Field Performance of High-Performance High-Capacity Heat Pumps

John Bush, OTS Energy
Ben Larson, Larson Energy Research
Christopher Dymond, Northwest Energy Efficiency Alliance
JJ Sawicki, TRC Companies

ABSTRACT

Cold climate heat pumps promise significant improvements in comfort, and savings of carbon and energy in heating climates. This study presents the result of a three-year field test of 57 high-performance high-capacity heat pumps in the Pacific Northwest. The goal of this study was to understand the real-world performance of top performance units available in 2022-3, the impact of design and system choices and how they compare to rated values.

The field study covers an equal mix of ducted and ductless whole home systems across all three Northwest climate zones. It includes a mixture of home types and installation practices. Instrumentation provided minute data on power, temperatures, auxiliary heat and for airflow in ducted systems allowing researchers to calculate real time COP.

The field study gained insight into sizing heat pumps for heating loads including with low or no electric resistance (ER) supplemental heating. It also examined impacts of thermostat setbacks on the use of ER heat. The research team also evaluated potential metrics to help select an appropriate heat pump system. The heat pumps produced average savings of 30% - 50% in whole home heating energy compared to baseline ER heating. The range of performance varied significantly with some systems clearly outperforming others, and notable performance differences between central ducted and whole home multizone systems.

Introduction

The electrification of space heating is widely recognized as a central strategy for achieving deep reductions in greenhouse gas emissions from the building sector while maintaining affordability and grid reliability. In cold and mixed-marine climates, however, concerns about performance, peak demand impacts, and customer acceptance have historically constrained the adoption of air source heat pumps. New generations of cold climate heat pumps (CCHPs) that are designed to maintain capacity and efficiency at low outdoor temperatures have entered the market with the potential to address these challenges.

The Northwest (NW), encompassing Oregon, Washington, Idaho, and Montana, with its distinctive electrical system characterized by a large share of low-carbon hydropower and a legacy of winter-peaking loads driven primarily by space heating has long been supportive of using air source heat pumps to displace or replace use of electric resistance heating. The region's utility support of heat pumps is largely the result of the Pacific Northwest Electric Power Planning and Conservation Act of 1980 (NPCC 1980), that mandates energy efficiency to be treated as a priority resource on par with generation. Space heating has consistently represented the largest source of residential electricity consumption in the region, making heating technologies a primary focus of the region's efficiency programs and research initiatives. As a consequence, the region has four decades of energy efficiency experience with heat pumps with nearly a million homes now using heat pumps as their primary space conditioning systems.

Early generations of air source heat pumps faced performance limitations in colder temperatures and were often perceived as unsuitable for the region's heating-dominated climate. At a high level, many of the utility programs have found heat pumps underperform expectations, for example Baylon et al (2005) found only 70-85% of expected savings were achieved. In response, many research studies and program evaluations have been conducted that explore equipment performance, design, ducting, and controls of heat pumps. Recent research on system performance (Winkler and Ramaraj 2023) has shown that cold climate heat pumps in closely controlled field trials still have inconsistent results. That research has shown that seasonal COPs do not always align with expectation, and use of auxiliary heat is sometimes excessive. An RDH study (2023) showed a wide range of seasonal efficiency in Western Canada, with many units operating near the rated efficiency level, but several significantly lower. The reasons for these efficiency departures can be difficult to pin down but excessive resistance heat is often a key contributor. Indeed, efforts by BPA and regional utilities to improve installation and configuration practice around controlling for electric resistance heat (and other installation and commissioning practice) have found poor compliance (Kelsven, 2019). Researchers have also focused on low-load performance of systems; one of variable speed systems' claimed advantages is the ability to run at high efficiency when loads are low, delivering greater savings than systems which cycle off and on at short duty cycles (Harley et al 2025). But it remains to be seen if systems deliver this performance reliably in a way that matches claimed performance.

The region's electric utilities are keenly interested in heat pumps that eliminate electric resistance heat where possible, and minimize it elsewhere, and in determining if peak impacts on the electric grid can be minimized through either system improvements or using demand response systems. A core question is if the latest cold climate heat pumps with advances such as variable-speed, inverter-driven compressors, enhanced vapor injection, and new refrigerants can maintain high capacities and coefficients of performance at temperatures the range of cold temperatures most relevant to the region's winter peak conditions. A system which could meet the goal of drastically reducing or eliminating resistance heat, while operating efficiently at low loads and in cooling operation, can benefit utility system operation and keep energy costs lower.

This study was a collaboration of researchers and utilities in the Pacific Northwest seeking to identify and test heat pumps expected to be among the best available at achieving these goals. The researchers sought to understand what are the essential design and installation criteria and what unique features and capabilities of new cold climate heat pumps should be emphasized in utility program specifications, contractor training and customer education.

Project Overview

In this 2023-2025 project, 57 cold-climate variable speed heat pumps were metered in locations across Washington, Oregon, Idaho and Montana. The sites included a mix of central ducted systems and whole-home, multi-zone systems. The sites were all formerly heated with electric resistance. Of the 57 systems, 38 were installed during the project, with contractors asked to size to the heat pump balance point to 20°F or below. In addition, 19 were systems which had been previously installed as part of a utility rebate program within the past three years or less. The systems prioritized for installation were identified after a modeling and analysis effort to use the performance curves provided in the Northeast Energy Efficiency Partnership (NEEP) cold climate air source heat pump (ccASHP) Product List, to identify the best expected energy and demand saving systems for the region. The key metrics for selecting eligible systems centered around:

- COP at minimum capacity, 47°F outdoor temperature
- COP at maximum capacity, 17°F outdoor temperature
- EER2
- Turn-down ratio (defined here as maximum capacity at 17°F divided by minimum capacity at 47°F)
- Capacity maintenance between the 17°F and 5°F maximum capacity data points

The details of the effort to identify filter criteria is outside the scope of this paper. The research team had the objective of developing a more predictive formula for system performance tuned to local conditions, but the study revealed that the data necessary to do so was not sufficient to support the effort. In particular, the research team concluded, after analyzing project data, the accuracy of the manufacturer-submitted data (particularly around the low-load operation of the heat pumps) varied widely from product to product. This portion of the study is described in more detail in the project report (BPA 2026). The resulting list of installed heat pumps is shown in Table 1.

Table 1: Description of Installed Heat Pumps

HZ		ID	Brand	Outdoor Unit	Indoor Unit	Nominal Capacity (tons)	SEER*	HSPF2*	EER2*
HZ 1	Study-Install Ducted	TAC 01	American Standard	4A6L9036A1000A	AMSTAM9A0B30V31DA	3	17	9	11.5
		TAC 02	American Standard	4A6L9036A1000A	AMSTAM9A0B30V31DA	3	17	9	11.5
		TAC 03	Trane	4TWL9036A1000A	1TEM8A0B30V31DB	3	17	9	10
		TAC 04	Trane	4TWL9036A1000A	TAM9A0C36V31DB	3	17	9	11.5
		TAC 05	Trane	4TWL9036A1	TAM9A0C36V31	3	17	9	11.5
		TAC 06	Trane	4TWL9048A1000AC	TAM9A0C42V41DAB	4	19	8.7	12
		TAC 08	Mitsubishi	SUZ-KA30NAHZ	SVZ-KP30NA	2.5	15.2	8.5	12.8
		TAC 09	Trane	4TWL9036A1	TEM8A030V3+0BA	3	17	9	10
		TAC 10	Carrier	25VNA036A00SID	FE4ANF003	3	18.1	10	11.1
		YAK 01	Lennox	MLB036S4S-2P	MMA036S4-*P	3	14.9	8.9	8.9
		YAK 61	Lennox	MLB036S4S-2P	MMA036S41P	3	14.9	8.9	8.9
	Study-Install Multizone	TAC 07	Mitsubishi	MXZ-SM36NAMHZ-U1	MSZ-FS12NA; SVZ-KP24NA	3	20.8	11.5	13.5
		TAC 21	Mitsubishi	MXZ-3C24NAH4Z		2	19	9	13.5
	Prior-Install Ducted	SNO 43	American Standard	4A6V8024A1000BD	TAM9A0B30V31	2	17.1	9.5	11.1
		SNO 44	American Standard	4A6L9060A1	TEM8A0C60V51+TDR	5	18.1	9.5	8.9
		SNO 41	Bosch	BOVA-36HDN1-M20G	BVA-36WN1-M20	3	19	9.5	12
	Prior-Install Multizone	TAC 11	Daikin	5MXS48TVJU		4	16.9	9.4	8.1
		TAC 12	Mitsubishi	MXZ-4C36NA2		3	18.2	10.5	8
		TAC 13	Carrier	38MGRQ24C--3		2	21.9	9.5	10.6
		TAC 14	Mitsubishi	MXZ-3C24NAHZ2		2	16.4	9	10
		TAC 15	Carrier	38MGRQ30D--3		2.5	21.3	9.5	9.6
		TAC 16	Carrier	38MGRQ30D--3		2.5	21.3	9.5	9.6
		SNO 11	Mitsubishi	MXZ-4C36NA2		3	18.2	10.5	8
		SNO 53	Mitsubishi	MXZ-4C36NAHZ2		3	19	10.7	11.9
HZ 2	Study-Install Ducted	INL 02	Daikin	DZ6VSA3010A*	DZ36FECC14A	2.5	17.1	8.2	9.9
		PLU 01	Mitsubishi	SUZ-KA24NAHZ	SVZ-KP24NA*	2	16	8.4	9.9
		PLU 02	Mitsubishi	SUZ-KA24NAHZ	SVZ-KP24NA*	2	16	8.4	9.9

		PLU 03	Mitsubishi	SUZ-KA24NAHZ	SVZ-KP24NA*	2	16	8.4	9.9
		PLU 04	Mitsubishi	SUZ-KA24NAHZ	SVZ-KP24NA*	2	16	8.4	9.9
		CEC 01	Lennox	SL25XPV-036-230A**	CBA38MV-036-230*+TDR	3	22	9.1	13.4
		CEC 02	Goodman	GSZS604210A*	AHVE42CP1400A*	3.5	16.7	8.1	8.5
		CEC 03	Goodman	GZV6SA2410A*	AHVE24BP1300A*	2	18	8.5	10.2
		CEC 04	Goodman	GZV6SA2410A*	AHVE24BP1300A*	2	18	8.5	10.2
		CEC 05	Goodman	GZV6SA3010A*	AHVE36CP1300A*	2.5	17.5	8.5	10
	Study-Install Multizone	CEC 31	Daikin	4MXL36WVJU*		3	20	9	11.7
		CEC 32	Daikin	4MXL36WVJU*		3	20	9	11.7
		CEC 33	Daikin	5MXS48WVJU*		4	20.6	9.3	10.5
		COL 30	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 31	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 32	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 33	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 34	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 35	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 36	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 37	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 38	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
		COL 39	Carrier	38MGHBQ24CA3		2	24.6	10	13.4
	Prior-Install Ducted	INL 41	Trane	4TWL9036A1	TAM9A0C42V41	3	17.5	8.7	11.7
		INL 42	Trane	4TWL9048A1	TAM9A0C48V41	4	19	9	11.7
		INL 44	Lennox	XP25-060-230-03	SLP00DFD0-XV60C-02	5	17.6	9.2	9.8
		INL 46	Carrier	25VNA436A00310	FE4ANB006L+UI	3	20.9	10.9	11.5
		CEC 41	Daikin	DZ17VSA42	DV42FECC14A	3.5	16.7	8.1	8.5
	Prior-Install Multizone	CEC 11	Mitsubishi	MXZ-4C36NAHZ2		3	19	10.7	11.9
		CEC 12	Mitsubishi	MXZ-3C30NAHZ3***		2.5	18	9.7	12.5
		CEC 13	Mitsubishi	MXZ-SM48NAMHZ		4	19.5	10.5	11.8
HZ 3	Study-Install Multizone	GEC 31	Fujitsu	AOU36RLXFZH		3	20.5	9.5	13
		GEC 32	Fujitsu	AOU36RLXFZH		3	20.5	9.5	13.0

**The transition from SEER/HSPF/EER to SEER2/HSPF2/EER2 overlapped with the project. SEER2, EER2 and HSPF2 converted from SEER, EER and HSPF if needed*

For each site, a home performance contractor was hired to perform an energy audit, occupant survey, and a heating and cooling load calculation. Separately, for all the study-install heat pumps, a HVAC contractor performed their own load calculation and system selection. The research team did not intervene unless the contractor had questions or the team had concerns about the sizing, which was rare. To capture power and seasonal performance and roughly understand whether indoor comfort was maintained, each site was equipped with a whole-home power meter, metering of the outdoor unit power and any known auxiliary heat circuits, and indoor and outdoor air temperature. For the central ducted systems, supply and return air temperatures were collected. Also, at the time of metering install, a measurement of supply air flow and blower motor current draw at several flow rates were captured and used to calculate air flow in the measured data. This allowed for capacity calculations for the central ducted systems, but not the multi-zone systems.

Site installations centered around eight different utilities and four Tribal Housing Authorities in the region. These are summarized in Table 2.

Table 2 Summary of Study Locations

Utility/Entity	Site ID Code	Location	Approx. Heating Design Temp	# of Installations
Tacoma Power	TAC##	Tacoma, WA area	24°F (HZ1)	17
Central Electric Cooperative	CEC##	Bend, OR area	4°F (HZ2)	12
Glacier Electric Cooperative and	GEC##	Babb, Montana	-16°F (HZ3)	2

Blackfeet Housing Authority				
Inland Power	INL##	Spokane, WA area	7°F (HZ2)	5
Yakama Power and Yakama Housing Authority	YAK##	Yakima, WA area	11°F (HZ1)	2
Okanogan Public Utility District and Colville Housing Authority	COL##	Colville, WA area	9°F (HZ2)	10
City of Plummer and Coeur d'Alene Housing Authority	PLU##	Plummer, Idaho	-1°F (HZ2)	4
Snohomish Public Utility District	SNO##	Snohomish, WA area	25°F (HZ1)	5

Energy Savings and Efficiency

The savings for each site were calculated by extracting pre- and post-install heating energy from utility bills using variable-base degree day methods (Hannas 2015). The savings were weather normalized with TMYx 2009-2023 temperatures for the nearest weather station. This approach allows site-level adjustment to a common weather basis, as well as filling gaps from missing data to ensure each site is compared on an equivalent basis.

The pre-heat pump energy consumption and the post-heat pump energy savings are plotted in Figure 1 and tabulated in Table 3. The majority of systems saved energy, and energy savings correlated strongly with pre-install, electric resistance energy consumption. Ten of the sites had negative or near-zero energy savings; of those, five had previously been heavy wood burning sites, meaning that an increase in metered energy was attributable to a shift in fuel usage, from wood to electricity. Two of the other systems with negative energy savings had system commissioning issues leading to excessive electric resistance heat. Another had an increase in heating season indoor comfort reported during the study.

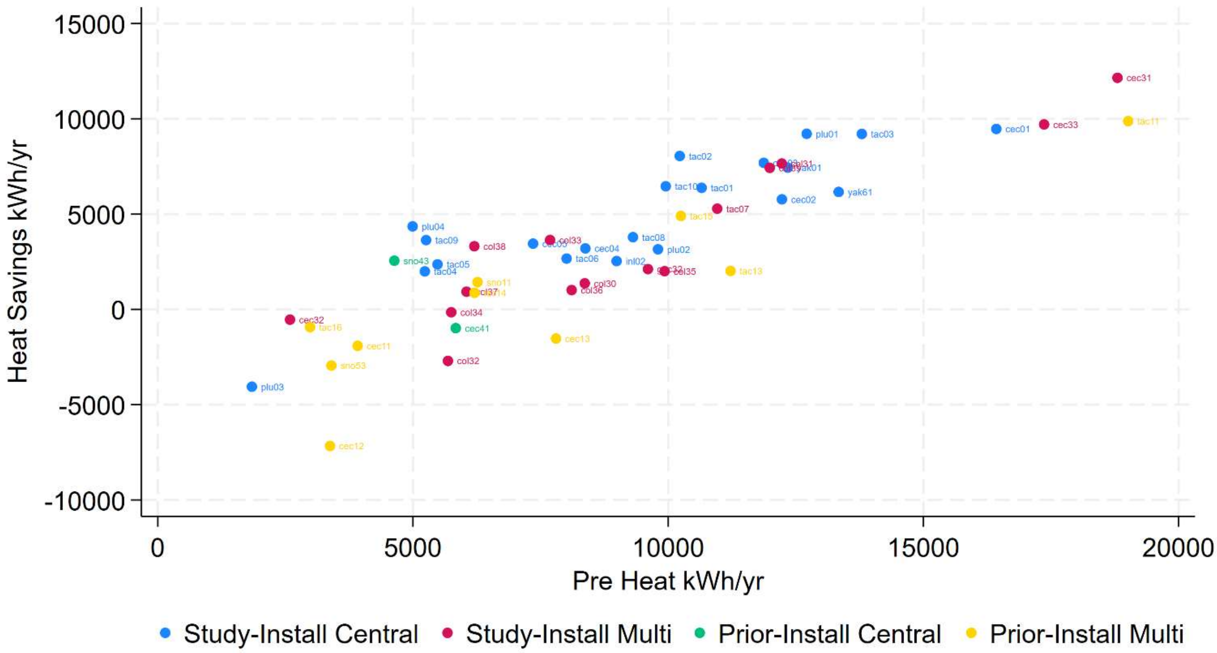

Figure 1 Weather-Normalized Heating Season Energy Savings vs. Heating Energy Consumption Before Heat Pump Install

Table 3 Heating Season Energy Savings by System Type and HZ

	Heat Savings HZ1 kWh/yr			Heat Savings HZ2 kWh/yr		
	Heat Pump Type			Heat Pump Type		
	Prior-Install HPs	Study-Install HPs	All HPs	Prior-Install HPs	Study-Install HPs	All HPs
Central Ducted						
Mean	2555	5285	5058	1719	5862	5344
Std. Dev.	.	2494	2505	.	2975	3120
n	1	11	12	1	7	8
Multi-zone Ductless						
Mean	3028	5283	3350	-1915	2981	2573
Std. Dev.	3857	.	3623	.	4385	4413
n	6	1	7	1	11	12
Total						
Mean	2960	5285	4429	-98	4101	3681
Std. Dev.	3525	2378	2988	2569	4065	4099
n	7	12	19	2	18	20

The team observed higher savings for the central ducted systems than the multi-zone systems. Since multi-zone system COPs were not measured, a comparison of efficiency is not possible. The team suspects a combination of sample bias and behavior caused this difference. The pre-heat pump estimated heating energy consumption for sites which had a multi-zone system installed was about 15% lower than for central ducted systems. Also, the central ducted systems were generally replacing a ducted electric resistance furnace (which was removed); the multi-zone systems generally went into homes which had existing zonal resistance heating, which was not always removed or properly set up to minimize operation.

For the central ducted systems, both a localized, seasonal-average efficiency and a weather-adjusted seasonal efficiency were calculated. The seasonal average COP is shown in Figure 2. This represents the efficiency as experienced in each site's climate and includes electric resistance heat. As expected, the systems in the milder HZ1 had higher average efficiency, with a median COP of approximately 2.5.

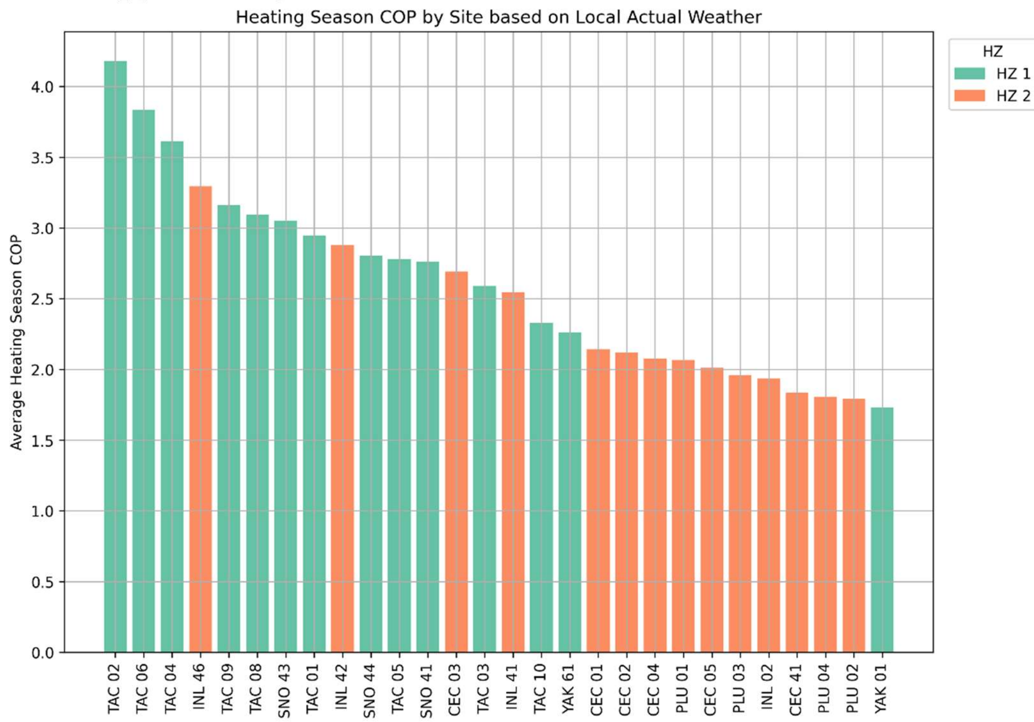

Figure 2 Heating Season Average COP for All Ducted Heat Pumps in Study

Another way to assess this seasonal efficiency is by using the bin data to calculate a field-measured HSPF2 and compare that value with the nominal HSPF2 as shown in Figure 3. To do this and compare all systems across equivalent conditions, we used the temperature bin weights (Region IV) from AHRI 210/240 for HSPF2 and applied the actual field COPs for each bin. The result is essentially a temperature-weighted COP (converted by a factor of 3.412 to match the units of HSPF2). As in the rated calculation, ER heat is included. In the milder locations in the study, not all the coldest bins in Region IV were encountered. For example, in Tacoma, there were no data in the bins from -8°F to 12°F outdoor temperature. In the following results, we re-normalized the values by dividing by the total number of hours for which data were present. This reflects only 6% of the hours so the re-normalization is minimal; if the assumption were instead to assume COP=1 (an overly-conservative assumption) for those hours, the field HSPF2 for those mild-climate sites would decrease by only about 3-5%. To reiterate, the data in Figure 3 show the seasonal efficiency of the product, with site-measured performance data re-mapped on to a standardized temperature distribution (Region IV). Site-measured efficiency is adjusted to a different range of weather to match Region IV.

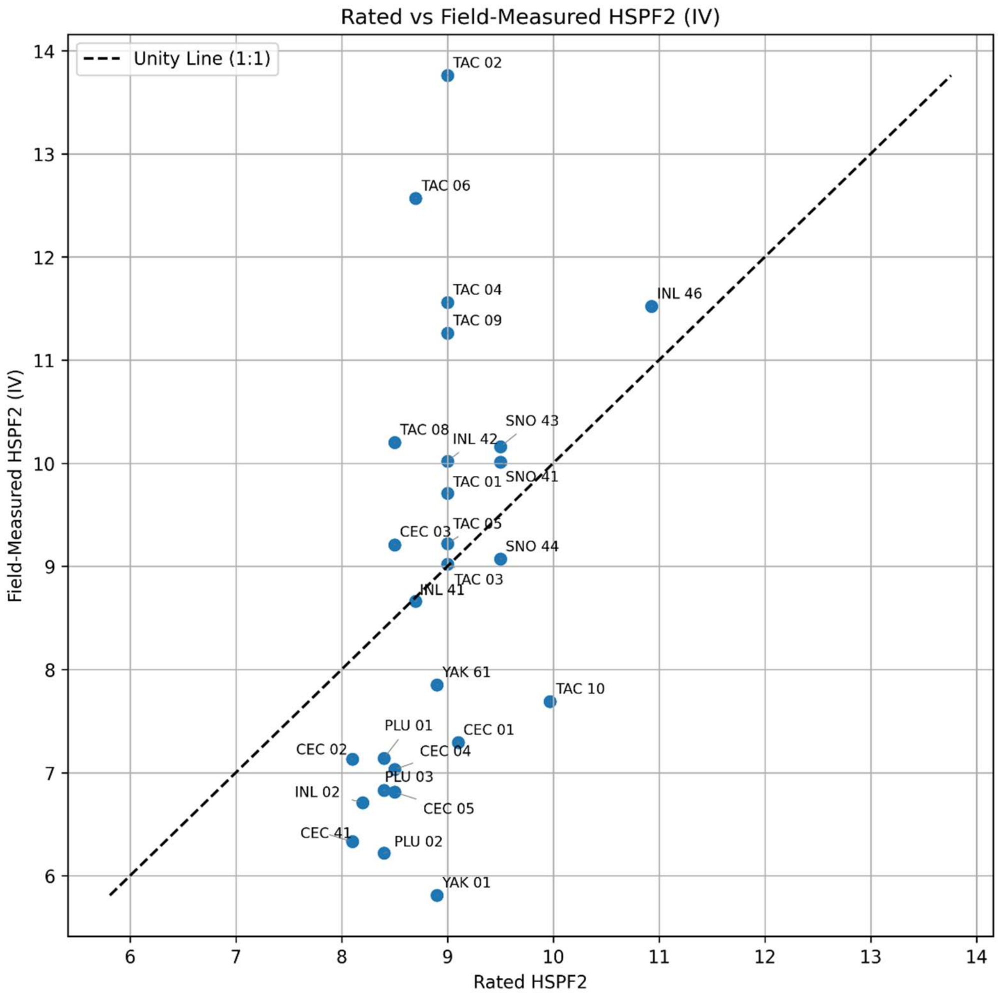

Figure 3 Field-Measured, Weather-Adjusted HSPF2 (IV) vs. Rated HSPF2 (IV) for All Ducted Heat Pumps in Study

This data shows system performance in the field varied widely, and HSPF2 was not predictive of field efficiency. This was verified with regression analysis, which showed nearly no correlation between HSPF2 and energy savings, including when normalized for square footage, weather, etc.

A key component of this study was seeking to understand if better predictive metrics could be developed by using manufacturer-submitted performance data as found in the NEEP ccASHP Product List. This was easiest addressed for the central ducted systems, which had measured capacity data. Four example sites, representative of the team's observations of the NEEP list's accuracy are shown to illustrate the range of results. The first shows a site where no backup ER is used. The second illustrates a system with moderate resistance heat usage. The third presents a site where input power data shows evidence of reaching the claimed power levels in the NEEP List, however, the capacity is consistently below expectation. The fourth shows a site with a different discrepancy: regular operation at both power and capacity levels outside of the manufacturer-reported data in the NEEP List.
For the next four figures, each dot represents one minute; the green dots are when only the heat pump was running, the purple dots are when the heat pump and the ER heat are running at the

same time, and the orange dots are when ER heat is running but the outdoor unit is not running. For simplicity, all of the following examples are in the same weather region (Tacoma, WA).

The data in Figure 4 – Figure 7 show minute-by-minute measurements of system capacity, power, and COP in 5°F outdoor temperature bins, along with the NEEP Product List maximum and minimum capacity at 47°F, 17°F, and 5°F. Under the dots is a box-and-whisker plot, for which the horizontal lines of the box represent the 25^{th}, 50^{th} (median), and 75^{th} percentile of measured power, and the whiskers extend to the 5^{th} and 95^{th} percentile. Power and capacity are shown for all minutes (including when the system was off). COP is only shown when the system is running. Four sites are shown, all in Tacoma, WA.

Figure 4 shows data for TAC 09, where backup heat was never used so all points represent "heat pump only" operation. At this site, the system was usually off (off at least 75% of the time as indicated by the entire box being at near-zero on the capacity and power graphs) for outdoor temperatures 45°F and higher; however when it did run, the data clusters within the range predicted by the NEEP data (the blue dots and dashed line). The capacity and power are both within the expected range, and COP tracks slightly below the predicted range. System capacity increased with decreasing outdoor temperature, but even in the 20-25°F temperature bin, the system was still off about 25% of the time, and when on was delivering heating capacity in the relatively narrow range of 20-25,000 BTU/h (in close agreement with the range expected from NEEP data). In the power graph, we can see similar observations. Overall, this shows a system which has sufficient capacity to meet the load across the range of temperatures encountered, operating at high efficiencies and very close to the manufacturer's claimed performance, using no electric resistance heat.

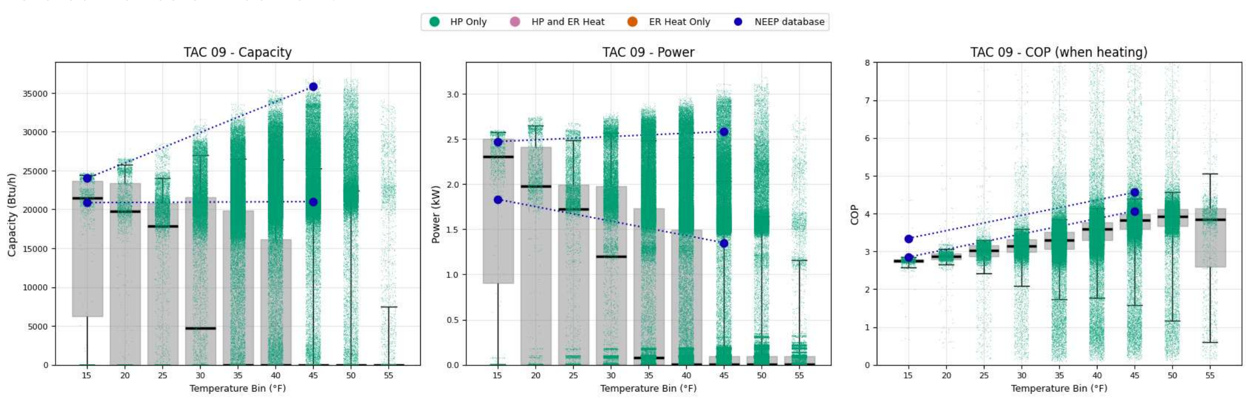

Figure 4 One-Minute Interval Capacity, Power, and COP, with Box and Whisker Plot and NEEP-Listed Values, for TAC 09

Figure 5 shows data for TAC 03. This system illustrates some moderate resistance heat usage. Detailed review at this site showed two use cases for ER heat: during defrost and in response to setback recovery. In this case, the purple dots can be seen both above and below the main cluster of delivered capacity starting first in the 40-45°F temperature bin. This often reflects defrost, where the measured capacity varies quickly in response to the compressor reversing cycle (capacity is briefly low), and through the continuation of electric resistance heating after the defrost has ended (capacity is higher than heat pump-only). We also see a small amount of electric resistance heating in the 15-20°F and 20-25°F temperature bins, although in both bins the majority of the measurements are heat pump-only heating, and there is still a meaningful amount of "off" time. For TAC 03, the system COPs again reflect fair agreement with the NEEP data, particularly in milder weather, and are slightly below the claimed performance at colder temperatures.

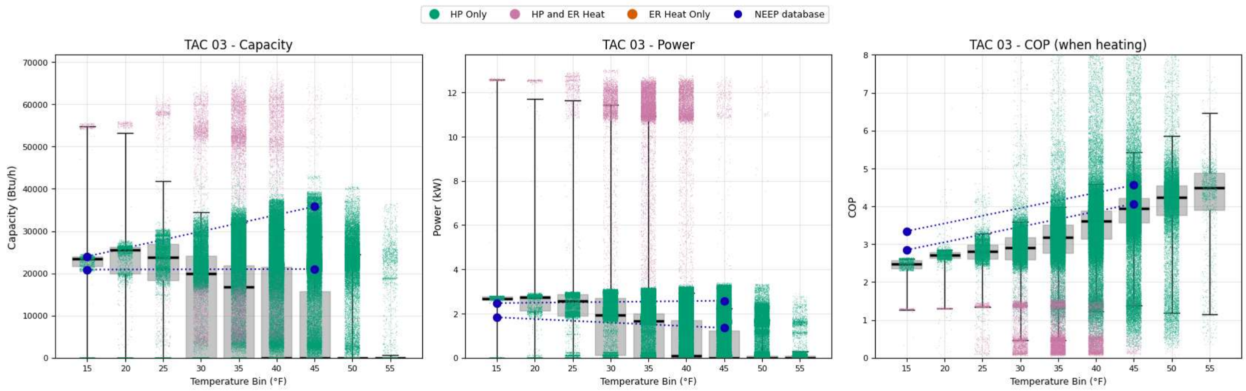

Figure 5 One-Minute Interval Capacity, Power, and COP, with Box and Whisker Plot and NEEP-Listed Values, for TAC 03

Figure 6 shows data for TAC 10. At that site, the power data shows evidence of reaching the claimed power levels in the database; however, the capacity is consistently below expectation. For example, in the range of the 25°F through 45°F temperature bins, there are regular clusters of heat pump-only power data significantly above the expected power consumption level, but the delivered capacity remains at or significantly below the expected values. In the 25°F bin, for instance, the 95th percentile power is heat pump-only power of approximately 3.5 kW, and in the same bin the 95th percentile of delivered capacity is only approximately 21,000 Btu/h, against a claimed maximum of approximately 32,000 Btu/h. As such, the COP is consistently significantly lower than claimed across all temperature bins.

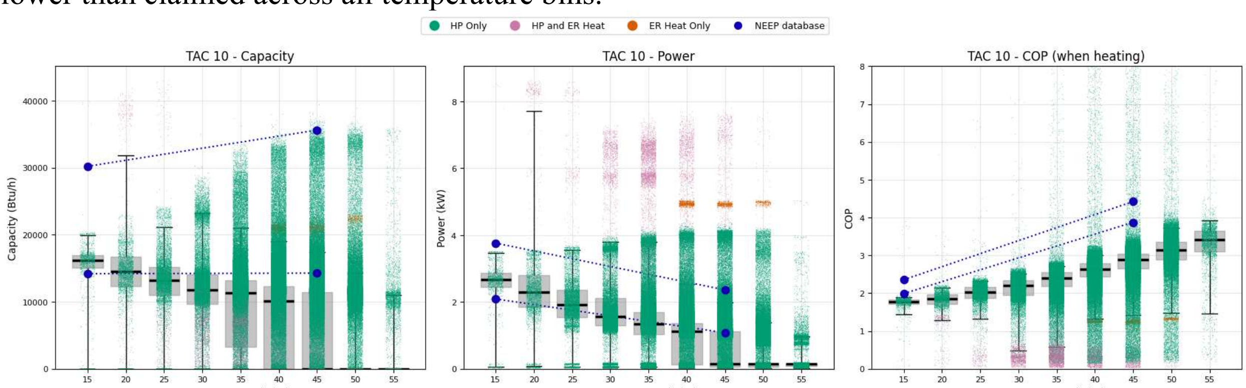

Figure 6 One-Minute Interval Capacity, Power, and COP, with Box and Whisker Plot and NEEP-Listed Values, for TAC 10

Figure 7 shows TAC 08. Where TAC 10 showed a regular under-delivery of capacity, TAC 08 has a different discrepancy. This system, which had no installed resistance heat, shows regular operation at both power and capacity levels outside of the manufacturer-reported data in the NEEP Product List. For instance, in the 15°F bin, the system regularly delivered approximately 40,000 Btu/h at approximately 5 kW of input power, despite the NEEP data suggesting a maximum power of 4 kW and approximately 30,000 Btu/h capacity. From the 35°F bin and lower, there is clear clustering of power in the range of approximately 2.5 kW, and no cluster of data near the minimum reported power level from the NEEP database. This site exemplifies a case where the manufacturer-reported data is fairly disconnected from what was observed in the field. Interestingly, the COPs were within the expected range across the span of outdoor temperatures.

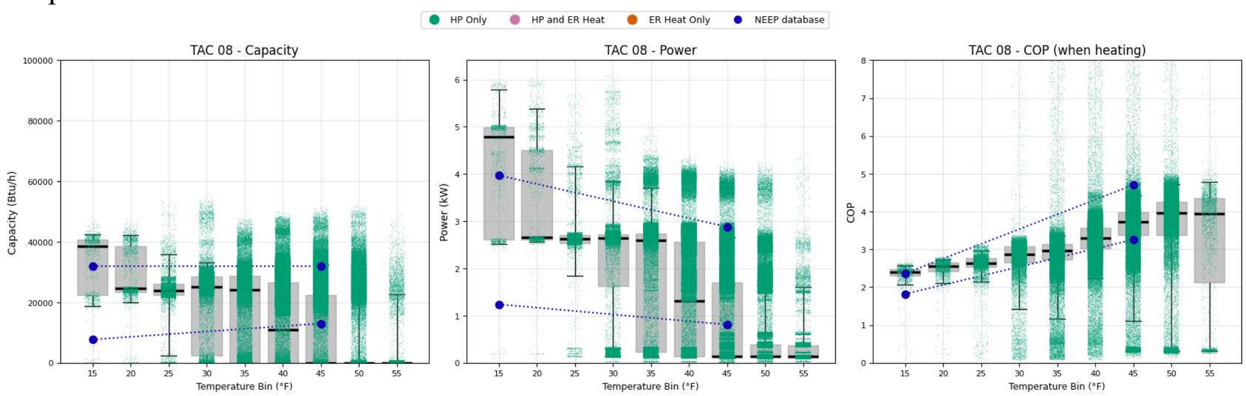

Figure 7 One-Minute Interval Capacity, Power, and COP, with Box and Whisker Plot and NEEP-Listed Values, for TAC 08

Load Shapes

The power profiles of the different system types are shown in Figure 8, and grouped by days with overnight low temperatures in the 30-45F, 15-30F, and 0-15F outdoor temperature range. There are fewer sites in the 0-15F bin. The average power profile for each type is shown with the thick black line. For all temperature bins, the average power profile of the centrally ducted systems was more peaky than for the multi-zone systems; this is reflected in the load factors (the ratio of peak-hour average power to daily average power), which were 2.06 for central ducted system and 1.29 for multi-zone systems, in the bin with 15-30°F overnight low temperatures. Central ducted systems have a single control point for the heat pump and backup heat, and any call for heat is a call to heat the entire house. The multi-zone systems on the other hand have multiple distributed heads and do not control auxiliary heat; heating with a multi-zone system is therefore inherently disaggregated. The graphs also show that a small number of sites have particularly prominent peaks (even in mild weather).

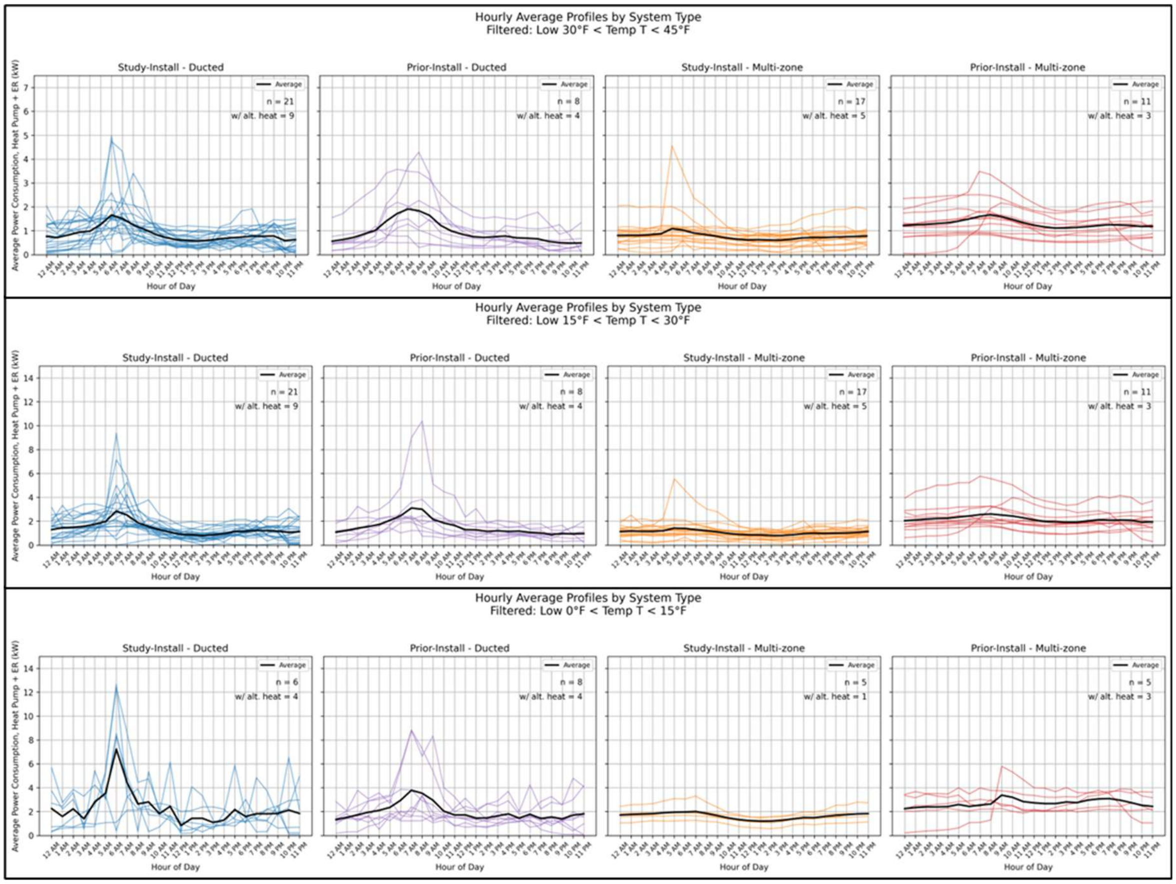

Figure 8 Average Hourly Power, by Bin of Daily Low Outdoor Temperature, for Each Installation Type. Average Power for All Systems in Category in Black

Impact of Setbacks

To more closely examine the impact of controls on power profiles, the research team recruited a total of 14 sites to participate in a controlled thermostat setback study in January and February, 2025. The participants were asked to use a thermostat setback of 3-5°F, with 8-hour duration, every day in January, and then to select a fixed temperature within the same range to maintain for all hours in February. Figure 9 shows the stacked, hourly average power for all 14 sites on the top axis and the average outdoor temperature on the lower axis; the vertical dashed line shows the date of change from "with setback" to "without setback" operation. February had more varied weather, a cold snap mid-month and mild weather at the end of the month, compared with more moderate weather for the duration of January. In February, the highest combined power was approximately 47 kW, on a day with significantly colder weather in the Yakima and Inland regions (outdoor temperatures between 0-5F), significantly colder than any day in January. However, in January the daily peak power regularly exceeded 50 kW in the second half of the month, despite outdoor temperatures seldom reaching below 15F.

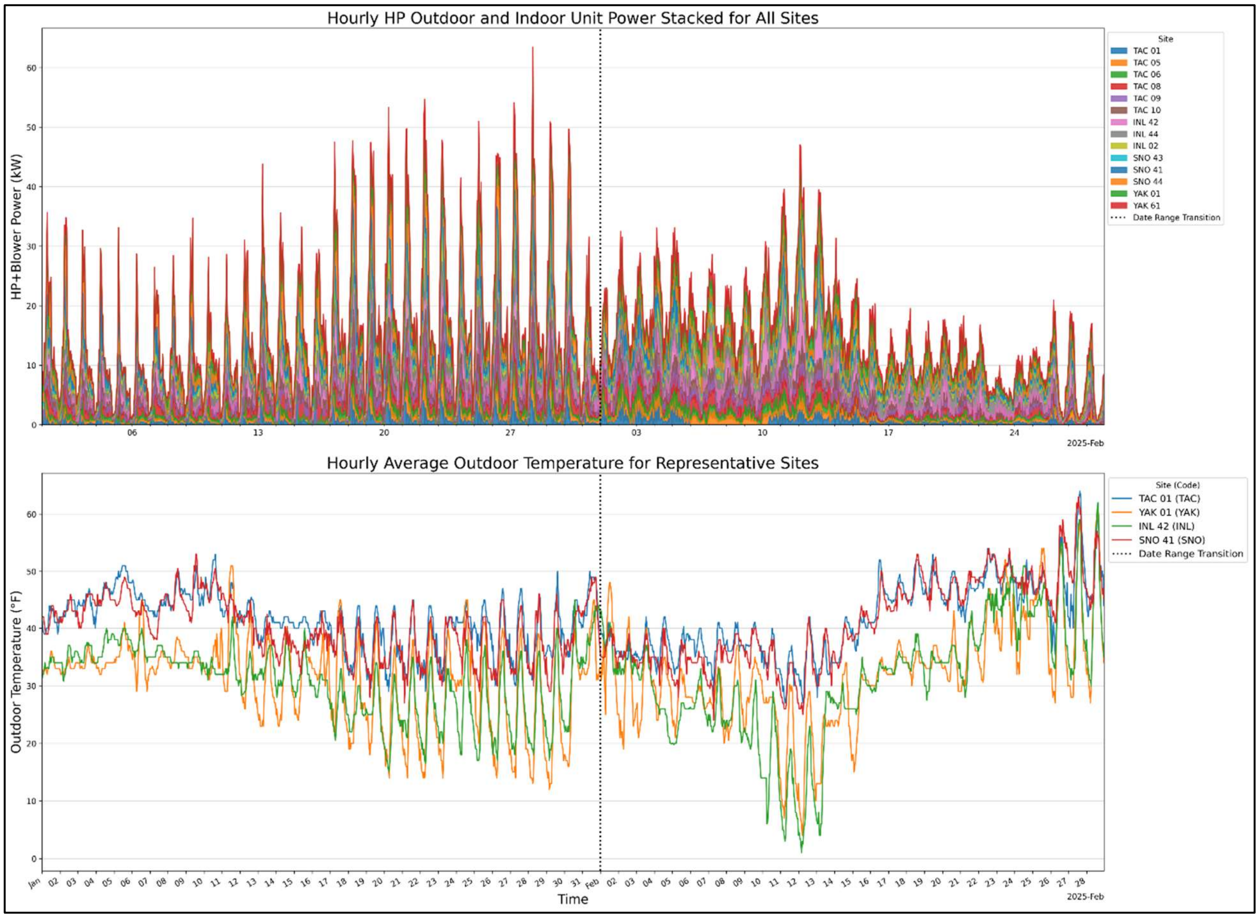

Figure 9 Stacked Average Hourly Power (Top) of Each Site Participating in Thermostat Setback Study, and Representative Weather (Bottom) at Each Location

Site-by-site inspection reveals the drastic change in power profile associated with the controls change. Figure 10 shows the hourly average power during the range of January 8-31, in blue, and February 1-28, in Orange, for each site. The shift is visible for all sites but is most pronounced at INL 02 and SNO 44, where very high electric resistance heat power causes the average hourly power during setback recovery to exceed 10 kW at each site in January, while average power is approximately 2 kW for the same hours in February. Other sites also show substantial changes to load shape, but the magnitude of the difference is much smaller. Review of per-site changes showed that all sites selected a flat temperature which was within the range of the setback temperatures selected for the setback period. TAC 08 was unique in selecting the highest temperature in their setback range as their 24-hour temperature during the no-setbacks period. TAC 08 is also noteworthy for having no electric resistance heat elements installed.

Figure 10 Average Hourly Power Profile of Each Site Participating in Thermostat Setback Study; With Setback (Blue) and Without Setback (Orange)

These findings regarding setbacks and resistance heat highlight a paradox: thermostat setbacks are intended to improve comfort and efficiency. Yet, setback recovery triggers the system's least-efficient operating state (resistance heat), in extreme cases causing power to be five times greater than needed, or more, during the recovery period. Additionally, the typical setback recovery period in winter happens to align with hours when the grid is most constrained.

Implications for Industry

The findings of this study suggest a few important implications relating to heat pump deployment and utility programs. The first is that current system efficiency metrics do not sufficiently represent field performance of variable capacity heat pumps and are not suitable for predicting the relative energy consumption of different systems. This suggests that programs should not have incentive or qualification tiers based on SEER2 or HSPF2. These metrics may be a convenient shorthand to prompt variable-speed heat pumps, but among high-rated systems, stakeholders should not expect a system with higher HSPF2 to reliably deliver better savings.

The findings also suggest an urgent need to unify and improve upon controls, particularly regarding the deployment of electric resistance heat. As this study showed, cold climate heat pumps are capable of delivering high heating capacity at low outdoor temperatures, such that they should be capable of heating without electric resistance heat during normal conditions to their design temperature. However, because of controls particularly around setback recovery, many of the heat pumps used resistance heat regularly even in mild conditions. The potential benefits of setbacks – energy savings and comfort improvements – are therefore undermined by these controls. This issue is not easily solved through controls: many variable capacity systems use proprietary integrated controllers, and the control strategies, features, and nomenclature vary from brand to brand. Examining thermostat user and contractor manuals from different brands suggests that it is not feasible today to identify a set of setting adjustments which could be applied across multiple brands to achieve a simple goal such as avoiding unnecessary resistance heat in setback recovery mode. This is an issue which cannot be solved only by contractors, homeowners or utilities, but will require manufacturer agreement and support.

Conclusions and Recommendations for Future Work

This study demonstrated performance of 57 heat pumps in the Pacific Northwest region. The results showed average energy savings in the range of 30% for multi-zone systems and 50% for central ducted systems, compared with the prior existing electric resistance. The research team suspected a combination of site sample factors and zonal resistance heaters remaining in place contributed to lower savings for the multi-zone systems. The multi-zone systems contributed significantly less to morning peak power draw than the central ducted systems on average. Thermostat setbacks were identified as a significant issue causing high electric resistance usage during winter morning hours; a sample of 14 sites in a controlled test showed a near-doubling of the average power consumption across the sites at 6AM on average. The data also showed the sites using about the same amount of energy, or more, with a setback than without a setback. Electric resistance heat triggered by setback recovery drove these issues.

The study highlights some additional research and industry coordination opportunities. In particular:

- A further in-depth examination of variable speed heat pump controls, to understand cause and effect of different settings as they exist across different brands. This may make it possible to unify recommendations to allow consistent guidance for energy savings, peak demand reduction, and comfort across brands.
- To help address winter peak load concerns, utilities and manufacturers should coordinate around thermostat controls. The current status quo is that thermostats overwhelmingly prioritize comfort, often unnecessarily at the expense of efficiency, and proprietary controls are distinct; if utilities wish to be able to address heat pump behavior during peak events, it would be necessary to make the desired outcome explicit and seek consensus with manufacturers on how to achieve those goals.
- This study included 57 systems but did not include a baseline sample of standard-efficiency, typical-practice heat pumps. A data set containing detailed performance of a tranche of "market baseline" systems would provide value in further quantifying the performance differences with high performance systems. Similarly, comparison with mid-market variable speed systems (non-cold climate) can help understand the major drivers of energy and demand savings.
- The results of this study suggest that manufacturer-provided data found in the NEEP Product List, was reasonably reliable for maximum capacity data (as might be used for

sizing) but not sufficiently reliable for low-load performance to be useful in energy modeling or other analysis. Improving this data could have substantial benefits for the industry, in better quantifying the expected benefits of heat pumps and identifying metrics to accurately predict performance for a given set of conditions.

- The results of this study suggested better energy savings and worse winter peak load contribution for central ducted systems compared to multi-zone systems. In this study, site selection differences may confound the true differences between the system types. Further investigation via larger and more representative samples of homes could help more clearly quantify the performance differences between these system configurations, and identify any causes and possible mitigation strategies.

References

Northwest Power and Conservation Council. *Pacific Northwest Electric Power Planning and Conservation Act (Northwest Power Act)* (PDF). (1980). Available: https://www.nwcouncil.org/sites/default/files/poweract.pdf [nwcouncil.org]

Baylon, David et al. (Ecotope for BPA/NW Power Council). "It Sounded Good on Paper: Field Performance of Residential Heat Pumps in the Pacific Northwest." (Report for BPA/NPCC).

RDH Building Science. BC Cold Climate Heat Pump Field Study. 2020. Available: https://www.rdh.com/app/uploads/2023/12/BC-Cold-Climate-Heat-Pump-Study-Final-Report.pdf

Winkler, J & Ramaraj, S Field Validation of Air-Source Heat Pumps for Cold Climates. 2023, https://doi.org/10.2172/1973090.

Kelsven, Phillip et al. "Air Source Heat Pump Commissioning, Controls & Sizing Baseline Field Study." (BPA Momentum Savings program, published 2019).

Harley, Bruce et al. "Low-Load Efficient Heat Pump Investigation: 2020 – 2025 Summary Report." (NEEA Report #E25-349).

Bonneville Power Administration. *High Performane High Capacity Heat Pump Field Study Report*. PDF (2026). Available: https://www.bpa.gov/-/media/Aep/energy-efficiency/emerging-technologies/2026-hphchp-finalreport_web.pdf

Hannas, Benjamin and Michael Logsdon, Comparison of Bayesian Billing Analysis to Pooled Fixed Effects and Variable-Base Degree-Day. 2015 2015 International Energy Program Evaluation Conference, Long Beach. https://www.iepec.org/wp-content/uploads/2015/papers/017.pdf

Disclosure template

AI-assisted tools used: Google Gemini
Used for: Copy editing
Extent: Suggested edits for readability and flow of arguments.
Human review: The authors reviewed all suggestions and manually incorporated the desired edits.