in a cold climate or in a conditioned space in a warm one. In the end, the project data and simulation were used to update the unit energy savings estimate at the RTF and guide future research.

References

Baylon, D., P. Storm, K. Geraghty, and B. Davis. 2012. *Residential Building Stock Assessment: Single-Family Home Characteristics and Energy Use*, Prepared for Northwest Energy Efficiency Alliance. Portland, OR. http://neea.org/docs/reports/residential-building-stock-assessment-single-family-characteristics-and-energy-use.pdf

Baylon, David, B. Larson, P. Storm, K. Geraghty, May 2012b. *Ductless Heat Pump Impact & Process Evaluation: Field Metering Report*. Prepared for the Northwest Energy Efficiency Alliance.

Becker, B.R. and K.E. Stogsdill, 1990. "Development of Hot Water Use Data Base." ASHRAE Transactions, Vol. 96, Part 2, pp. 422-427. American Society of Heating, Refrigeration and Air Conditioning Engineers, Atlanta, GA. 1990.

Bedney, Kacie. *Heat Pump Water Heaters: The Journey*. American Council for an Energy Efficiency Economy Hot Water Forum. May 22, 2012. http://www.aceee.org/files/pdf/conferences/hwf/2012/6B-Bedney-Final.pdf

BPA 2012. *Heat Pump Water Heater Demonstration*. October 2012. Bonneville Power Administration. http://www.bpa.gov/energy/n/emerging_technology/pdf/HPWH_demonstration-project_overview.pdf

DOE. 2014. US Department of Energy 10 CFR Parts 429, 430 and 431. *Energy Conservation Program for Consumer Products and Certain Commercial and Industrial Equipment: Test Procedures for Residential and Commercial Water Heaters*. Retrieved from: http://energy.gov/sites/prod/files/2014/06/f17/rwh_tp_final_rule.pdf

Ecotope. 2014. *Residential Building Stock Assessment: Metering Study*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: http://neea.org/docs/default-source/reports/residential-building-stock-assessment--metering-study.pdf?sfvrsn=6

Evergreen Economics. 2013. *Northwest Heat Pump Water Heater Market Test Assessment*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: http://neea.org/docs/default-source/reports/northwest-heat-pump-water-heater-market-test-assessment.pdf?sfvrsn=6

Fairey, P and D Parker. "A Review of Hot Water Draw Profiles Used in Performance Analysis of Residential Hot Water Systems." FSEC-RR-56-04. Florida Solar Energy Center. Cocoa, FL. 2004.

Fluid Market Strategies. 2013. *NEEA Heat Pump Water Heater Field Study Report*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: http://neea.org/docs/default-source/reports/heat-pump-water-heater-field-study-report.pdf?sfvrsn=5

Hanford, Jim, Mike Kennedy, Mary Jane DeLaHunt, and Larry Palmiter, 1985. *Heat Pump Water Heater Field Test*. Ecotope Inc. June 1985. Prepared for the Bonneville Power Administration.

Hendron, R. and J. Burch. "Development of Standardized Domestic Hot Water Event Schedules for Residential Buildings." NREL/CP-550-40874. National Renewable Energy Laboratory. Golden, CO. 2008.

Larson, B., A. Hadley, K. Bedney. 2011a. *Heat Pump Water Heaters: Provisional UES (pre) Proposal Presentation to the Regional Technical Forum.* 30 August 2011. http://rtf.nwcouncil.org/meetings/2011/0830/HPWH_Provisional%20UES%20preProposal_v2_4.pptx

Larson, B., and Logsdon, M. 2011b. *Residential Heat Pump Water Heater Evaluation: Lab Testing & Energy Use Estimates*. Prepared for Bonneville Power Administration Retrieved from: http://www.bpa.gov/energy/n/emerging_technology/pdf/HPWH_Lab_Evaluation_Final_Report_20111109.pdf

Larson, B., and Logsdon, M. February 2012. *Laboratory Assessment of AO Smith Voltex Hybrid Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=888

Larson, B., and Logsdon, M. February 22, 2012. *Laboratory Assessment of AirGenerate ATI66 Hybrid Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=887

Larson, B., and Logsdon, M. September 2012. *Laboratory Assessment of General Electric GeoSpring Hybrid Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=1183

Larson, B. March 2013. *Laboratory Assessment of AirGenerate ATI80 Hybrid Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=1522

Larson, B. June 2013. *Laboratory Assessment of Rheem HB50RH Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=1646

Larson, B. and Logsdon, M. September 2013. *Laboratory Assessment of Sanden GES-15QTA Heat Pump Water Heater*. Prepared for the Northwest Energy Efficiency Alliance. Retrieved from: http://neea.org/docs/default-source/reports/laboratory-assessment-of-sanden-ges-15qta-heat-pump-water-heater.pdf?sfvrsn=8

Lutz, J.D., G. Ghatikar, E. Edleson, S. Meyers. *Hot Water Draw Patterns: Findings from Field Studies*. Environmental Energy Technologies Division. Lawrence Berkeley National Laboratory. University of California. Berkeley, CA. 2006.

Lutz, J. and M. Melody. 2012. *Typical Hot Water Draw Patterns Based on Field Data*. Lawrence Berkeley National Laboratory. https://publications.lbl.gov/islandora/object/ir%3A159117

NEEA 2012a. Northwest Energy Efficiency Alliance. 2012. *Northern Climate Heat Pump Water Heater Specification.* Retrieved from: http://neea.org/northernclimatespec/

NEEA 2012b. Northwest Energy Efficiency Alliance. *Northern Climate Specification Media Backgrounder*. Retrieved from: http://neea.org/docs/northern-climate-heat-pump-water-heater-specification/northern-climate-specification-media-backgrounder.pdf?sfvrsn=4

NPCC 2010. *Sixth Northwest Conservation and Electric Power Plan*. Northwest Power and Conservation Council. February 2010. Portland, OR. http://www.nwcouncil.org/energy/powerplan/6/default.htm

Palmiter, L. June 1982. *Potential for Conservation of Electrical Energy in Northwest Residences.* Prepared for Pacific Northwest Laboratory.

Perlman, M., and B.E. Mills. 1985. *Development of Residential Hot Water Use Patterns*. ASHRAE Transactions, Volume 2a.

Quaid, M., R. Kunkle, and B. Lagerberg. August 1991. *RCDP Cycle 2 Appliance Analysis*. Washington State Energy Office.

Roos, C. and D. Baylon. May 1993. *Non-Space Heating Electrical Consumption in Manufactured Homes, Residential Construction Demonstration Project Cycle II.*

RTF 2011. Regional Technical Forum. 2011. Unit Energy Savings Measures. Residential: DHW – HPWH. http://rtf.nwcouncil.org/measures/measure.asp?id=176

Widder, S., J. Petersen, G. Parker, M. Baechler. July 2014 *Impact of Ducting on Heat Pump Water Heater Space Conditioning Energy Use and Comfort*. Pacific Northwest National Lab. PNNL-23526. http://labhomes.pnnl.gov/documents/HPWH_SpaceConditioning_Report_PNNL_23526_FINAL.pdf

Appendix A: Metering Protocol

The metering protocol collects the information needed to model the home's heat loss rate, duct characteristics and heat pump water heater performance. It also collects sound level measurements. The steps are:

- Perform site audit
    - House characterization
    - Duct characterization
    - Sound levels
    - Site sketch
- Install heat pump water heater
- Install sensors & datalogger
- Set datalogger to log 5 second readings every minute (5 second sampling, 1 minute logging)
- Confirm sensors respond and measure accurately
- Confirm U30 cellular reception
- Confirm heat pump water heater works
- Educate homeowner on heat pump water heater programming and maintenance
- Take pictures/make notes

The house characterization gathers information on insulation levels, windows, heated/unheated area and overall house leakiness.

The duct characterization gathers information on dimensions, location and insulation levels. It also collects information on duct system leakage and airflow.

Sound level measurements are taken in the room where the heat pump water heater is installed and in an adjacent room frequented by the occupants. For each location, two sets of measurements are recorded. The first set measures the ambient noise level. The second set, taken when both the heat pump water heater fan and compressor are running, measures the overall heat pump water heater noise level.

The monitored data record the air and water temperature entering and exiting the heat pump water heater and the amount of energy required to create those temperature differentials. They monitor the energy use of the HVAC system and the house as a whole. Outdoor temperature is also measured.

One-Time Measurements:

Data Point	Measurement Method
Heat Pump Water Heater Fan Airflow	Digital pressure gauge and static pressure tap; take measurement as close to fan outlet as possible
Noise Level Measurements in Room with Heat Pump Water Heater and Adjacent Room	Sound level meter
House Characteristics (for modeling)	Blower door and other tools as needed
Duct Characteristics (for modeling); evaluation includes duct area, location(s), insulation values, and tightness	Duct pressurization fan, digital pressure gauge, and other tools/materials as needed
External Static Pressure (record both supply and return plenum static pressures)	Digital pressure gauge and static pressure tap
System Airflow (combination of duct tightness, external static pressure precursors, and system airflow results in leakage fraction, a primary SEEM input)	Energy Conservatory TrueFlow meter and digital pressure gauge

Continuous measurements

Data point	Sensor type
Cold (Inlet) Water	Wired Veris Thermistor
Hot (Outlet) Water	Wired Veris Thermistor
Heat Pump Water Heater Exhaust	Temperature Sensor
Air Entering Heat Pump Water Heater	Temperature Sensor
Outdoor Temperature	Temperature Sensor
Service Drop	CTs, WattNode and Pulse Counter
HVAC	CTs, WattNode and Pulse Counter
Heat Pump Water Heater Power	CTs, WattNode and Pulse Counter
Water Flow	Water Flow Meter and Pulse Counter

The datalogger is an Onset Computer Corporation U30 with cellular data connectivity.

Site Form:

Name:	**Site ID:**
Address:	**Date:**
Phone:	**Technician(s):**
Utility:	

House type:	Rambler 2 story Split level Attached gar. Duplex/townhome Has conditioned basement	**Year house built**	
		Indicate major remodel details/dates (especially if weatherization occurred):	
Primary Heating System Type Gas Forced air furnace Electric FAF Heat Pump Dual Fuel HP DHP Zonal Electric Other:	**Location of air handler** Garage Inside Crawl Attic Other **Does site have central AC?** Yes No **Does site use non-utility fuel?** Wood Oil Propane Other ___________ Quantity/yr:	**HPWH Make:** [ ] GE GeoSpring 50gal [ ] AO Smith Voltex 60gal [ ] AO Smith Voltex 80gal [ ] AirGen 66gal – unducted [ ] AirGen 66gal – ducted **HPWH Serial:** **Where installed:** **Location of exhaust exit:** **Water temperature:**	
Large unusual loads (well pump, spa, shop, etc):			

Occupants:

Occupant age	Total number of occupants	Occupant age	Total number of occupants
Under 1		19-45	
1-6		46-64	
6-10		65+	
11-18			

House Audit and U-value Tables

We need to know enough about the house to estimate its heating load. ***You therefore need to calculate a house UA.*** *The purpose of this is to compare the load with the heat pump size. Areas can be reported to the nearest 50 ft^2. Accuracy is more important in poorly insulated houses. For windows, the big break is between single and double glazed units; within double-glazed units with metal frames, older units have smaller air spaces and non-thermally improved frames.* ***Calculate house volume; you do not need to calculate infiltration UA.***

Above Grade Walls

Uninsulated	0.25
R-11	0.09
R-19	0.065

Doors

Hollow wood*	0.50
Panel wood*	0.40
Solid wood*	0.35
Insulated metal	0.20

*subtract 0.15 from U-value if storm door installed. If more than half glass, use appropriate glass U-value.

Below Grade Walls (fully below grade; assumes uninsulated slab)

Uninsulated	0.2
R-11	0.06
R-19	0.04

Floor Over Crawlspace

Uninsulated	0.12
R-11	0.055
R-19	0.04
R-30	0.03

Slab Floors (use lineal feet, not ft^2)

Uninsulated on grade	0.75
Uninsulated below grade	0.50
Insulated on grade	0.55

Attics/vaults

Uninsulated	0.3
R-11	0.06
R-19	0.05
R-30	0.04
R-38	0.03

Windows

Single glazing	1.1
Double glazing metal	0.75
Double glaze metal improved	0.65
Double with wood/vinyl frame	0.55
Dbl wood/vinyl low e	0.40
Modern high-performance	0.30

House UA Calculation Page

(show sketch or use separate sheet of graph paper)

Record house UA (no infiltration) here: ___________________Btu/ft^2 ºF

Record conditioned floor area here: _______________ ft^2

Record conditioned house volume here: _________ft^3

Duct Audit

We need enough information to estimate the system efficiency of the ducts. This means getting the length and diameter and insulation level of the ducts in unconditioned spaces such as garage, attic and crawlspace. If the ducts run between-floors, also note this. ***Ducts fully inside the conditioned space do not need to be measured.*** *Measure diameters to nearest inch and lengths (overall) to nearest 3'. Estimate as needed to save time by pacing off runs inside the house, using stud spacing as an estimating device, etc. If insulation is damaged or missing, note as needed. The duct audit should take no more than 30 minutes.* ***Describe both supply and return sides of system.***

Supply ducts (list all unique dimensions/insulation levels)

Duct type (metal/flex)	Duct Zone Location (garage, attic, crawl, other)	Dimension (LxW or inside diameter if round)	Length (feet)	Area (ft^2) (convert dimension to ft first)	Insulation (best guess on R-value)*	UA to Duct Zone

*R-value/inch is about 3 for fiberglass; derate if damaged or missing

Return ducts (list all unique dimensions/insulation levels)

Duct type (metal/flex)	Duct Zone Location (garage, attic, crawl, other)	Dimension (LxW or inside diameter if round)	Length (feet)	Area (ft^2) (convert dimension to ft first)	Insulation (best guess on R-value)*	UA to Duct Zone

*R-value/inch is about 3 for fiberglass; derate if damaged or missing

If any ducts in crawl, check if crawl is vented (more than 4 open vents) ______
If any ducts in attic, check if attic vented (soffit and ridge or gable vents) ______

Notes on duct system:

2-Point Blower Door Test

Depressurize to near 50 and 25 Pa with respect to outside. **Note the house pressure WRT outside doesn't have to be exactly 50 or 25 Pa; the actual values will be corrected to 50 Pa during analysis.**

Make and model of blower door used ____

Blower Door (BD) Depressurization Test Procedure:

1. *Close all windows and doors to the outside. Open all interior doors and supply registers.*
2. *Close all dampers and doors on wood stoves and fireplaces**.** Seal fireplace or woodstove as necessary to prevent ash disaster.*
3. ***Make sure furnace and water heater cannot come on during test. Put water heater and/or gas fireplace on "pilot" setting. Make sure all exhaust fans and clothes dryer are off. Make sure any other combustion appliances will not be backdrafted by the blower door**.*
4. ***Make sure doors to interior furnace cabinets are closed**. Also make sure crawlspace hatch is on, even if it is an outside access. Check attic hatch position. Put garage door in normal position.*
5. *Set fan to de*pressurize house. Run pressure tap out through door shroud.*
6. *Depressurize house to –50 Pa or thereabouts. Record house pressure, BD flow pressure, and BD ring (below). If you cannot reach –50 Pa, get as close as possible and record information.*
7. *Now take the house down to –25 Pa WRT outside and record information.*

Blower Door Tests	House P near 50 Pa (P_{50})	BD fan pressure	BD Ring	BD flow near 50 Pa (Q_{50})	House P near 25 Pa (P_{25})	BD fan pressure	Ring	BD flow near 25 Pa (Q_{25})
Test 1								
Test 2								

8. *To check test, calculate the flow exponent, n. Use the following formula, n = $\ln(Q_{50}/Q_{25})/\ln(P_{50}/P_{25})$. Note Q_{50} and Q_{25} are the flows through the blower door at the testing pressures (which are denoted P_{50} and P_{25}. Depending on the test, you may not get the house to exactly –50 or –25 Pa WRT outside. Use the exact ΔP you measure when checking the flow exponent. For example, if the house gets to –48 Pa for the high ΔP, use this as the P_{50} in the equation. If the flow exponent is not between 0.50 and 0.75, repeat the test.*

Note testing conditions (if windy, inaccessible room(s), garage door open or closed, etc):

Exterior Duct Leakage Test

Performing exterior duct leakage test:

1. *Exterior house doors and garage doors should be closed for exterior duct leakage test.*
2. *Pressurize the house to about 50 Pascals WRT outside.*
3. *Pressurize tested part of duct system to about 50 Pascals with smallest flow ring possible.*
4. *Measure pressure of ducts WRT house. Make sure blower door flow does not impinge on pressure tap measuring house pressure.*
5. *Adjust duct tester speed controller so that duct pressure WRT house is zero or very close.*
6. *Re-check pressure of ducts WRT outside.*
7. *Measure duct tester fan pressure. Look up flow in table, use gauge* ***(make sure gauge is paired with the right duct tester)*** *or use flow equation. Record duct pressure WRT out, DB fan pressure, DB fan ring.*
8. ***If you cannot reach 50 Pa or 25 Pa, test to the highest pressure you can reach and enter this in the 50 Pa column. Use a test pressure of half this pressure for the low pressure test.***
9. *Repeat steps 2-7 with house and ducts at about 25 Pa WRT outside.*
10. *Check flow exponent (as above).*
11. *Note any unusual testing conditions (wind, etc.):*

Duct Leakage to Outside Data (note duct pressure WRT outside may not be exactly 50 or 25 Pa)

	Both sides		Supply or Return (circle one)	
	50 Pa	25 Pa	50 Pa	25 Pa
Duct P	____	____	____	____
Ring	____	____	____	____
Fan P	____	____	____	____
Flow	____	____	____	____

12. *To check test, calculate the flow exponent, n. Use the following formula,* $n = \ln(Q_{50}/Q_{25})/\ln(P_{50}/P_{25})$. *Note* Q_{50} *and* Q_{25} *are the flows through the blower door at the testing pressures (which are denoted* P_{50} *and* P_{25}. *Depending on the test, you may not get the house to exactly –50 or –25 Pa WRT outside. Use the exact* ΔP *you measure when checking the flow exponent. For example, if the house gets to –48 Pa for the high* ΔP*, use this as the* P_{50} *in the equation. If the flow exponent is not between 0.50 and 0.75, repeat the test.*

TrueFlow Test

Set-up: *Turn on air handler (by using fan-only switch or by turning on heat/AC). Drill access hole as needed and point hooked end of static tap into airflow. Do not drill into the duct at any point where you are concerned with hitting something.* **Repeat test if needed to get flows at both low and high stage; record first stage readings to left of "/" in blanks below and second stage readings to right of "/".**

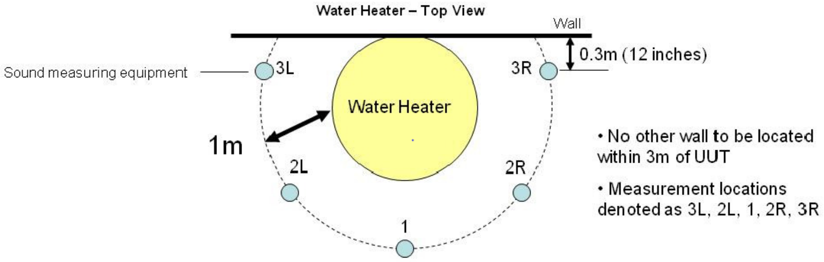
Measure pressure in return plenum and record:_____/_____ Measure pressure in supply plenum. Record pressure below as Normal System Operating Pressure (NSOP). Place appropriate plate and spacers into filter slot. Turn on air handler and record supply static pressure with TrueFlow in place (TFSOP) and pressure drop across plate.

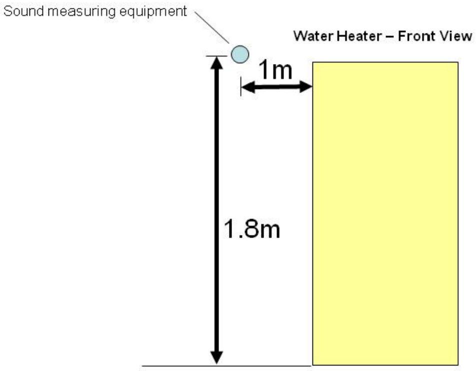
Plate used (14 or 20) _______/________
Normal System Operating Pressure (NSOP) ____/_____Pa
True Flow System Operating Pressure (TFSOP) ____/_____Pa
Correction Factor* √(NSOP/TFSOP) _____/______

Plate pressure drop ____/_____Pa
Raw Flow (CFM)______/______
Corrected Flow (CFM)______/______

Sound Measurement Test Method

Measure the sound level in the room where the HPWH is installed and in an adjacent room frequented by the occupants. For each location, record two sets of measurements. The first set measures the ambient noise level. The second set, taken when both the HPWH fan and compressor are running, measures the overall HPWH noise level. For the room with the HPWH, take 5 measurements of the sound according to the specs in the figure. If 5 measurements can't be taken, take as many as possible:





Sound Levels in Room where HPWH is Installed

Location	Ambient dBA	HPWH dBA
3L		
2L		
1		
2R		
3R		

For the room adjacent to the HPWH, measure the sound level in two locations: (1) measure in the center of the room at ear level and (2) measure at the wall adjacent to the room with the HPWH.

Sound Levels for Room Adjacent to HPWH

Location	Ambient dBA	HPWH dBA
Middle of Room		
Adjacent Wall		

Water heater exhaust duct flow: use static pressure tap and measure with digital pressure gauge at exhaust port, before any elbows in duct:	

Datalogger Information

Device	S/N	Notes
U30 [ ] AT&T or [ ] Verizon?	Device keycode: RSSI:	
WattNode Model:		
Temp. sensor 1 (Cold (inlet) water)		
Temp. sensor 2 (Hot (outlet) water)		
Temp. sensor 3 (HPWH exhaust)		
Temp. sensor 4 (air entering HPWH)		
Temp. sensor 5 (OAT)		
Pulse 1 (SERV)		CT size: Parallel install? Y N
Pulse 2 (HVAC)		CT size: Parallel install? Y N
Pulse 3 (HPWH power)		CT size: Parallel install? Y N
Pulse 4 (Water flow)		

Notes:

Exit Checklist

○ Water heater tested and working
○ Datalogger channels tested and working
○ HPWH set to auto mode
○ Water temperature set no higher than 120
○ Homeowner informed what temperature was on arrival
○ Homeowner educated about HPWH use
○ Form and photos complete

Appendix B: Details of Data Annualization

We sought to generate annualized summaries of the following variables: flow, intake air temperature, inlet water temperature, heat pump water heater total input energy, and fraction of heat delivered through resistance elements. The regression modeling was most straightforward for the temperatures, as they are continuous, real-valued readings. For flow and water heater energy use, an ordinary linear model was found inadequate, as these values are constrained to be non-negative, and hence the regression residuals are not constant variance and the “predicted" values can actually be negative. Asserting negative energy use leads to a bad annualization! Similarly, the fraction of resistance heat is a proportion between zero and one, leading to similar problems.

Broadly, when confronted with regression responses that do not easily fit into the framework of the linear model (constant variance, real-valued error terms), there are two options: transform the response, or consider a generalized linear model (GLM) in which the probability distribution on the outcome is not Gaussian (the normal bell curve). Transformations were more common before generalized linear models became available computationally, but the GLMs are often preferable because they typically offer “nicer" interpretations of coefficients. For example, the variance-stabilizing transform (the canonical transformation to apply before using a linear model) for counts of a rare event is to take the square root. However, it is awkward to interpret coefficients from a regression when talking about the effect on the square root of the response.

The initial approach we took was to model seasonal trends as follows for three main types of outcome:

1. Temperatures – ordinary least squares regression,
2. Water flow and water heater energy use – Gamma GLMs (the GLM with Gamma distribution on the outcome has the property that the standard deviation is proportional to the mean, which seemed to be a pretty good variance model for flow in particular), and
3. Fraction of input energy provided by resistance elements – binomial GLM (also known as logistic regression).

These regression models worked well on the well-behaved sites, but, as stated above, running individual regressions on each end use at each site often led to unstable curve fits that were not entirely believable. In short, where it worked it worked very well in terms of specifying both a mean model and variance model that closely tracked the data. But it didn’t work for all sites. This led to consideration of linear mixed effects models.

A “mixed effects model" is a regression model that combines the usual sorts of terms – known in this context as “fixed effects" – along with so-called “random effects". The distinction is that fixed effects are assumed to be fixed yet unknown numbers. Random effects are assumed to be drawn from some specified probability distribution. This in effect places a constraint on the estimated values associated with those coefficients (a stickler statistician will tell you that these quantities cannot be “estimated,” only “predicted,” as estimation refers to guesses for the value of fixed yet unknown quantities, whereas we have specified the random effects as random variables, not fixed quantities). This type of model can also be interpreted through the lens of penalized regression; there is an equivalency between specifying Gaussian random effects and specifying fixed effects with a ridge penalty[9].

The generic form for the annualization model, somewhat arbitrarily written for intake air temperature Tintake, is given below. The equation denotes that the intake air temperature for site *i* at time *j* is represented as a linear function of an overall intercept β_0, a site specific intercept b_{0i}, a linear combination of sines and cosines of time in the year – dictated by overall slopes β_1 and β_2, and site specific slopes b_{1i} and b_{2i} – and some error term ε. The b_i terms and the error term ε are assumed independent draws from normal distributions.

$$\text{Tintake}_{i,j} = \beta_0 + b_{0i} + \beta_1 \sin(t_{i,j}) + \beta_2 \cos(t_{i,j}) + b_{1i} \sin(t_{i,j}) + b_{2i} \cos(t_{i,j}) + \varepsilon$$

$$b_0 \sim N(0, \sigma_0^2);\ b_1 \sim N(0, \sigma_1^2);\quad b_2 \sim N(0, \sigma_2^2);\ \varepsilon \sim N(0, \sigma_\varepsilon^2)$$

Using random effects in this way – essentially as a smoothing tool – introduces bias, but with the benefit of much lower variance, such that the “predictions" should have lower mean squared error (if we got to hypothetically observe the unobserved data points).

It would have been preferable to simply add site-level random effects to the GLM regressions described initially, but the mixed effects model computational machinery is finicky and fickle, especially with slightly unusual distributions such as gamma. It proved more prudent to apply a linear mixed effects model to transformed outcomes (for example, the logarithm of flow), than fight with unexpressive R errors while attempting to fit generalized linear mixed effect models.

The annualizations were ultimately performed through linear mixed effects models on the relevant transformed outcome scale. These scales were mainly chosen to map the actual range of possible values for a measurement to the entire real line. For example, water heater flow must be non-negative, and one very obvious way to map positive numbers to the entire real line is taking a logarithm. Similarly, the resistance fraction lies between zero and one, and a method of mapping the real interval [0,1] to the real line is the arcsine of the square root (this is also the variance-stabilizing transformation for a proportion). One potential problem, however, with transforming away from the measurement scale is that a linear mean model may be no longer plausible. Asserting linearity in the logarithm of some variable implies exponential in the actual value of that variable, which may or may not be a good assertion. However, all we’re really estimating for each site is an amplitude and a phase of the seasonal trend, and so unusual functional forms for the mean model result in very small practical differences in predicted values.

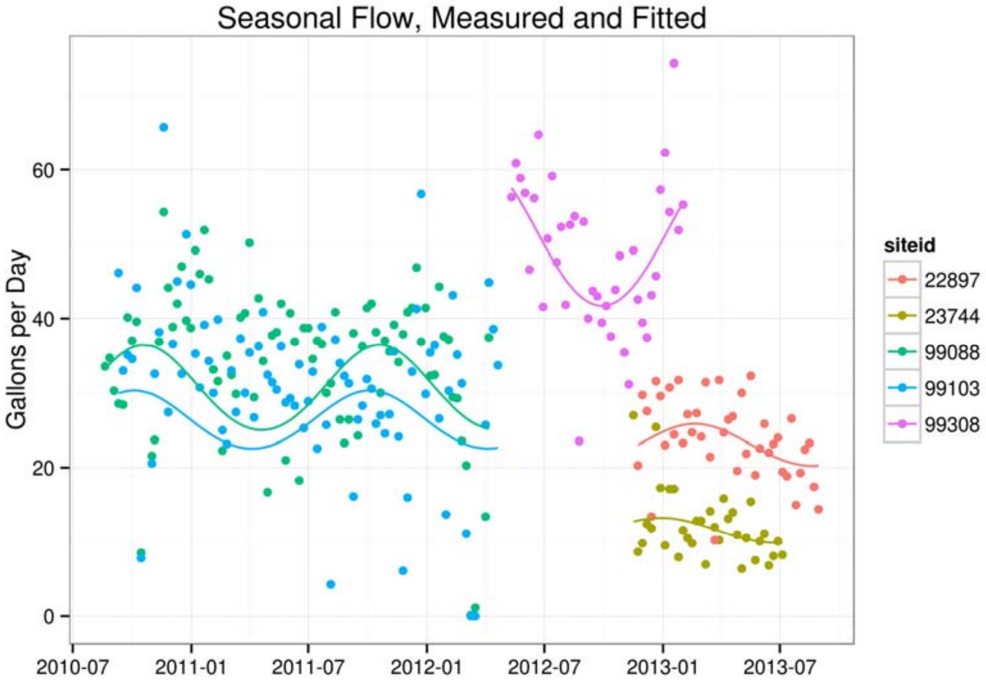
The final transformations used were as follows:

1) Temperature – no transformation.
2) Water flow and water heater energy use – natural logarithm scale.
3) Fraction of input energy provided by resistance elements – arcsin square root scale[10].

---

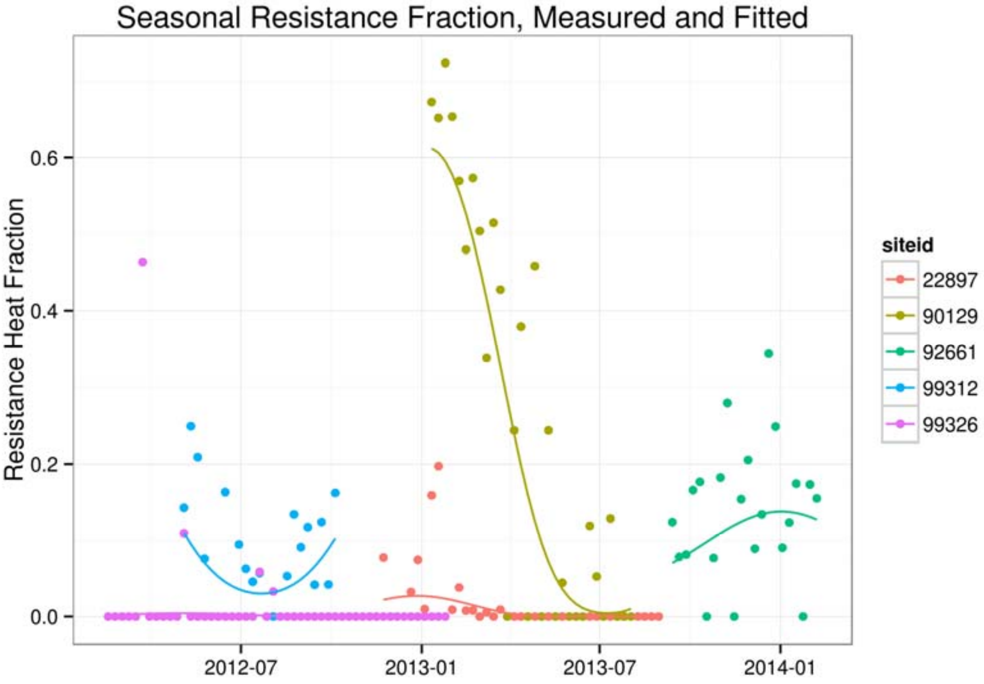
[9] Basically, a ridge penalty optimizes the usual sum of squared errors criterion, subject to a constraint on the sum of squared coefficients.

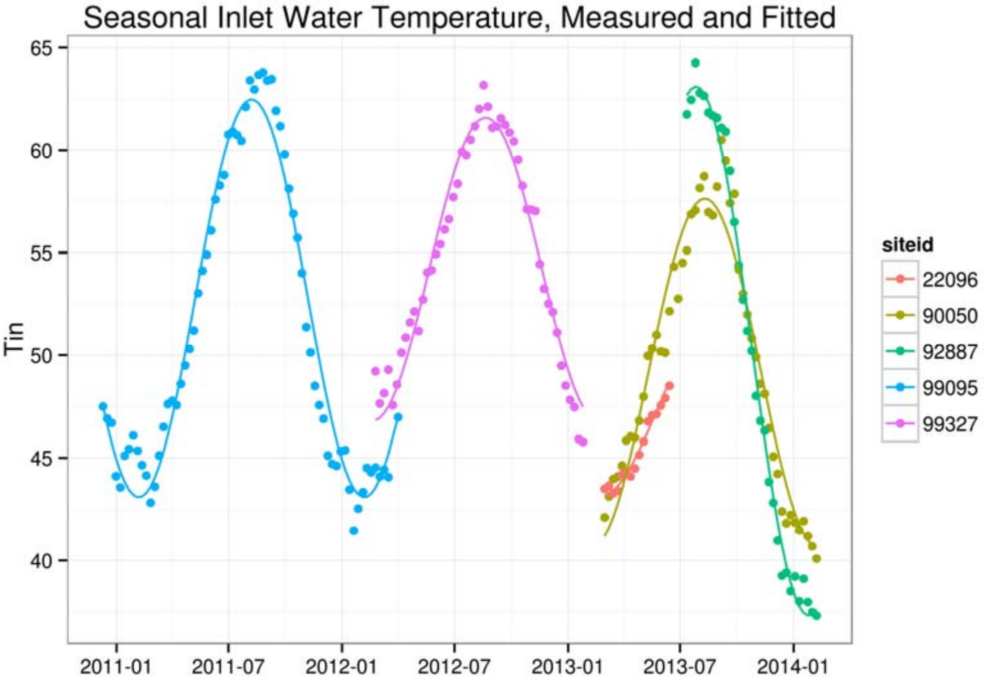
On the appropriate scale, a linear mixed effects model was fit with site level random intercepts and site level random slopes with reference to the trig terms. In addition, relevant covariates were included for adjustment. For example, the flow regression further adjusted for occupancy, the resistance fraction regression adjusted for water heater make, etc… See below examples plots, showing measured and fitted values for a handful of sites each.



---

[10] Luckily we're just trying to fill in a few months of missing data to create an annualized estimate, and not interpret coefficients, because it is not intuitive to talk about changes to the arcsine of the square root of the proportion of resistance heat.


Seasonal Resistance Fraction, Measured and Fitted
Resistance Heat Fraction
0.0
0.2
0.4
0.6
2012-07
2013-01
2013-07
2014-01
siteid
22897
90129
92661
99312
99326


Seasonal Inlet Water Temperature, Measured and Fitted
Tin
40
45
50
55
60
65
2011-01
2011-07
2012-01
2012-07
2013-01
2013-07
2014-01
siteid
22096
90050
92887
99095
99327

Appendix C: Diagnostic Regression Model

The diagnostic regression model – that ferreted out performance and data logging anomalies – was similar in flavor to the annualization model. Both use so-called "random effects" as a shrinkage tool, in order to avoid overfitting, and smooth unit-specific results to the overall mean. In addition, both considered data aggregated weekly. Further, we assessed the entire field data in the study with tool described within this appendix.

Several models were considered, with goodness of fit assessed through the Akaike Information Criterion (AIC). The AIC "penalizes" the likelihood according to the number of parameters in the model. A more complicated statistical model will always fit the data better, but the key question is whether the boost in model fit was substantial enough to justify including another term (fewer terms are always preferable). Philosophically, the AIC is motivated by attempting to minimize the Kullback-Leibler divergence (a sort of distance between probability distributions) between the true data-generating process and the statistical model. Obviously the true data-generating process is unknown, but one can approximate the relative information loss for two models by comparing their AIC scores, which are calculated as AIC = 2k – 2 ln(L): twice the number of parameters minus twice the log likelihood. There are other methods for this type of non-nested model selection – such as the Bayesian Information Criterion (BIC), which seeks the model with highest posterior probability of being the true model[11] – but for this exercise AIC is as good as any. You'd have to be a bit naïve to really believe any of them, anyway – I mainly use it merely to have some criterion with which to compare models.

The ultimate functional form treated transformed weekly compressor on-time as a linear function of flow, transformed resistance element on-time, inlet water temperature, intake air temperature, outlet water temperature, an indicator for whether the unit was ducted, and a unit-specific random effect. This model form is displayed in the equation below, denoting the expected transformed compressor on-time for unit i at time j. As before, specifying the random effect allowed unit-specific deviations, which could themselves be evaluated for anomalies. The on-time transformations were again arcsin square roots of the fraction of respective time spent running. The model was fit using the "lmer" function in the "lme4" package with the statistical software R, separately for each water heater make.

$$\boldsymbol{E}[sin^{-1}(\sqrt{comp\_on_{i,j}})] = \beta_0 + b_{0i} + \beta_1 Flow_{i,j} + \beta_2 sin^{-1}(\sqrt{res\_on_{i,j}}) + \beta_3 Tinlet_{i,j} + \beta_4 Tintake_{i,j} + \beta_5 Toutlet_{i,j} + \beta_6 I(unit\ i\ ducted)$$

$$b_0 \sim N(0, \sigma_0^2)$$

- comp_on = fraction of the week spent with compressor running
- Flow = weekly total hot water draw
- res_on = fraction of the week spent with resistance element running
- Tinlet = average measured inlet water temperature during flow events
- Tintake = average measured intake air temperature during compressor operation

---

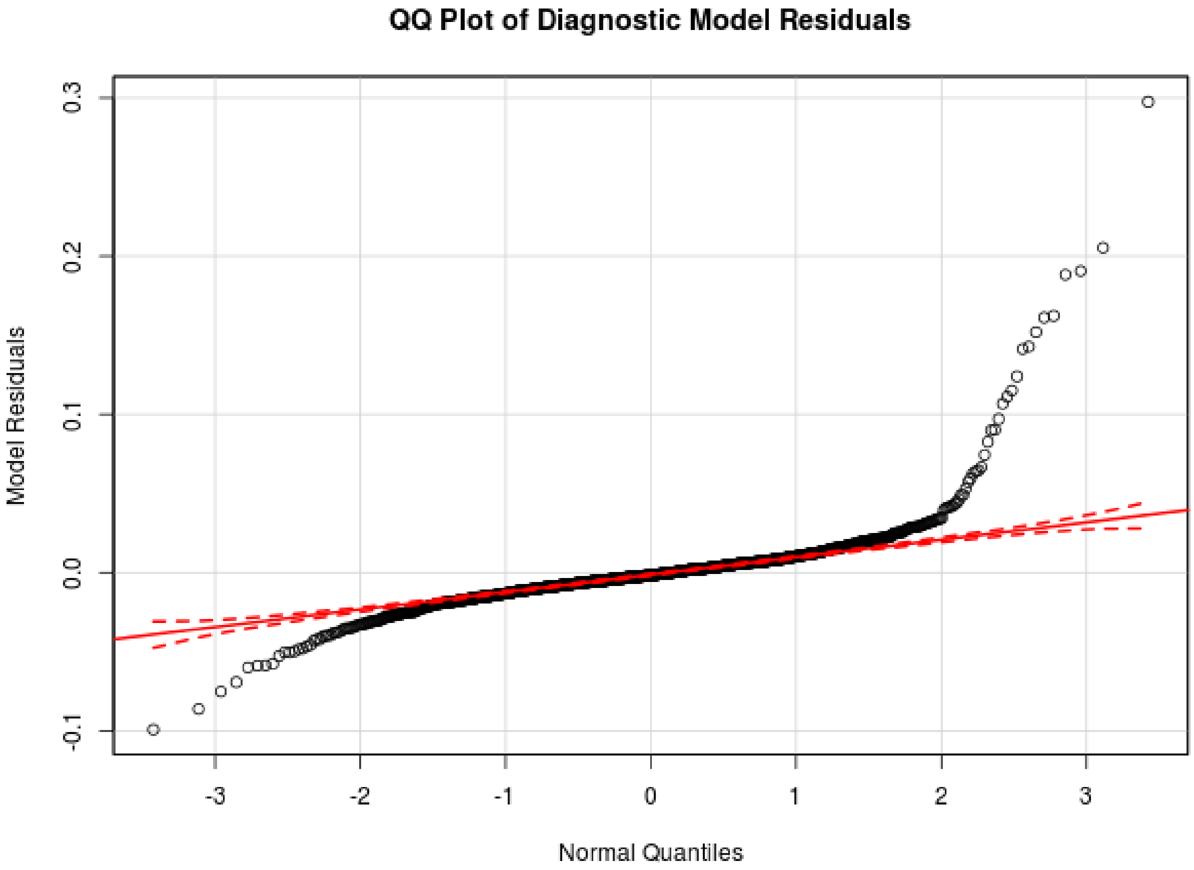
[11] Treating the statistical model as a nuisance parameter is a horrifying thought for frequentist statisticians.

- Toutlet = average measured outlet water temperature during flow events
- I(unit i ducted) = a binary indicator for whether the unit in question was ducted

Figure 39 shows a QQ plot of the model residuals, and is mildly frightening. QQ plots help visualize how closely some set of numbers follows a given distribution: here we hope that the regression residuals look Gaussian, which is only believable if the points fall roughly along the line. That is obviously not the case. I would hesitate to trust p-values or standard errors from this model, but it should be plenty sufficient as just a tool to call attention to egregious performance anomalies.

Note that the diagnostic model is necessarily *hypothesis generating*. We cannot observe the output and immediately reach dramatic conclusions as to improperly performing units. We may only note oddities as requiring further investigation. Units flagged as unusual by the diagnostic model were closely scrutinized, and it was under that extra scrutiny that the ATI performance problems became apparent.

Figure 39. QQ Plot for Diagnostic Regression Model



Probably also of note, with respect to the diagnostic model, are the complicated ideas that didn't work. The simple linear form for the model seems unlikely to closely describe the workings of a heat pump. However, none of the more complicated approaches yielded a better AIC score: adding complexity to the model did not increase fit by enough to make it worthwhile.

Generalized Additive Models:

Generalized Additive Models (GAMs) can be useful for regression modeling non-linear data, where the functional form is not only unknown, but may not even be important[12]. The idea is to model the response not as a sum of linear terms in the covariates, but as a sum of arbitrary functional forms. The functional fits may be derived in many ways, although splines are popular and splines are what I used. Basically a spline fits a smooth curve to the data according to some criteria to minimize overfitting, such as a penalty on curvature (the integrated second derivative) or a penalty on the magnitude of the coefficients (a ridge penalty). These models allow fairly arbitrary functional forms of the covariates.

None of the various GAMs investigated boosted the AIC beyond the simple, completely linear model. We allowed for non-linearities in each of the terms separately, as well as some interaction-type terms with two-dimensional thin plate splines, but none of these more complicated fits explained the data with greater clarity and parsimony than the linear model.

The Flow Severity Score:

Another idea we had was scoring how tightly clustered the hot water draws were. Since heat pump water heaters add heat to cold water much more quickly and efficiently than to hot water, it seems logical that units with a small number of rapid draws should perform more efficiently than units with many small, diffuse draws. An obvious way of calculating the “spread” of the hot water draws is through a statistical entropy: build an empirical distribution of water draw time of occurrence, then calculate the entropy defined as $-\sum_i p_i log(p_i)$, where p_i denotes the flow in some interval i, scaled by the total flow (the probability of flow occurring in that interval). Notice that the minimum of this function – zero – occurs when all probability piles up at a single point. The maximum occurs when all time periods *i* have the same probability.

As an example, consider a hypothetical home with hot water draw of 48 gallons per day. The maximum entropy draw schedule would draw two gallons per hour, every hour of the day. The minimum entropy draw schedule would draw all 48 gallons in a single pass. Think of it as a measure of how diffuse the draw schedule is.

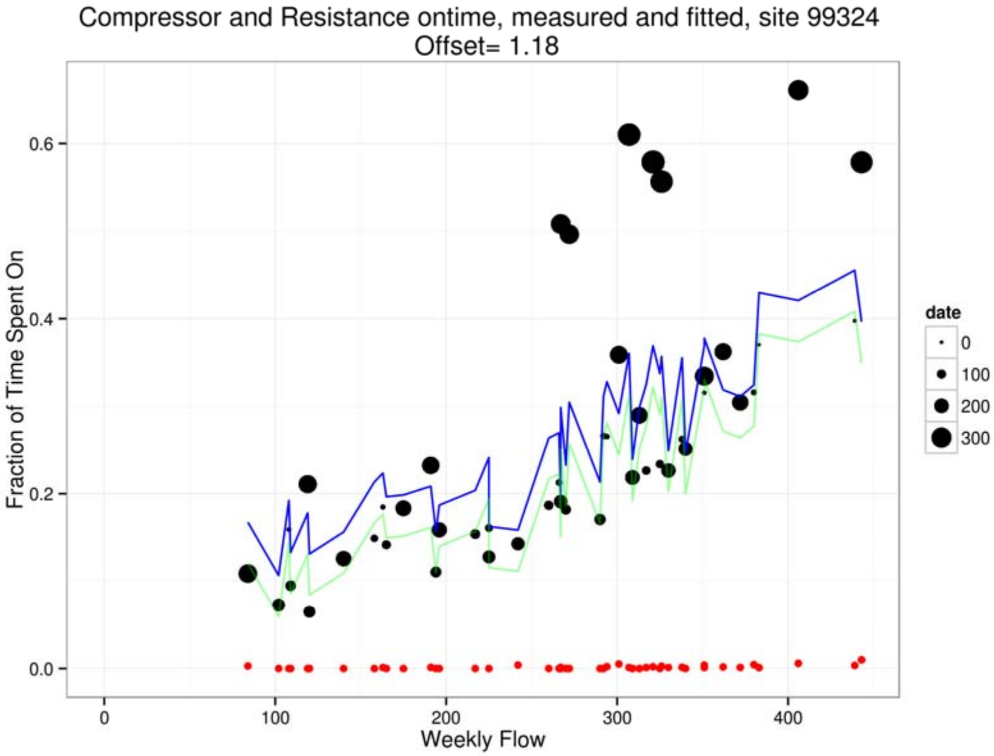
Interestingly, the flow severity score – essentially the entropy of the empirical draw schedule – correlated with nothing after adjusting for obvious variables like total flow and operating conditions. The flow severity score seemed like a good idea, but we found it ultimately uninformative.

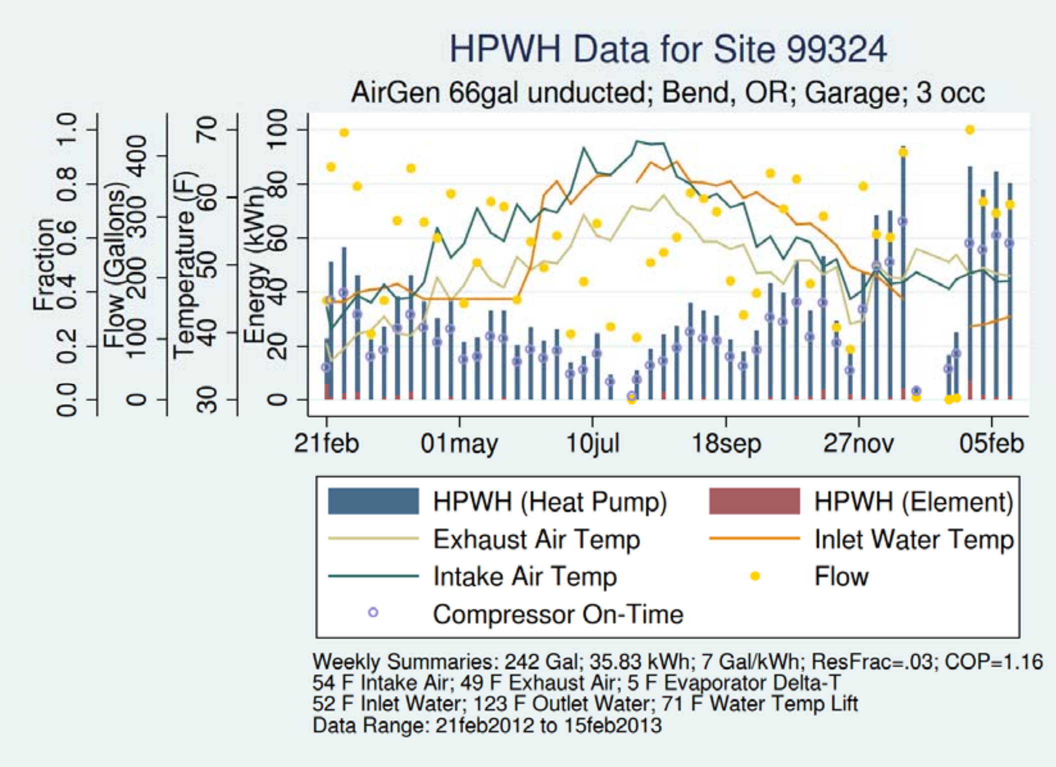
Finding the malfunctions:

---

[12] A good reference for non-parametric regression like splines and GAMs is *The Elements of Statistical Learning,* by Trevor Hastie, Robert Tibshirani, and Jerome Friedman out of Stanford.

As a final example on how this model was used to find potential refrigerant problems in the datasets, consider the following two figures: in the first, we see runtimes much higher than expected in the final approximately two months of monitoring. In the second we see – a few months from the end of monitoring – the measured exhaust air temperature abruptly rise to equal or even exceed the measured intake air temperature. This was thought highly indicative of refrigerant loss.





HPWH Data for Site 99324

AirGen 66gal unducted; Bend, OR; Garage; 3 occ

Weekly Summaries: 242 Gal; 35.83 kWh; 7 Gal/kWh; ResFrac=.03; COP=1.16
54 F Intake Air; 49 F Exhaust Air; 5 F Evaporator Delta-T
52 F Inlet Water; 123 F Outlet Water; 71 F Water Temp Lift
Data Range: 21feb2012 to 15feb2013

Appendix D: Space Conditioning Interaction

A handful of sites were selected for tests known often as “co-heat”, or more informally “flip-flop”, in which the space heating impact of an interior HPWH is assessed by forcibly switching the water heater between heat pump and resistance modes. The idea was to develop a heating signature for the home in each operating mode, and use the difference between those heating signatures to assess the impact of the heat pump water heater. We hoped for exploratory and illuminating findings on the magnitude of the penalty, suspecting that each Joule removed from the interior space is not made up with a full Joule output from the heating system. To facilitate the exercise, the water heaters were manually switched twice during the heating season by the occupants (after receiving phone calls from project staff requesting the switch), once into resistance mode, and once back into heat pump mode.

To a first approximation, space heating energy is linear with outdoor temperature. Colder temperatures require more heat to keep the house at the same indoor temperature. Figure 40 shows this relationship for a house with a heat pump in Yakima, WA. Each point on the graph is the total daily HVAC system energy use for a given day in December. The horizontal axis plots the daily average outside temperature. The red line is the linear fit of heating energy to outdoor temperature. Figure 41 presents the same house but for the month of April. In the warmer days of April, the heating system sometimes doesn’t run, as shown by blue points at the bottom of the graph.

Comparing Figure 40 and Figure 41 show a different slope to the fitted heating line depending on the time of year. Therefore, to develop an accurate model of heating system energy versus outdoor temperature, we sought to observe the house over a range of temperature conditions. Both cold and warm temperatures are necessary, (deep winter and mild spring or autumn conditions).

Figure 40. Example Wintertime HVAC Use

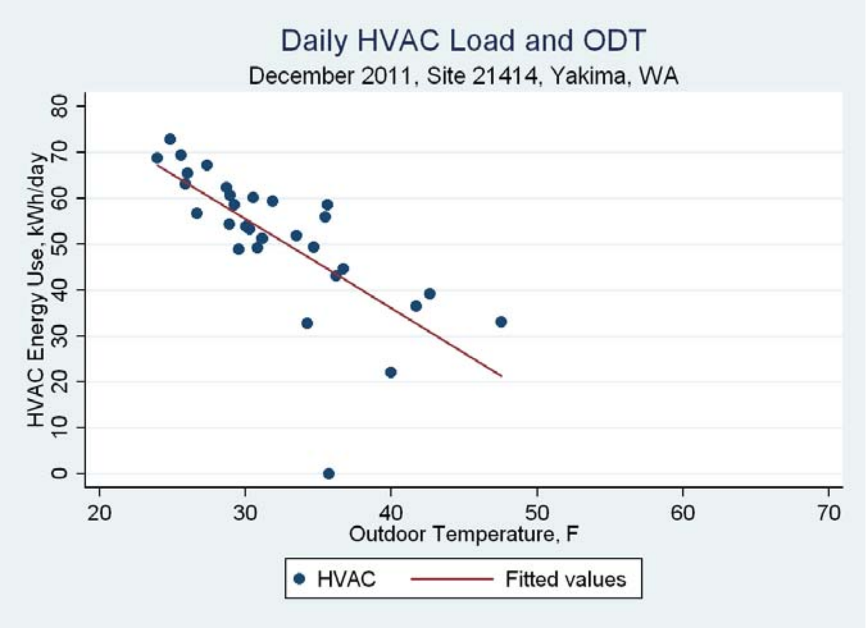
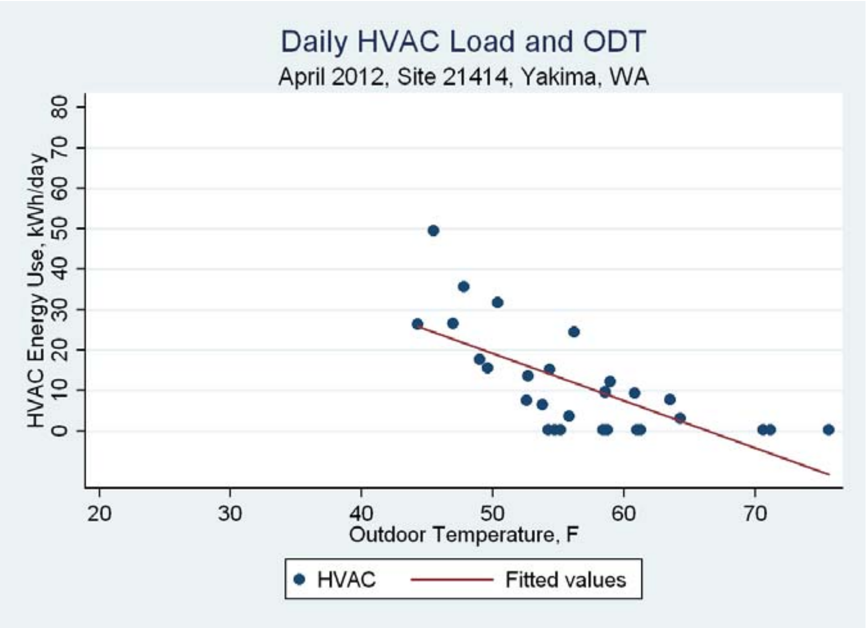
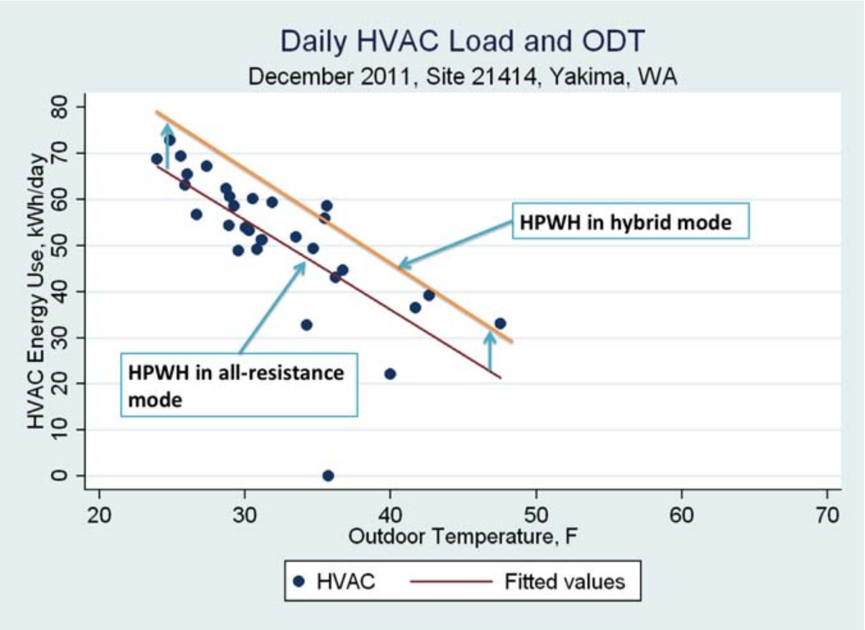


Figure 41. Example Springtime HVAC Use



A HPWH installed completely inside the conditioned space of the house has the potential for heating interaction. When the compressor runs, it cools the house, whereby imposing an added heating load on the house (a negative internal gain). Conceptually, the added heating load will shift the line in Figure 40 upward. Figure 42 displays a hypothetical version of the result. The amount of the vertical shift in the graph corresponds to the added heating load imposed by the HPWH. Additionally, a slight shift should theoretically occur in the heating slope as well, as the water heating load increases during the coldest months due to colder inlet water. This alteration of space heating ultimately counts against the energy savings of such a device.

Figure 42. Conceptual Example of Heating under both hybrid and all-resistance operation



The difference in heating load in all-resistance vs hybrid mode may be small and therefore difficult to detect. It may also be nonlinear. Moreover, only a fraction of the added negative heat gain may contribute to added heating. The typical noise created by variable weather patterns and human behavior may also swamp the signal we are trying to measure.

Due to these constraints and issues, five sites in the Ecotope study were selected on the basis of an orderly relationship between heating and weather, one that suggested no unusual occupant behavior or unmetered heating sources. Ideally these sites would have been selected at random, but the more targeted approach was justified for two reasons:

1. With only five sites for this test, the results are necessarily exploratory and suggestive – we could not declare a definitive answer from such a small sample.
2. The estimated magnitude of the heating interaction (given the physics of the situation) is somewhat small relative to the normal variation in daily heating at the observed time span of a partial heating season. We felt it unjustified to perform the test at sites where we believe strongly a priori that the space heating impact will be obscured by the natural variation in heating energy.

Flip-Flop Analysis Process and Findings

The idea was to estimate the annual impact on space heating by learning a heating signature through degree day regression in both operating modes, and then examining the difference when applied to a typical meteorological year (TMY). Degree day regression is somewhat unusual in the world of regression methods, as the data themselves are a function of one of the parameters to estimate. Degree day regression estimates a degree day base and a balance point, but the degree days themselves are a function of the degree day base. This quirk invalidates the analysis of covariance (ANCOVA) method that is typically used to assess the statistical significance of a treatment effect, adjusting for a real-valued variable (here the "treatment" is installation of a HPWH, and the adjustment variable is degree days).

As such, we resort to the more numerically-motivated permutation test, to investigate whether the estimated space-heating impact is statistically distinguishable from noise in the heating signature. The somewhat unusual sounding null hypothesis in this case is that the heat pump water heater does not affect space heating – before expounding on an estimated impact, we want to make sure that it is beyond the bounds of what could happen only by chance. The most convenient method to test this null hypothesis with degree day regression is to permute the water heater status labels. If the space heating impact is indistinguishable from regression noise, then the labels of operating mode are irrelevant, and can be rearranged without consequence. We may then build a null distribution by iteratively permuting the operating mode labels, and estimating the annual space heating difference.

Figure 43 shows the degree day regression for site 90051, an 80 gallon AO Smith Voltex water heater, installed in a Spokane conditioned basement. This is the one flip-flop site that yielded conclusive, credible results. The plot is shown with degree days base 60° F on the x-axis, and electric resistance daily kWh on the y-axis. The points are color-coded according to whether the HPWH was operating in heat pump mode, or resistance mode. It appears the energy use was slightly higher, adjusting for degree days, in heat pump mode as compared to resistance mode, as one would expect. Figure 44 shows the estimated annual space heating difference, and significance. The dashed red line shows the estimate calculated from observed data, and the shaded density represents the empirical null distribution. We estimate the space heating impact of the HPWH at this site to be 1,500 kWh for a typical meteorological year, and that estimate is clearly distinguishable from noise in the regression.

Figure 43 shows the degree day base for the compressor-only days was higher than the resistance mode days, as expected. The vertical axis is heating energy in kWh/day. However, the estimated space heating slope for the compressor-only days was estimated as smaller than that for the resistance mode days. This is a counterintuitive, and likely a relic of sampling variability. We know that, theoretically, the addition of an interior unducted heat pump water heater should act like negative internal gains; the addition of an interior ducted heat pump water heater should look like a large exhaust fan. Both cases should theoretically increase the degree day base – either directly removing heat, or indirectly increasing the heating load. In addition, heat pump water use is highly seasonal, with greater energy demands in the winter than the summer. Since heat pump water heater runtime is correlated with cold outdoor temperatures – high space heating load – we also expect the degree day regression heating slope to increase with the HPWH in compressor-only mode. Thus, it is odd that in the flip-flop sites the space heating slope was estimated as lower in compressor mode, but with so few sites it is likely attributable to sampling variability.

The flip flop test, while an interesting exercise, apparently lacks the power to declare emphatic results, save for in a study of prohibitive size and expense. Due to the large natural variation in heating energy, relative to the magnitude of the space heating impact of a heat pump water heater, a conclusive statistical study of space heating impact would require a much larger sample of units observed for much longer monitoring periods; and the nature of the flip flop test makes the data unusable for most other purposes. In addition, great care must be taken to ensure a full range of outdoor temperatures in both operating modes. The flip flop testing in this project ultimately had a single interval of resistance heat, and so the regressions were not “anchored” at low degree days as well as they could have been (which likely led to the estimates of decreased heating slope). There is a tradeoff, though, between rapid switching to ensure a broad distribution of outdoor temperatures, and switching so rapidly as to not accurately assess the impact, due to transient effects of thermal mass. Switching operating modes on a weekly basis for a full heating season is probably the optimal configuration for the test.

Figure 43. Flip-Flop Degree Day Regression

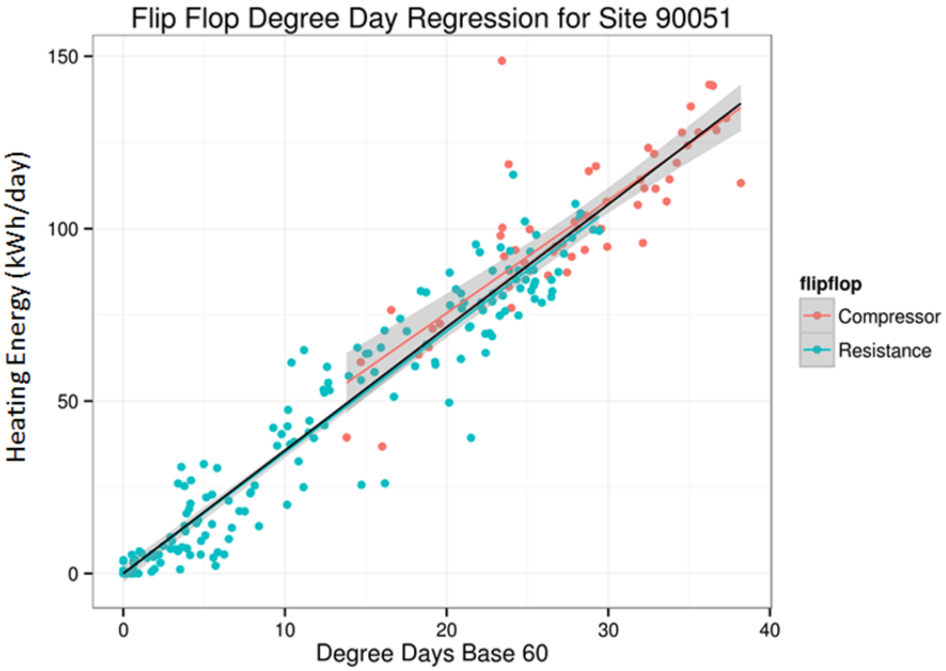
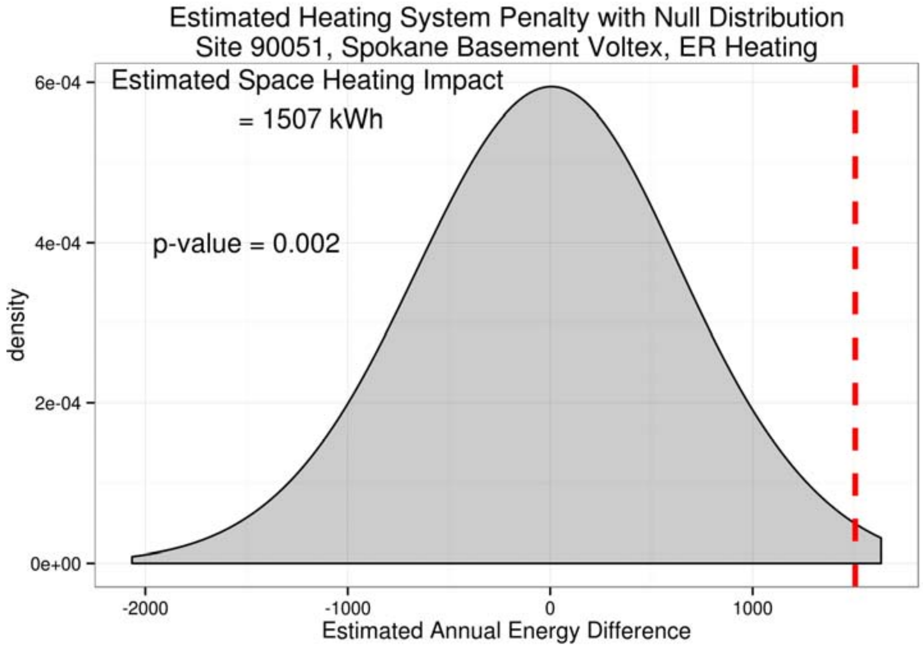
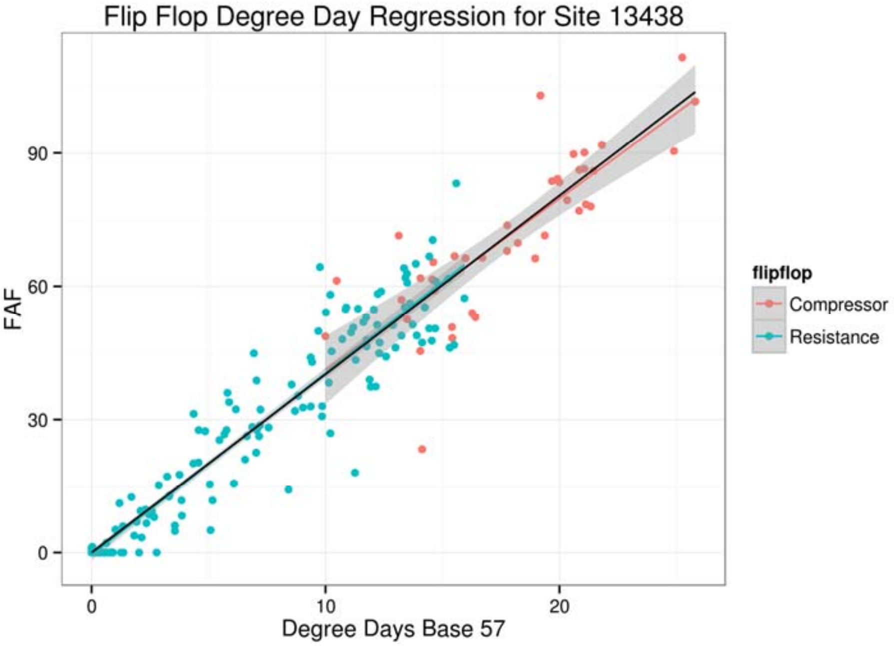
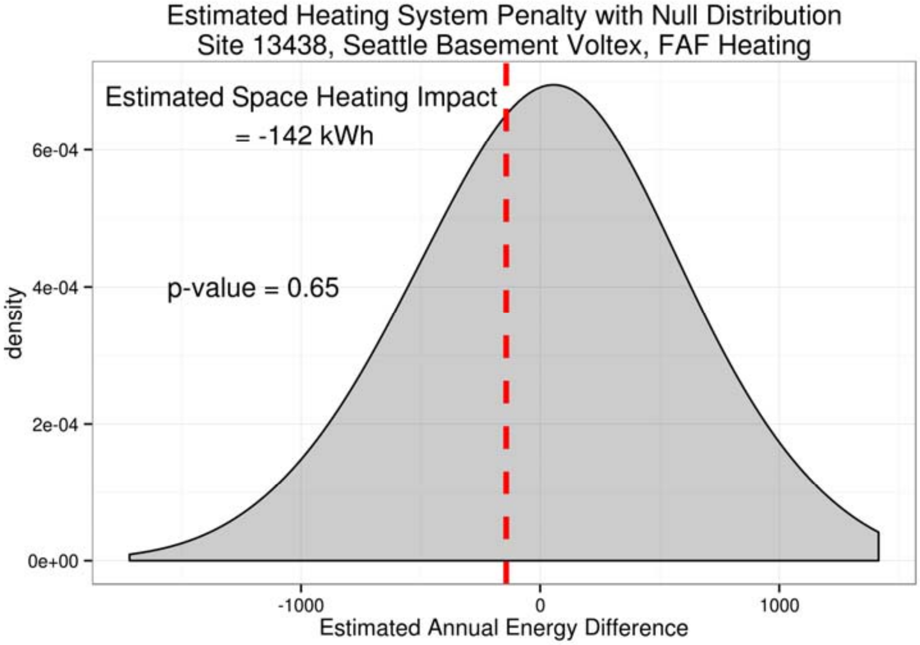


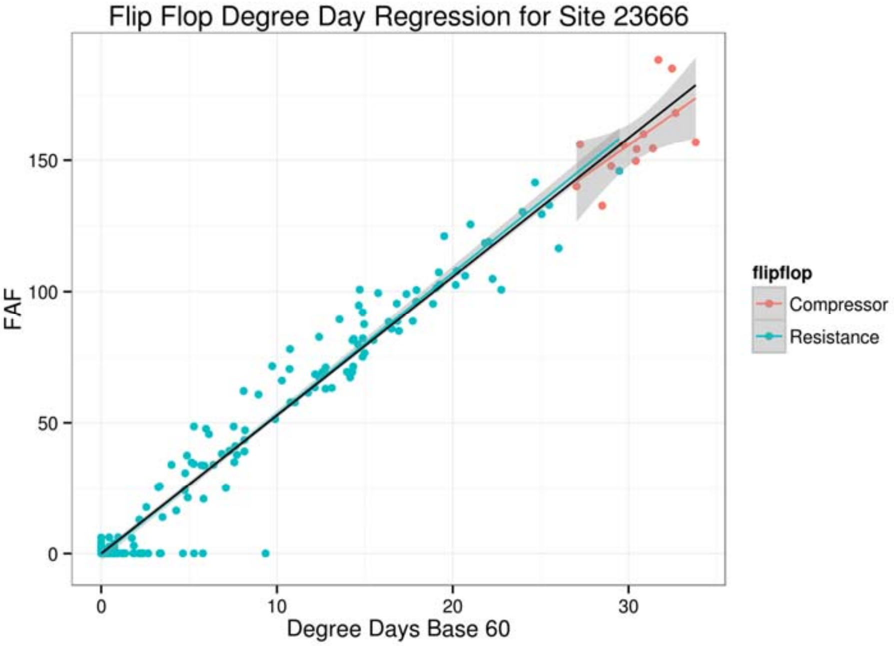
Figure 44. Estimated Space Heating Impact, One Site

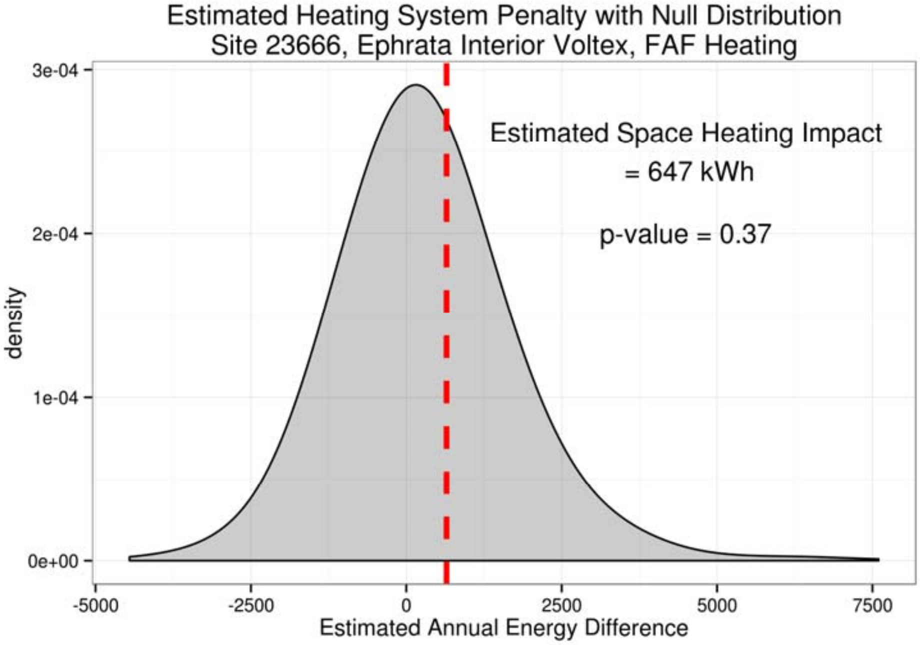


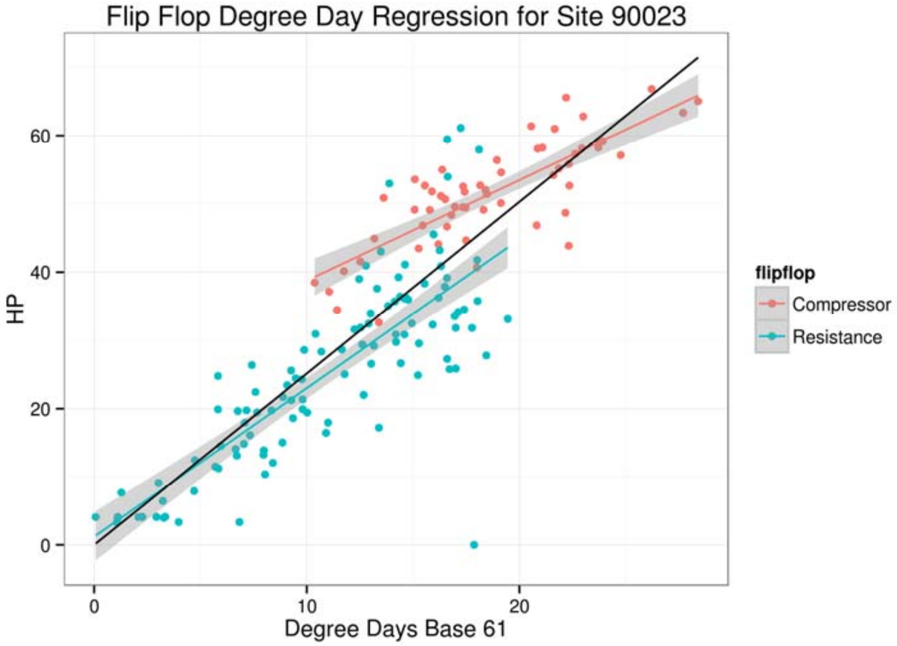
Similar graphics for the remaining flip flop sites follow.

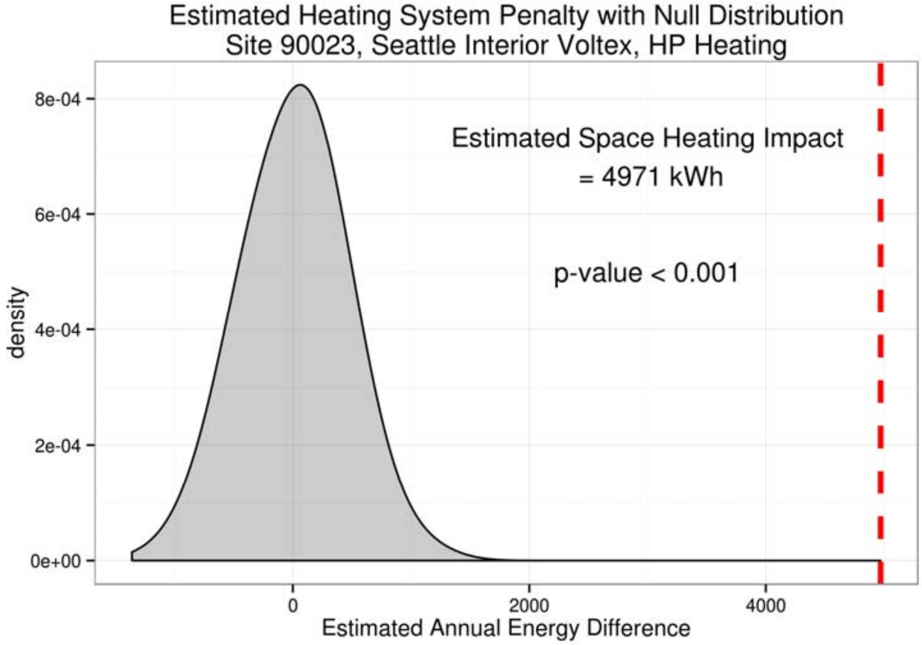

Flip Flop Degree Day Regression for Site 13438
FAF
90
60
30
0
0
10
20
Degree Days Base 57
flipflop
Compressor
Resistance


Estimated Heating System Penalty with Null Distribution
Site 13438, Seattle Basement Voltex, FAF Heating
Estimated Space Heating Impact
= -142 kWh
p-value = 0.65
density
6e-04
4e-04
2e-04
0e+00
-1000
0
1000
Estimated Annual Energy Difference


Flip Flop Degree Day Regression for Site 23666
FAF
150
100
50
0
0
10
20
30
Degree Days Base 60
flipflop
Compressor
Resistance


Estimated Heating System Penalty with Null Distribution
Site 23666, Ephrata Interior Voltex, FAF Heating
Estimated Space Heating Impact
= 647 kWh
p-value = 0.37
density
3e-04
2e-04
1e-04
0e+00
-5000
-2500
0
2500
5000
7500
Estimated Annual Energy Difference


Flip Flop Degree Day Regression for Site 90023
HP
0
20
40
60
0
10
20
Degree Days Base 61
flipflop
Compressor
Resistance


Estimated Heating System Penalty with Null Distribution
Site 90023, Seattle Interior Voltex, HP Heating
density
8e-04
6e-04
4e-04
2e-04
0e+00
Estimated Space Heating Impact
= 4971 kWh
p-value < 0.001
0
2000
4000
Estimated Annual Energy Difference

Appendix E: Draw Profiles

Determining the Draw Profiles

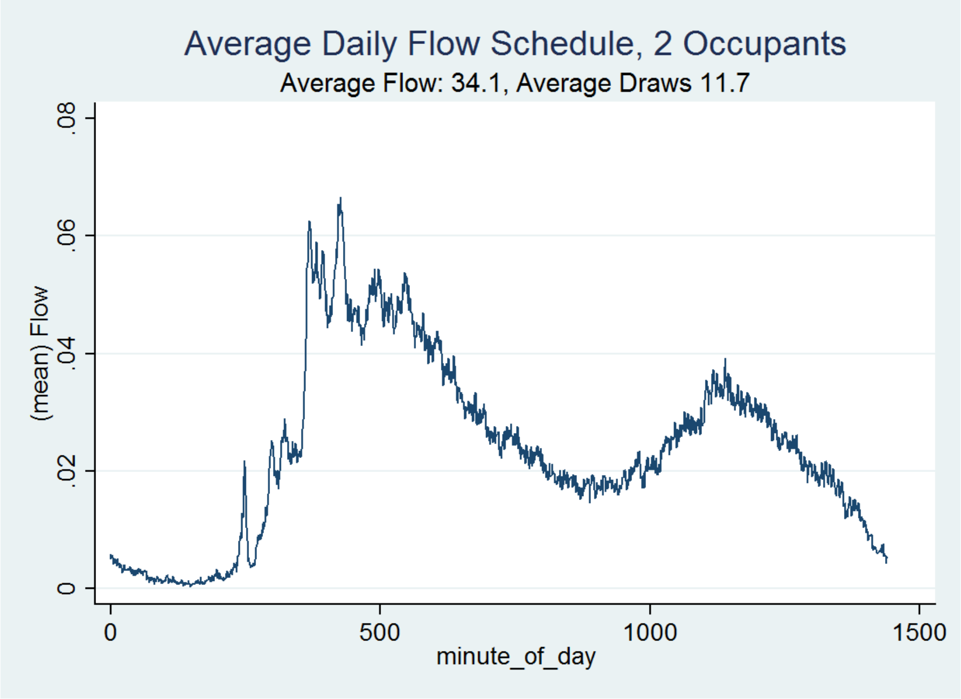
Occupant hot water draw patterns were observed with the main purpose of creating representative profiles to be used in water heater simulations. Daily hot water draws are known to be highly variable across households and within households. For example, a 4-person household will often use more water in a given day as the bathing requirements are greater than a 2-person household. Further, appliances such as dishwashers and clothes washers don't operate every day giving a great deal of variation to draws within a household.

The draw data available in the dataset has a resolution of 1 gallon at 1-minute intervals.[13] A one gallon resolution is enough to capture showers, appliance use and dishwashing events but it misses the smaller draws common to uses like hand washing. To be clear, the flow meters are totalizing – that is, they count one gallon after it has flown past even if that occurred in three, one-third gallon events. Fortunately, in the scheme of a large tank (50-80 gallons) of stored hot water, the finer resolution is not necessary. Water tanks do not respond to every, small flow event. Instead, they generally operate in response to two temperature sensors installed in the lower and upper third of the tank. In other words, the water heater doesn't turn on every time 0.5 gallons of hot water is used. Instead, the tanks operate on an apparent delay. They wait for enough cold water to build up in the tank before turning on to heat it. To understand this behavior, one gallon resolution is small enough.

The primary objective was to describe how occupants use their hot water in such a way that could be passed through a simulation. It was apparent that any method simply attempting to average the activity of several data sources would fall short of a realistic approximation. The following figure shows the draw activity averages across all days of all 2 occupant households. The mean Flow on the y-axis is given as gallons per minute. Overall, the shape is representative of average household use but not of any one day at one house. Consequently, we chose to ascertain the descriptive characteristics of draw patterns on a daily basis, summarize those, and craft new, "typical" profiles matching those characteristics.

---

[13] Some of the studies had finer resolution. For simplicity, we opted to work at the smallest, common scale.



Ecotope determined the input required for a simulation to be the time, duration, and magnitude of draws. It is important that these draws be temporally distinct in order to account for water tank recovery and strain. For this reason draws were subcategorized into small, medium, and large draws at 1-2, 3-9, and 10+ gallons, respectively.

While small, single draws, in isolation, are not particularly important to storage water heater behavior, the clustering of many draws is. It is the eventual accumulation of draws that stresses the water heater and forces it to reheat the cold water accumulating at the bottom of the tank. Consequently, the analysis technique focused on identifying distinct clusters (or windows) of water draws.

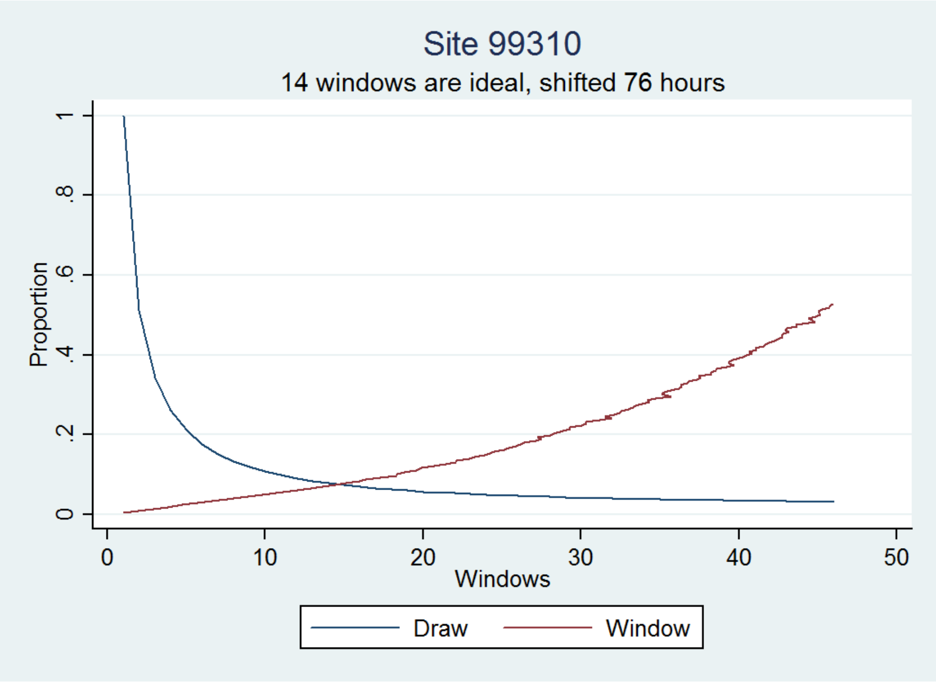
The following describes the method in detail:

Based on histograms of water flow it was determined that the draws would need to be assigned to distinct windows of activity. Each window of activity would have its own draw profile.

The first task was to assign a number of windows of activity to a given time interval (e.g. how many periods of use are there in a given day?). For this process each number of windows was equally sized in such a way that encompassed the entire interval in question. For example, three windows fit over a day would each be eight hours long; the first window beginning at midnight, the second at 8am, and the third at 4pm. Additionally, to account for the realistic probability that a window of activity would encompass midnight, the process was applied to artificially shifted days to begin at the hour of lowest mean use. All results were shifted back after windows were assigned.

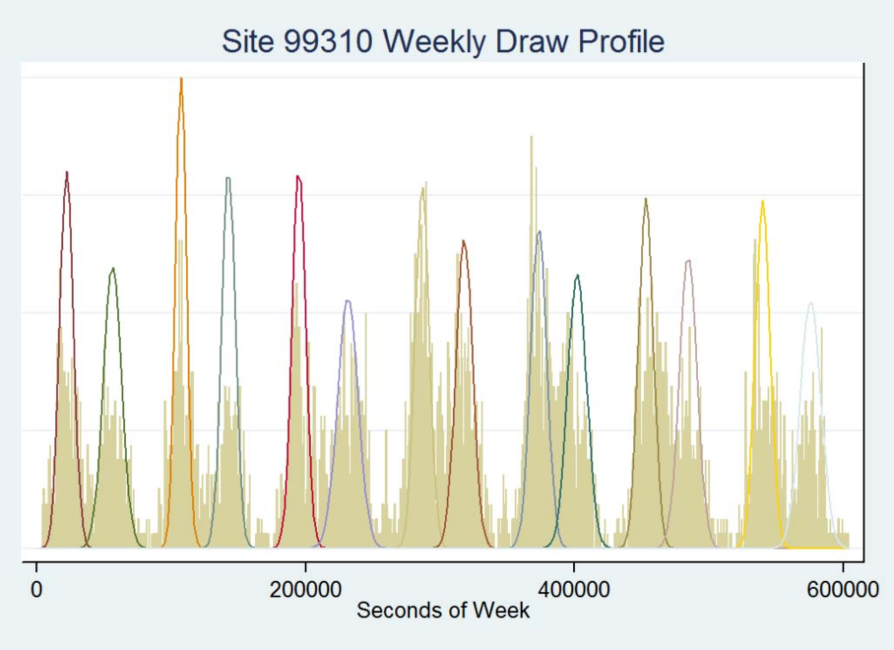
Each potential number for window of activity was assigned a ratio of draws per window to draws per time interval, as well as a percentage of windows with activity. Both of these heuristics serve to assess the integrity of the choice of windows. The ratio of draws per window to draws per time interval will decrease as the numbers of windows increase. The percentage of windows with activity will increase as the numbers of windows increase. Both heuristics range from 0 to 1 so the point at which they intersect is believed to be the optimal number of windows for that time interval.

Figure 45. Clustering Water Draws in to Windows of Activity



Having decided on the number of windows allows for further manipulation of those windows. In order to obtain an accurate description an iterative process allows the windows to move their temporal center to the time of median activity as well shrink the duration of each window to more closely encapsulate the draws. This iteration continues until the window centers stop adjusting according to a threshold. This process also allows windows to be dropped if insufficient activity is found at the window's new time and span. Figure 46 shows the final results of the process for site 99310 over an entire week. Each colored line is the window identified with the technique and overlaid on top of the actual draws. With the number, temporal center, and span of each window we can easily summarize draw count, size, and volume.

Figure 46. Identifying Clusters of Draw Activity – Weekly Basis.



Now comes the time to force these statistics into resembling what could be recognized as useful information. After rounding the average number of draws in each window to be an integer; large, medium, and small draws are fabricated according to their proportionate representation. Then assign individual draws using a normal distribution centered on the window center with a third of the time span as the standard deviation. Each draw is diffused so as not to draw at a physically impossible rate in an observational period.

A short sketch of the activity follows:

Selecting the number of windows:

1. Bridge one observation to find semi-uninterrupted flow.
2. Mark beginning of flows as draws.
3. Shift time of interval to begin at hour of lowest use.
4. Create each number of windows, each with equal size summing to the whole time interval.
5. Test each number of windows for ratio of draws per window to draws per time interval.
6. Test each number of windows for percentage of windows with activity.
7. Select number of windows where the two previously described heuristics best align.

Describing window activity:

An iterative process of:

1. Adjusting center of windows at temporal median of draws.
2. Adjusting window span to first and last draw within window.
3. Removing windows with insufficient activity.

4. Return the following stats for each window:
    a. Mean temporal center
    b. Mean time span
    c. Mean flow
    d. Mean number of draws
    e. Mean small, medium, and large flow
5. Returns daily draws and flow as well

Fabricating typical draw profile:

1. Force the number of draws in each window to be an integer.
2. Assign large, medium, and small draws at 10, 3, and 2 gallons respectively according to their proportion per window. Giving priority by size. Unaccounted gallons are assigned to medium draws.
3. All draws are forced to integer gallon flow rates per minute.
4. Assign individual draws using a normal distribution centered on the window center with a third of the time span as the standard deviation.
5. Disperse each draw so that no one minute observation draws more than 4 gallons.

Draw Characteristics

The typical draw characteristics are summarized on a daily basis in Table 24 and a weekly basis in Table 25. The total flow scales directly by a factor of seven when moving from the daily to weekly tables. Due to the variability in water use, the other quantities do not necessarily scale directly. It is precisely this variability that drives the need for water heater energy use simulations to use more than just one day's worth of draw profiles.

Table 24. Daily Draw Characteristics

Occupant Count	Clusters per Day	Gallons per Day				Draw Count per Day				Sites	Days Metered
		Total Flow	Small Flow	Mid Flow	Large Flow	Total Draws	Small Draw	Mid Draw	Large Draw		
1	3	23	6	5.5	11.5	6.4	4.5	1.1	.7	7	2160
2	5	34	10.4	7.7	16	12.6	9.7	1.7	1.1	32	10602
3	5	46	13.8	10.7	21.5	15.3	11.9	2.2	1.3	14	5193
4	5	57	13.8	12.5	30.9	14.6	10.7	2.3	1.7	13	4440
5+	5	72.4	14	14.7	43.9	18.5	12.6	3.2	2.7	10	2448

Table 25. Weekly Draw Characteristics

Occupant Count	Clusters per Week	Gallons per Week				Draw Count per Week				Sites	Weeks Metered
		Total Flow	Small Flow	Mid Flow	Large Flow	Total Draws	Small Draw	Mid Draw	Large Draw		
1	11	161	42.3	38.7	80.9	50.7	36.1	9.1	5.4	7	279
2	17	238	74.2	54.8	114.5	105.6	81.3	14.7	9.6	32	1325
3	17	322	97	75	150.8	132.8	102.6	19.2	11	14	619
4	17	399	97.9	88.7	219.9	124.6	91	19.4	14.2	13	550
5+	18	506.8	100.9	106.4	317.7	139.5	95.2	24.2	20.1	10	349

Table 26 shows how the clusters and the events within each cluster are distributed within the typical day described by Table 24. Likewise, Table 27 shows the same information on a weekly basis.

Table 26. Daily Draw Characteristics by Event Cluster

Occupant Count	Total Flow (Gal)	Median Time (Hr of Day)	Cluster Span (Minutes)	Draws per Cluster (count)	Flow per Cluster (Gal)
1	23	7	74.1	2.6	12.3
		13.9	57.1	1.7	5
		20	58.3	1.9	4.8
		.	.	0.3	0.9
2	34	6.6	35.7	1.9	7.8
		9.4	61.3	2.9	9
		13.5	65.7	2.5	5.5
		18.9	75.7	3.2	6.8
		21.8	40.1	1.9	4.5
		.	.	0.3	0.5
3	46	7.1	48.5	2.6	10.9
		10.5	71.5	3.1	9.2
		15.5	75.1	3	7.3
		19.1	81.6	4	11.2
		22.2	43.5	2.1	5.8
		.	.	0.6	1.6
4	57	7.4	63.4	3.1	18.4
		10.5	67.3	2.9	12.4
		15.4	76.2	3	8.6
		19	72	3.5	11.8
		21.8	29.5	1.6	4.3
		.	.	0.5	1.7
5+	72.4	6.9	47.2	2.5	14.6
		9.7	78.6	3.8	16.9
		14	91.7	3.7	11
		18.2	104.9	4.9	17
		21.7	75.1	3.2	12.3
		.	.	0.4	0.8

Table 27. Weekly Draw Characteristics by Event Cluster

Occupant Count	Total Flow (Gal)	Median Time (Hr of Wk)	Cluster Span (Minutes)	Draws per Cluster (count)	Flow per Cluster (Gal)
1	161	11.3	432.5	7.6	20.9
	161	27.7	324.3	4.2	17.2
	161	40.6	242.1	3.9	9.4
	161	54.1	195.3	3.6	16.5
	161	65.9	171.6	3.1	7.3
	161	78.9	252.3	4.1	17.6
	161	101.9	214.5	3.9	16.9
	161	113.8	179.1	3.2	7.7
	161	127.9	282.2	4.8	18
	161	150.3	169.3	3.5	8.9
	161	159.1	252.1	4.3	11.2
	161	.	.	4.5	10.4
2	238	6.4	238.6	6.9	20.2
	238	17.1	269.6	6.9	13.5
	238	28.9	102.7	4	12.9
	238	33.6	137.5	4.4	9.4
	238	41.5	224.8	6.1	11.5
	238	54.9	256.3	7.5	20.3
	238	64.9	271.1	6.6	11.9
	238	78.9	147.7	5.8	15.8
	238	84	140	4.8	9.4
	238	90.1	183.8	5.1	10
	238	104.1	278.8	8.6	21.7
	238	112.9	285.3	7.9	14.9
	238	126.5	222.3	6.7	20.1
	238	135	166.3	5	9
	238	139.5	89.9	3.3	6.7
	238	150.5	248.3	6.9	20.1
	238	161.2	254.5	6.7	12.3
	238	.	.	2.3	3.9
3	322	7.7	303.3	10.8	26.6
	322	15.9	343.4	12.4	26.8
	322	28.9	166.7	5.9	17.1
	322	35.8	170.4	5.3	11.6
	322	41.3	224.3	7.9	17.3
	322	53.4	238.9	6.9	19.3
	322	64.3	331.1	10.3	22.5
	322	77	197.5	6.3	19.3
	322	85.5	187.2	6	12.4
	322	90.2	157.8	5.4	13.6
	322	101.1	228.7	7	20.9
	322	112.5	295.9	8.8	19.7
	322	125.5	230.7	7	20.3
	322	134	187.9	6	12.4
	322	138.8	114.9	3.9	8.6
	322	151.6	326.1	10.7	27.7
	322	160.4	301.1	9	19.3
	322	.	.	3.4	7.4
4	399	5.3	249.7	8.3	36
	399	15.3	321.2	10.2	27.8
	399	28.8	154.8	6.1	26.6
	399	35.8	130.1	4.3	11.2
	399	40.8	180.6	6.5	17.1
	399	53.3	221.2	7.4	34

	399	63.7	289	9.2	24.5
	399	77	195.1	6.9	31.1
	399	85	154.7	5.2	11.1
	399	89.4	137	4.9	12.8
	399	101.4	244	7.9	33.3
	399	111.6	271.7	7.4	17.5
	399	126.6	221.1	8.4	30.1
	399	133.2	166.5	6.4	16.6
	399	138.4	92.7	3.2	7
	399	151.5	293.2	10.7	38.5
	399	159.7	293.2	9.9	27.6
	399	.	.	1.7	3.6
5+	506.8	6.6	257.3	8	36.3
	506.8	14.6	202.7	7.3	19.9
	506.8	19.2	136.4	5.5	18.9
	506.8	31.1	294.8	9.1	40.6
	506.8	41	315.5	10.6	35.1
	506.8	54.4	209.8	7	35.5
	506.8	61.8	167.4	5.7	17.5
	506.8	67.2	165.4	5.4	17
	506.8	80.4	271.4	8.6	37.1
	506.8	89.2	298.5	8.8	34.3
	506.8	102.5	117.3	5.2	26.1
	506.8	107.9	157.3	5.5	17.3
	506.8	114.2	234.7	8.1	27.1
	506.8	127.1	284	8.9	38.3
	506.8	136.7	328.3	11.2	37.3
	506.8	149.3	115.1	4.6	25
	506.8	154.5	169.2	5.8	17.3
	506.8	161.1	291.6	10.3	35.4
	506.8	.	.	3.9	9

Draw Profiles

The draw profiles crafted based on the characteristics above are provided in the following tables. The exception is for the “Five+ Occupant” schedules which are additionally informed by the need to have at least one, large draw pattern for simulation purposes. The average daily draw volume of 5+ occupant households in the field sample was 72 gallons was for that specific distribution of household sizes. The distribution in the general population is different and can have more people and more water use. Further, a calibration exercise overseen by the RTF HPWH evaluation subcommittee concluded that having a larger draw available would more accurately simulate the amount of resistance heat use observed in the field data. Consequently, the 5+ person draw given on both the daily and weekly basis below uses 85 gallons per day on average.

Daily Profiles

One Occupant	
Minute of Day	Flow (Gallons)
394	2
408	2
430	3
431	3
432	4
835	3
844	2
1199	3
1202	2

Two Occupants	
Minute of Day	Flow (Gallons)
384	3
385	3
404	2
555	2
569	2
571	3
572	3
798	3
800	2
1091	2
1112	3
1113	2
1306	3
1309	1

Three Occupants	
Minute of Day	Flow (Gallons)
418	2
427	2
428	3
429	4
603	3
604	2
614	2
630	2
919	2
925	3
934	2
1141	2
1143	2
1151	3
1152	4
1162	2
1333	4
1343	2

Four Occupants	
Minute of Day	Flow (Gallons)
415	2
421	3
422	3
438	2
444	3
445	3
446	4
615	2
630	3
631	3
632	2
651	2
924	3
925	2
933	2
947	2
1117	2
1127	3
1128	3
1137	2
1158	2
1310	1
1318	3

Five+ Occupants	
Minute of Day	Flow (Gallons)
410	3
411	2
415	2
422	3
423	3
424	4
557	2
583	2
592	3
593	3
601	3
602	3
603	4
803	2
832	2
855	2
862	3
863	4
1065	2
1076	2
1091	3
1092	2
1100	2
1153	3
1154	3
1155	4
1264	3
1265	3
1266	4
1282	2
1309	2

Weekly Profiles

One Occupant			
Minute of Week	Flow (Gallons)	Minute of Week	Flow (Gallons)
642	1	3225	2
675	4	3245	3
686	4	3917	4
690	1	4680	1
698	1	4704	3
730	1	4705	3
758	1	4706	3
785	1	4707	3
797	3	4714	1
798	3	4716	3
799	4	6059	3
1582	1	6087	1
1644	1	6100	1
1659	3	6114	3
1660	2	6115	3
1663	1	6116	3
1695	1	6117	2
1698	3	6812	4
1699	3	7662	2
1700	3	7673	3
1701	3	7674	3
1702	3	7675	3
1703	2	7676	2
1712	1	7704	2
1758	1	7763	3
2383	2	8988	2
2406	2	9030	3
2430	2	9031	2
2450	2	9075	2
2454	4	9528	2
2478	2	9541	2
3160	3	9586	2
3161	3	9646	3
3162	3	9647	2
3163	2		

Two Occupants			
Minute of Week	Flow (Gallons)	Minute of Week	Flow (Gallons)
326	1	5031	3
346	1	5044	2
354	4	5399	2
365	3	5404	2
366	3	5424	2
367	3	5427	4
368	2	6164	2
404	1	6213	2
991	3	6214	2
992	2	6226	3
1006	2	6227	3
1049	2	6228	3
1075	2	6229	3
1098	2	6230	3
1721	3	6231	3
1722	3	6232	2
1723	4	6285	2
1758	1	6307	3
1768	1	6308	3
1969	3	6707	2
1970	2	6757	2
2021	2	6759	2
2057	2	6765	2
2426	2	6773	4
2479	2	6781	2
2501	2	6791	2
2524	2	7539	4
3231	1	7557	3
3247	1	7558	3
3254	1	7559	3
3272	1	7560	3
3294	3	7593	1
3295	3	7605	1
3296	3	7608	1
3297	3	8055	3
3310	3	8099	1
3311	3	8377	3
3331	1	8981	3
3842	4	8982	3
3869	2	8983	3
3888	2	8984	2
3929	2	8996	1
3967	2	9018	4
4725	1	9054	1
4728	3	9073	1
4729	3	9660	2
4730	4	9673	2
4734	3	9689	4
4744	1	9702	2
5004	2	9735	2
5005	2		

Three Occupants					
Minute of Week	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**
395	1	2510	3	6090	3
422	3	2511	4	6128	2
423	3	2524	2	6654	2
424	3	3199	3	6660	2
425	3	3200	2	6673	2
426	4	3223	1	6738	2
458	1	3231	1	6741	2
463	3	3238	1	6778	2
464	2	3261	3	6784	3
472	1	3262	3	6785	3
485	1	3263	4	6786	2
489	1	3283	1	7458	3
491	1	3802	1	7459	3
500	3	3810	1	7460	3
501	2	3823	2	7461	2
514	1	3872	1	7491	1
524	1	3890	1	7504	4
528	1	3910	1	7549	1
876	4	3942	3	7552	1
892	1	3943	3	7593	1
933	1	3944	4	7994	2
951	1	3947	3	8034	4
971	1	3948	2	8037	2
975	1	4574	4	8069	2
976	1	4581	1	8080	2
985	1	4605	1	8307	1
1008	4	4610	1	8332	3
1029	1	4641	3	8990	1
1041	3	4642	3	9011	1
1042	3	4643	2	9040	1
1043	3	4644	2	9050	3
1044	3	5070	2	9051	3
1045	3	5103	2	9073	3
1046	3	5136	2	9074	3
1047	3	5167	2	9075	3
1048	3	5171	4	9076	3
1081	1	5372	2	9077	2
1718	4	5375	2	9087	1
1726	1	5410	3	9113	1
1742	1	5411	3	9152	1
1750	3	5412	2	9186	1
1751	3	5430	2	9610	2
1752	4	6031	1	9622	2
1762	1	6037	3	9623	2
2156	2	6038	2	9638	2
2159	4	6071	1	9644	2
2425	2	6074	1	9675	3
2433	2	6087	3	9676	4
2438	2	6088	3	9714	2
2467	2	6089	3		

Four Occupants							
Minute of Week	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**
271	4	2432	2	6034	1	9100	3
292	1	2439	3	6073	3	9101	3
299	1	2440	4	6074	4	9102	3
308	1	2449	2	6104	1	9103	3
317	1	3161	3	6106	1	9104	3
319	1	3162	3	6114	1	9105	3
325	1	3178	1	6121	3	9106	3
329	1	3183	1	6122	3	9107	3
347	4	3203	3	6123	3	9108	3
348	3	3204	3	6124	3	9109	3
349	3	3205	3	6125	3	9110	2
350	3	3206	3	6126	3	9114	2
351	3	3207	3	6127	3	9134	2
352	3	3208	3	6157	1	9486	1
353	3	3209	4	6634	2	9517	1
354	3	3213	1	6689	2	9520	4
355	3	3231	1	6696	2	9525	1
873	1	3244	1	6713	3	9572	1
884	3	3748	1	6714	2	9584	1
885	3	3754	3	6732	2	9591	1
894	1	3755	3	6736	2	9617	1
905	1	3756	4	6751	2	9626	3
930	3	3764	3	7506	1	9627	3
931	3	3765	3	7542	3	9628	3
932	3	3800	1	7543	3	9629	3
933	4	3805	1	7582	1	9667	4
941	1	3821	1	7607	1		
953	1	3851	1	7611	3		
971	1	3858	1	7612	3		
1035	1	4571	3	7613	3		
1714	3	4572	3	7614	3		
1715	3	4573	3	7615	3		
1716	3	4574	3	7616	3		
1717	3	4575	3	7620	1		
1718	3	4576	3	7639	1		
1719	2	4577	3	7963	2		
1728	1	4613	1	7978	2		
1741	1	4628	1	7985	2		
1746	1	4634	1	7996	2		
1771	3	4679	1	8001	3		
1772	2	4685	3	8002	3		
2133	2	4686	3	8007	2		
2148	2	5050	3	8314	3		
2156	3	5084	2	9010	4		
2157	2	5339	2	9024	2		
2169	2	5361	2	9055	4		
2352	2	5362	2	9058	2		
2378	2	5389	3	9061	2		
2391	2	5390	4	9067	2		

Five+ Occupants									
Minute of Week	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**	**Minute of Week**	**Flow (Gallons)**
337	1	1922	3	3722	3	6450	2	8224	3
351	3	1923	2	3723	3	6457	2	8225	3
352	4	1941	3	3724	2	6485	3	8226	3
386	1	1942	2	4003	3	6486	3	8227	2
390	1	1945	3	4004	3	6833	1	8233	2
392	1	1946	3	4034	2	6845	1	8260	2
416	1	1947	3	4073	2	6859	1	8935	1
423	3	1948	3	4744	3	6860	1	8949	1
424	3	1949	3	4745	3	6861	1	8951	3
425	3	1950	3	4746	3	6896	3	8952	3
426	3	1951	3	4747	3	6897	4	8953	3
427	2	1952	4	4748	3	6904	3	8954	3
432	3	2414	3	4749	3	6905	3	8955	3
433	3	2415	2	4750	3	6906	3	8956	3
434	3	2418	1	4751	3	6907	3	8957	3
435	3	2422	1	4752	4	6908	4	8958	4
436	2	2427	1	4763	1	6948	1	8970	1
844	3	2433	1	4796	1	7559	3	8972	1
845	2	2444	1	4807	1	7560	2	9213	4
851	3	2464	1	4817	3	7572	1	9226	1
852	3	2485	3	4818	4	7602	1	9229	3
853	3	2486	3	4822	1	7605	1	9230	3
854	2	2487	3	4837	1	7608	1	9231	3
857	1	2488	3	4927	1	7613	1	9232	2
858	1	2489	3	5246	4	7638	3	9271	1
868	1	2490	3	5327	1	7639	3	9279	1
929	1	2491	3	5331	1	7640	3	9291	1
950	1	2546	3	5340	1	7641	3	9296	1
1129	2	2547	2	5360	1	7642	3	9566	3
1137	3	2567	1	5369	1	7643	2	9567	2
1138	3	3224	1	5421	3	7644	3	9608	3
1148	3	3241	1	5422	3	7645	3	9609	2
1149	3	3246	1	5423	3	7646	3	9619	1
1150	4	3255	1	5424	3	7647	3	9621	1
1172	2	3263	3	5425	3	7648	3	9636	1
1197	2	3264	3	5426	3	7649	2	9644	1
1773	1	3307	3	5427	3	7654	3	9653	1
1812	1	3308	3	5428	2	7655	2	9664	1
1828	1	3309	3	5437	4	7675	1	9679	1
1855	3	3310	3	6124	1	7724	1	9725	1
1856	3	3311	2	6129	3	8135	3	9782	3
1857	3	3349	3	6130	2	8136	2	9783	3
1858	3	3350	3	6156	1	8147	3	9784	3
1859	3	3351	3	6167	1	8148	2	9785	3
1860	3	3352	3	6186	3	8154	2	9786	3
1861	3	3353	2	6187	3	8182	2	9787	3
1862	4	3684	2	6188	3	8215	2	9788	3
1875	1	3686	2	6189	3	8220	3	9789	2
1892	1	3696	2	6190	3	8221	3		
1902	1	3715	4	6191	3	8222	3		
1919	1	3721	3	6192	3	8223	3		

Appendix F: Measuring Tank Heat Loss

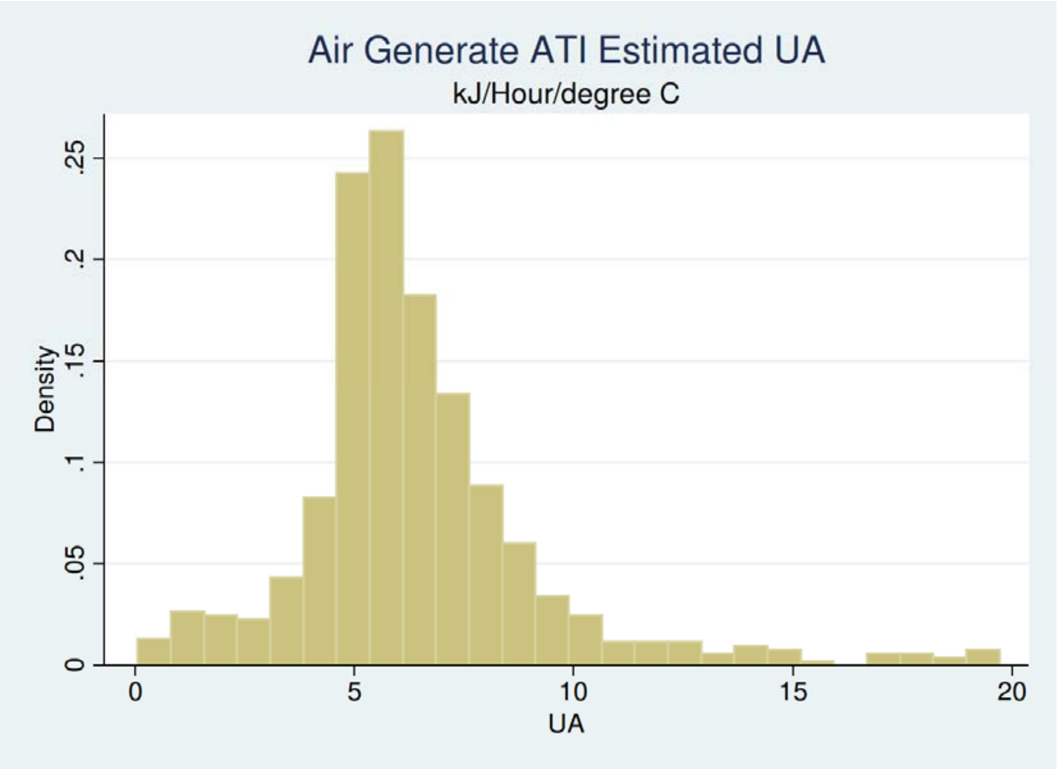
The tank and pipe system heat loss rate is necessary to calculate standby losses, which are necessary to compute the heat pump COP. However, estimating the tank and pipe system heat loss rate from field data offered challenges. In contrast, lab measurements provided computations of heat loss rate under much more controlled conditions. Nevertheless, investigation of the field data showed noisy patterns that more or less boiled down to the same estimates as derived in the lab. In light of this, we proceed with the measurements of UA (kJ/hour/degree C) from the more controlled lab conditions.

Theoretically it is possible to estimate the tank and pipe system heat loss rate from the field data by finding water draws separated by a somewhat long duration of no added heat. If the tank is at roughly uniform temperature, then the temperature difference between the last delivered water of the first draw and the first delivered water of the second draw should inform a UA calculation.

As an example, consider two water draws one hour apart with no intervening added heat. The difference in temperature between the last outlet measurement of the first draw, and the first outlet measurement of the second draw, can be used for a single estimate of the heat loss per hour with the usual formula $Q = mc_p\Delta T$, using the mass of water in the tank, the heat capacity of water, and the observed temperature difference. Because this is field data, however, nothing is that easy. Instantaneous readings of outlet water temperature are tough to derive, since the temperatures “float” in the absence of flow events. During a draw event, at some point the stagnant water in the pipe – which had drifted to a temperature different from that of the internal tank water – passes the sensor and the first hot tank water contacts the sensor. That’s the measurement necessary for the UA assessment, but it’s basically impossible to capture. Discard too little of the beginning of the draw and you end up with the pipe system floating temperature; discard too much at the beginning of the draw and you may no longer be measuring the water temperature from the level of the tank of the outlet water (which is desired for the most fair comparison). Further, the data resolution is only at one minute, so we lack the granularity to approach this with finesse and nuance.

Figure 47 shows a distribution of estimated heat loss rate for the Air Generate ATI water heaters. Draws were found separated by at least one hour, with no intermediate added heat. The last recorded temperature from the first draw was compared to the second recorded temperature from the second draw (discarding the first minute of the second draw). The distribution is extremely scattered and noisy. It ultimately clusters around approximately 6 kJ/hr°C (3.2 Btu/hr°F), which is essentially identical as the lab measurement for the 66 gallon ATI tank (Larson and Logsdon 2012a). In light of these explorations, we decided to proceed using the lab measurements of tank heat loss rate for all cases.

Figure 47. Attempted Tank Heat Loss Rate Estimates



Appendix G: Assessing Noise Satisfaction

Average satisfaction aligns well with the measured decibels in the adjacent room, however, this finding needs investigation to determine if it is significant or coincidental. We took two approaches to investigate if and how sound satisfaction varies by make and install parameters: the first was a Chi-Squared test of independence in a contingency table of make and satisfaction. The second was a cumulative logit regression model.

Table 28. Counts of Occupant Satisfaction Selection by Make

	Occupant Satisfaction (1-5)					
Make	**1**	**2**	**3**	**4**	**5**	**Total**
Voltex	0	0	10	7	3	20
ATI	1	0	3	4	7	15
GeoSpring	0	0	1	5	4	10
Total	1	0	14	16	14	45

The number of homeowners selecting each level of noise satisfaction by make is displayed in Table 28. Typically, a Chi-Square test is used to investigate hypotheses of independence in two-way contingency tables such as this one. The basic idea is that, under the null hypothesis of no relationship between water heater make and occupant noise satisfaction, the counts in the table should follow a multinomial distribution with cell probabilities equal to the product of the marginal probabilities. For example, with 16 selections of satisfaction level 4 out of 45 total, and 20 Voltex water heaters out of 45 total, the expected counts in the Voltex satisfaction 4 cell would be 16 * 20 / 45 = 7.1. The observed number in that cell was 7. The observed counts and the expected counts calculated as such can be combined to develop a statistic with Chi-Square distribution in large samples under the null hypothesis of independence[14].

Under such a test of independence between occupant satisfaction and make, the calculated Chi-Squared statistic was 8.94 on 6 degrees of freedom, which led to a p-value of 0.11. Due to the somewhat small counts, the large sample Chi-Square distribution may be inappropriate, and we also computed significance through a resampling type test, which led to a p-value of 0.08. This is a somewhat statistical grey area. Probably there are actual differences in sound satisfaction between HPWH makes, but those differences are small enough to not strongly distinguish themselves in a sample of 45 responses.

In addition to testing for differences in noise satisfaction between makes, we can also investigate the effects of the two most obvious aural irritants: volume and duration. As well as the measured decibels of the HPWH in an adjacent room (presumably where the occupants spend time), homeowners would logically find a water heater that runs nonstop to be more irritating than one that mostly sits idle. Regression modeling with ordered outcome categories is a bit more difficult than with continuous outcome data. The validity of standard errors and p-values from an ordinary linear regression rely on the satisfaction of the constant variance assumption, which in general will not apply with categorical data. One convenient alternative, which often shows

---

[14] http://en.wikipedia.org/wiki/Pearson's_chi-squared_test

up in social sciences scenarios with this type of ordered survey responses, is a cumulative or ordered logit model[15].

Whereas ordinary linear regression models a relationship between a continuous outcome and some explanatory variables, an ordered logit model models the logarithm of the odds of appearing in the various categories. The basic form of the model is to presume that the log odds of appearing in a one unit higher category are a linear function of some explanatory variables.

The results of the ordered logit model were baffling. The estimated effect of an additional hour of weekly runtime was about the same as that of an additional decibel: roughly 0.97 fold lower odds of selecting a one unit higher satisfaction category. Neither effect was statistically significant, though. The fit of the model with decibels and runtime was statistically no better than simply guessing the average satisfaction level every time. However, weekly average flow was highly associated with occupant sound satisfaction. A difference of 10 gallons higher in weekly flow was associated with .95-fold lower odds of selecting a one unit higher satisfaction category. A 100 gallon higher difference in weekly flow was associated with 0.59-fold lower odds of selecting a one unit higher satisfaction category. The p-value of this association was 0.005. Figure 48 shows noise satisfaction and flow. The relationship clearly declines (this is true even without the one homeowner who selected satisfaction of 1). The figure shows a linear fit for presentation – the above statements referred to the output of an ordered logit model, not a linear regression model, but it is much easier to visualize a line through the data than visualize the odds ratios of the ordered logit model.

It seems as though flow is only a proxy for runtime, which should be the true causal factor for occupant dissatisfaction. It is not clear why the relationship with the proxy is significant, while the actual causal determinant did not significantly explain variation in occupant-reported noise satisfaction. Nevertheless, we believe that these findings suggest at the very least consideration of noise and noise mitigation in high occupancy or high flow volume households.

The effects of volume were not strong enough to achieve significance in a sample of 45 households, although it is likely that homeowners are less satisfied with noisier units regardless of runtime.

---

[15] http://en.wikipedia.org/wiki/Ordered_logit

Figure 48. Weekly Flow and Noise Satisfaction