Mapping the Way to Higher Performing Heat Pumps

Neal Kruis and Elyse Malherek, Big Ladder Software
Christopher Dymond and Aaron Ingle, Northwest Energy Efficiency Alliance

ABSTRACT

Heat pumps often underdeliver on efficiency because of improper specifications, installation practices, and controls, which are sensitive to local climate. New technologies such as advanced vapor injection, onboard diagnostics, and low speed optimization of fans and compressors make heat pump selection more nuanced, with conventional laboratory test-based performance ratings not necessarily providing enough information on how a heat pump will perform in a home. This study quantifies efficiency opportunities from specific improvements to heat pump installation, controls, and performance maps that are undervalued in the current market.

The study developed heat pump performance maps from field performance monitoring data from BPA, NREL, NEEA and Efficiency Maine, combined with expanded heat pump performance data. These performance maps were then used in the California Simulation Engine building energy modeling software to generate relative performance differences and explore which heat pump attributes are most impactful across diverse applications.

This study had two main objectives: 1) provide field-calibrated energy savings estimates resulting from heat pumps with improved attributes. Savings rates could be used within utility programs to justify system requirements or contractor education. 2) Inform and support manufacturers' pursuit of improvements not revealed under industry standard lab testing and rating procedures. The authors present how results are being used to accelerate market transformation through increasing market knowledge about which systems optimize energy savings and customer benefits for various applications.

Introduction

This study aims to bridge the gap between published heat pump performance and the performance experienced in the field. Traditional models assume equipment operate as published by the manufacturer; however, that is rarely the case as numerous field testing studies have indicated (Winkler and Ramaraj 2023; Efficiency Maine 2026; BPA 2026). The study expands the modeling scope to apply observed heat pump performance trends across a variety of scenarios and locations to determine anticipated energy savings. Representations of the best- and worst-case observed performance provide bounds for the expected range of savings.

The workflow included creating new and existing home definitions for energy models covering ASHRAE climate zones 4-6 with typical loads, exploring permutations on HVAC configurations like ducted and ductless systems with heat pump sizing alternatives, as shown in Figure 1.

Four heat pump cohorts were analyzed: 1) unimproved; 2) cold climate capable (CCC); 3) low load efficient (LLE); and 4) high performance high capacity (HPHC), which combines the characteristics of CCC and LLE. Performance maps referred to as "idealized" were created to represent manufacturer published data for each cohort. This study utilized an innovative approach to use observed field and laboratory data to modify idealized performance maps to create "*observed top*" and "*observed bottom*" performance maps that reflect the top quartile and bottom quartile of field performance for each cohort. The performance of these four cohorts, each with three performance map variations, were then simulated across climate and sizing permutations to determine relative savings of heat pumps.

The heat pumps modeled for this work have Region IV HSPF2 values in the range of 8.5 to 9 which are most likely to be impacted by utility incentives and market transformation programs.

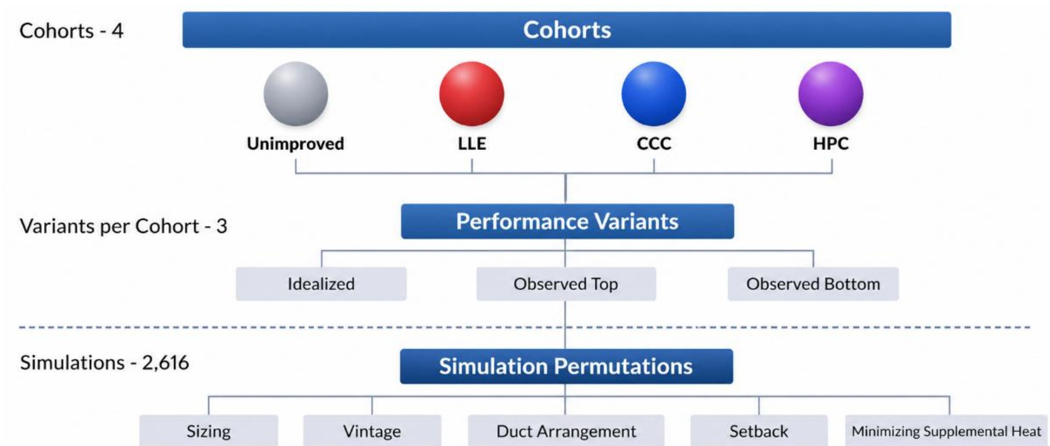

Figure 1. Study workflow diagram and modeled heat pump variations.

Modeling Residential Heat Pumps

Residential energy modeling has expanded, fueled by HERS ratings, code requirements, and RESNET-supported code compliance; however, accurately modeling heat pump performance remains a challenge and an abundance of evidence exists that field performance consistently underdelivers compared to published ratings. As heat pump adoption is growing rapidly, improved modeling will be necessary for market actors investing in this technology. Proper heat pump modeling requires special attention to heat pump performance, defrost energy, supplemental heat usage, and distribution losses.

Heat pump power and capacity performance is sensitive to operational conditions, most strongly correlated to the temperature across the evaporator (during heating) and compressor speed. While polynomial curve fits are common, performance maps were utilized in this project to allow a more nuanced characterization of performance with multivariable interpolations to more accurately reflect the heat pump operation across the entire operational range of the unit.

An additional consideration for heat pumps is operation outside of the mapped conditions and the degradation associated with compressor cycling. Single-speed machines are optimized for cyclic operation and have a cycling degradation coefficient of approximately 0.08 while variable capacity heat pumps are designed to reduce compressor speed to meet load down to a minimum. Below that, the cycling degradation coefficient is approximately 0.4.

Heat pump defrost energy, as seen in the field, can be consequential. There is, however, limited understanding of the interplay between local humidity and frost formation, so energy modeling assumptions rely on generalized mathematical approximations that simply can't reflect the complexity or usage seen in the field.

Additionally, most models underestimate the supplementary heat energy use because they fail to account for the thermal comfort impact of reverse-cycle defrost. Manufacturers often use supplementary heat to maintain thermal comfort during and after these cycles while models simplify inputs to include the type of defrost (reverse cycle or resistive), timed or on-demand control, and cut-in/cut-out temperatures. Besides defrost thermal comfort, the most impactful supplemental heat consumption is with large instantaneous loads, such as during thermostat setback recovery. (Setback recovery results are omitted from this paper for brevity, but are available in the final report.)

For ducted systems, the distribution system characteristics should be reviewed as they can account for 30% or more heat loss throughout the heating season (ASHRAE 2023). The losses are particularly sensitive to duct location. If the location is not within the conditioned space, the duct length, size, insulation, and leakage become driving factors for distribution effectiveness. This is especially true for variable capacity systems as was identified in a UC Davis study (Krishnamoorthy 2018).

Performance Maps

Performance maps in this study define the heat pump's capacity and power at combinations of multiple outdoor temperatures and compressor speeds. This study is based on three performance variants per cohort, for a total of 12 maps for each ducted and ductless unit. Table 1 provides brief descriptions of the three performance variants.

All performance maps are ultimately derived from the idealized unimproved heat pump map, which is defined as an 8.5 HSPF2, variable capacity heat pump. The capacity and power at each combination of outdoor temperature and compressor speed (min/max) are foundational map inputs and are derived from statistical analysis of real units in the NEEP Cold Climate Heat Pump List that do not meet the criteria for either LLE or CCC (see Table 2). The ResDX python package (Big Ladder 2026) was utilized to dynamically adjust the performance map to match the targeted HSPF2 of 8.5 without changing the overall shape of the map defined by the manufacturer reported data obtained from the 2024 NEEP cold climate heat pump list (2026).

Table 2 indicates the minimum criteria for the three improved cohorts. Any point not modified is denoted with a black fill and assumed equivalent to the unimproved map value. These minimum thresholds were applied to the idealized unimproved map to create the idealized improved cohort maps (LLE, CCC, and HPHC). Although the improved cohorts may express additional variations, only the thresholds of the improvement were applied to isolate the effects of the minimum definition.

Table 1. Performance map descriptions for this study.

Performance Map	Description
Idealized	Aggregated performance map derived from manufacturer published data and normalized to meet criteria near the federal tax credit threshold 8.5 HSPF2.
Observed Top	Performance map calibrated to observed lab and field data. Statistical multipliers weighted to the 75th percentile of deviations from manufacturer published within discrete temperature bins were applied to the idealized map. Represents the top quartile of observed heat pumps.
Observed Bottom	Performance map calibrated to observed lab and field data. Statistical multipliers weighted to the 25th percentile of deviations from manufacturer published within discrete temperature bins were applied to the idealized map. Represents the bottom quartile of observed heat pumps.

Table 2. Cohort minimum definition improvements for idealized performance maps.

		COP						Capacity Ratio					
		Min Speed			Max Speed			Min Speed			Max Speed		
Cohort	Abbreviation	5°F	17°F	47°F	5°F	17°F	47°F	5°F	17°F	47°F	5°F	17°F	47°F
Unimproved	Unimproved												
Low Load Efficiency	LLE		2.55	4.5									
Cold Climate Capable	CCC										0.8	0.91	
High Performance, High Capacity	HPHC		2.55	4.5							0.8	0.91	

Performance Map Calibration

The concept of creating a calibrated map presented a unique challenge. Calibration data was distilled from observed field or third-party laboratory test data. A limited amount of studies had sufficient measured data, and those studies featured a range of seasonal efficiencies and uneven cohort sample sizes. To normalize and calibrate the results from these disparate sources, the heat pump performance was expressed as a percent deviation from the published performance spec for each map data point. That relative difference was applicable regardless of the heat pump seasonal performance HSPF2. Cohort-representative deviations were applied to the idealized map to represent the calibrated field performance, representing two variations: top and bottom observed heat pumps.

Five field studies were found to have significant enough raw measured data to derive useful performance deviations from the published performance. Overall, the studies covered 151 installations, with 114 having sufficient heating season data without sensor malfunctions. The installations were then categorized as exhibiting variable capacity behavior or cyclic behavior. Figure 2 shows a bubble chart visualization utilized in this study to determine behavior. Short duration in a linear line indicates cyclic behavior. A “hockey stick” shape and longer duration operation (above minimum capacity) indicates variable capacity behavior.

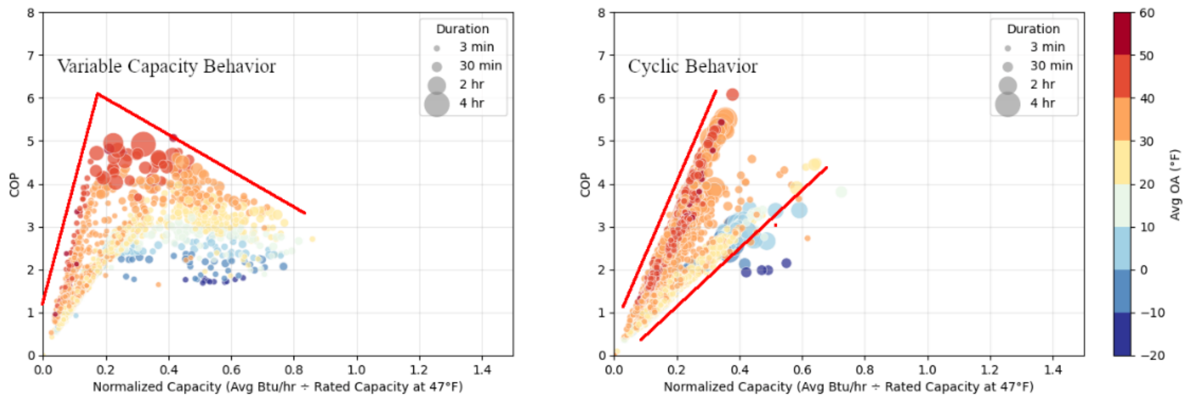

Figure 2. Sample bubble charts exhibiting variable capacity and cyclic behavior.

Figure 3 presents the categorization of the 151 units. Almost 1/3 of units demonstrated cyclic behavior. The amount of units within each cohort varied widely, but studies focused more on CCC and HPHC heat pumps. Note that this figure should be viewed as only representative of the units referenced in this study rather than a generalized industry benchmark.

Figure 3. Categorization of units referenced in this study by operation and cohort.

Only one study had laboratory test data (Wilcox and Conant 2018) that provided complete datasets to compare the performance data. It replicated the test conditions seen in the AHRI database or the NEEP ccASHP product list. A simple relative deviation from the advertised values was calculated.

For the remaing sources that were conducted as field studies, the determination of minimum and maximum capacity and the corresponding effiency was more of an art and often provided incomplete datasets. From the bubble charts, the data was grouped into 5°F bins. The minimum and maximum capacity and corresponding efficiency were determined as shown in Figure 4. Some installations had limited operational data in low ambient temperature bins, particularly relating to those at 5°F or less. In addition, field studies rarely capture the heat pump maximum rated capacity at 47°F (and higher) ambient temperature bins because the instantaneous building load seldom reaches the system's output ceiling under mild weather conditions.

The values were checked against statistical distribution of capacity and power by percentile to validate the data by temperature bin.

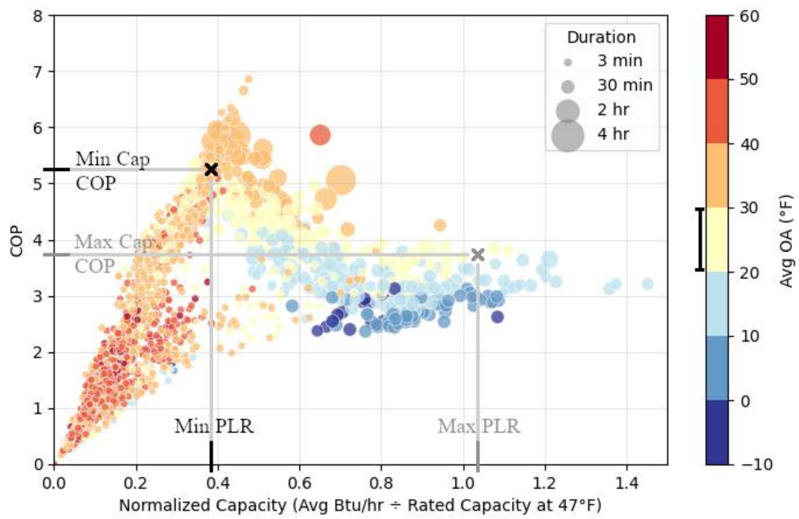

Figure 4. Example of bubble chart to show derivation of minimum and maximum capacity (part load ratio, PLR) and corresponding efficiency for a temperature bin between 20°F and 30°F. Note, the actual study used 5°F bins.

The deviations from temperature bins for minimum and maximum capacity and corresponding efficiency were aggregated for field and laboratory studies. For each aggregated temperature bin, the data was organized by percentiles; the 75th percentile was chosen to represent the *observed top* units and 25th percentile to represent the *observed bottom* units.

Sparse data points for conditions not often found in the field, such as the maximum capacity and corresponding efficiency deviations for mild conditions, were guided exclusively by laboratory data sources. For cohorts without sufficient sample size, heat pump aggregated results for individual, applicable conditions were applied. For example, LLE only units utilized the minimum capacity and efficiency at 47°F for LLE and HPHC units. For lower temperature bins, LLE only units utilized maximum capacity and efficiency at 5°F for LLE and unimproved units. Additionally, ResDX requires operational values for the compressor minimum temperature. For extreme low temperature bins with insufficient data (e.g. -10°F and -20°F), the deviations from the lowest temperature bin from the aggregrated values were applied.

Figure 5 displays the operational maps representing the idealized, *observed top*, and *observed bottom* by cohort and for capacity and efficiency. For capacity maintenance relating to the maximum capacity at lower ambient temperatures, the *observed top* were capable of meeting the manufacturer published values at 5°F and 17°F. CCC showed a significant drop in meeting the maximum capacity at the minimum temperature; however, there weren't significant data to provide high confidence at -20°F. The weather of the modeled locations does not operate at the compressor minimum temperature, so the overall simulation results will not be significantly affected. The *observed bottom* units, particularly for CCC and HPHC, showed a capacity ratio of 0.55 compared to the CCC idealized target 5°F capacity ratio of 0.80, implying the bottom range of CCC and HPHC systems exhibit comparable capacity maintenance to the bottom range of unimproved systems.

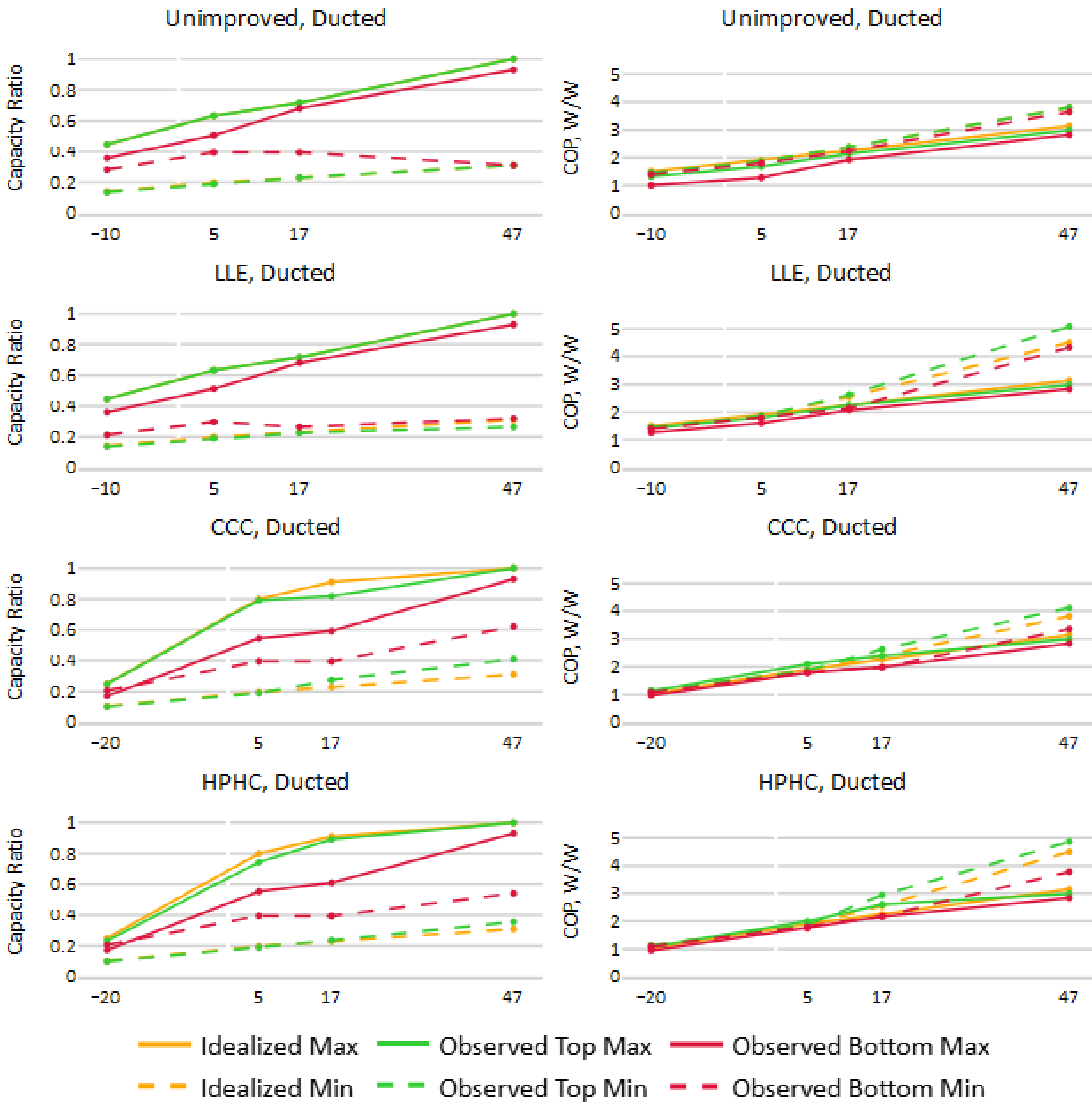

Figure 5. Performance map capacity and efficiency variations by ambient temperature for each idealized cohort.

Variable capacity heat pumps have an ability to turn down the compressor and fan speeds to achieve higher efficiency. *Observed top* units often closely matched the expected turndown, while *observed bottom* units, especially CCC and HPHC units, consistently showed a limited range of turndown relative to the published data.

In terms of efficiency, the maximum capacity COP, which is most relevant at 5°F and 17°F, for the *observed top* units performed as expected, even exceeding the published data for CCC and HPHC units. The unimproved *observed top* showed a 12% penalty at 5°F, while the unimproved *observed bottom* units showed a particularly steep decline in efficiency, a 33% penalty at 5°F. The improved cohorts maintained efficiency quite well across the observed range.

For minimum capacity COP, which is most relevant at 47°F, *observed top* units from all cohorrts at least matched the published performance, while the top LLE exceeded that efficiency by 13%. LLE oberved bottom units were still within 4% of published. *Observed bottom* units for CCC and HPHC; however, showed performance 22% and 26% lower than published, respectively.

A seasonal rating of Region IV HSPF2 was calculated based on AHRI procedures for each performance map. Figure 6 shows HSPF2 values of the idealized maps compared to the top and bottom observed maps calibrated from the field and lab study data. Of the four cohorts, the unimproved top and bottom values are both less than the idealized HSPF2 values, whereas the LLE showed the greatest consistency, and notably, the CCC and HPHC bottom values were substantially lower than the idealized values.

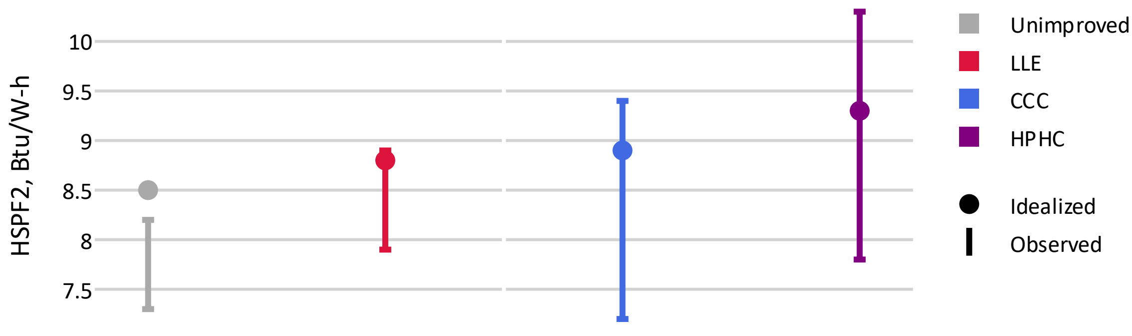

Figure 6. Rated HSPF2 values and those calculated from the performance maps after applying observed variations.

Building Assumptions

This study focused on new and existing homes around 2,000 square feet. The homes are representative of single-story, slab-on-grade ranch style. Attics and wood framed walls (24 inches on center) were modeled. The characteristics of the single-family new and existing construction homes were informed by the residential building stock assessment (NEEA 2024). For new construction homes, the location-specific energy code current to 2026 was referenced. See Table 3 for a summary of residential building characteristics.

Building loads were defined independently for lighting, people, TV, set top box, computers, refrigerator, dishwasher, cooking appliances, clothes washer, dryer, and other miscellaneous electrical loads, with default values and schedules based on Title 24 CBECC (CEC 2025).

Table 3. Building characteristics used within the study.

Building	Climate Zone	Roof U-Value	Wall U-Value	Window U-Value	Infiltration, ACH50
New Home, Seattle, WA	4C	0.024	0.056	0.3	4
New Home, Spokane, WA	5B	0.024	0.056	0.3	4
New Home, Boise, ID	5B	0.026	0.060	0.3	3
New Home, Bozeman, MT	6	0.024	0.045	0.3	4
Existing Home (NEEA 2024)	All	0.08	0.12	0.49	8.8

HVAC sizing was based on a building, vintage, and location combination to meet the design cooling load for the location. This sizing represents common practice in the field in the Northwest. Exact sizing was not essential as each system configuration would be modeled with a sizing variation of a ton larger and a ton smaller. Supplemental heat was sized to meet the heating load under design conditions. Table 4 indicates the initial energy model heat pump capacity.

Table 4. Heat pump initial sizing, in cooling tons, as determined from the simulation engine. Sizing variations of +/- 1 ton were modeling permutations.

Location	New Home	Existing Home
Seattle	2	3
Spokane	2	3.5
Boise	3	4.5
Bozeman	2	4

The California Simulation Engine (CSE) was utilized for all modeling (The CSE Authors, 2026). CSE is a rigorous and up-to-date simulation engine that provides detailed representation of air distribution losses and nuanced temperature-dependent heat pump performance maps. These capabilities are more advanced than those available in other common simulation engines.

Results

Figure 7 documents the primary findings of the study. Each improvement cohort and idealized and observed variations were simulated in a residential home in Seattle, Boise, Spokane, and Bozeman. New homes were simulated as ductless and with ducted, while existing homes were simulated as ductless, with legacy ducts, and with replaced ducts.

The results reveal that real-world heat pump performance consistently differs from published promises, often falling short of theoretical expectations.

LLE heat pumps are the most reliable for savings in real installations and the preferred heat pump cohort to achieve savings, particularly in Seattle and Boise. For Spokane and Bozeman, the potential upside for adding CCC shows additional savings; however, both cities still have significant hours within the low load region suggesting that HPHC units should be preferred to maximize savings while minimizing the risk for bottom performers. CCC machines had the largest variance in performance, with the *observed bottom* representation showing no characteristics of a cold climate heat pump and potentially producing less savings than the unimproved cohort.

Figure 8 provides a comparison of modeled savings as a percent of annual heating consumption for the idealized and *observed top* and *bottom* simulations. Only charts for new construction were included as existing homes exhibited similar results. The savings are relative to the average performance of the unimproved *observed top* and *bottom* performers. Note that the unimproved cohort with idealized performance is indicated by the gray dot. This adds some perspective to the savings that could be achieved if heat pump equipment performed as-advertised which averages around 5% in Seattle, but up to 10% in Bozeman.

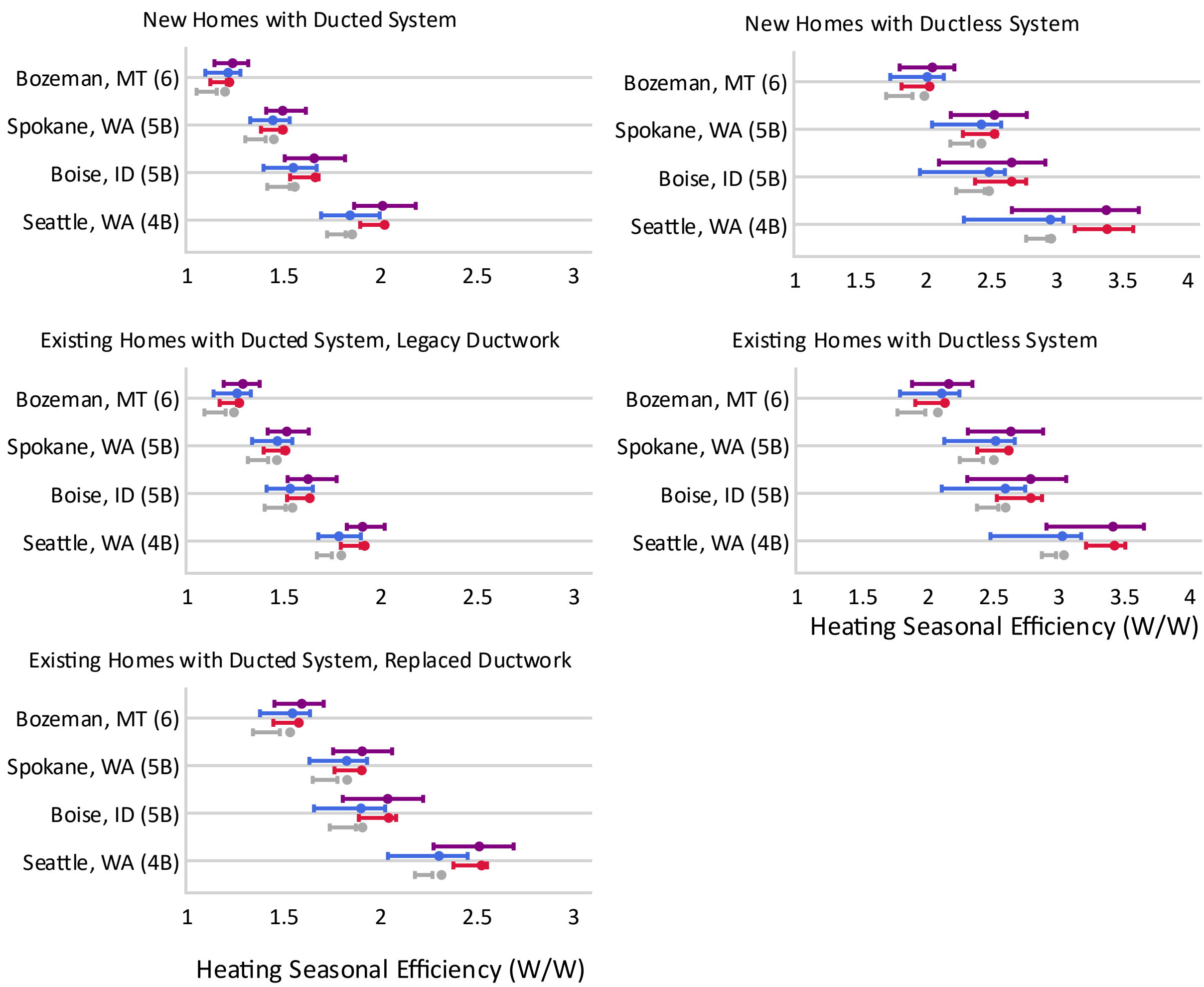

Figure 7. Heating seasonal efficiency (including distribution losses) for idealized and observed cohorts by location for new homes and existing homes by duct type.

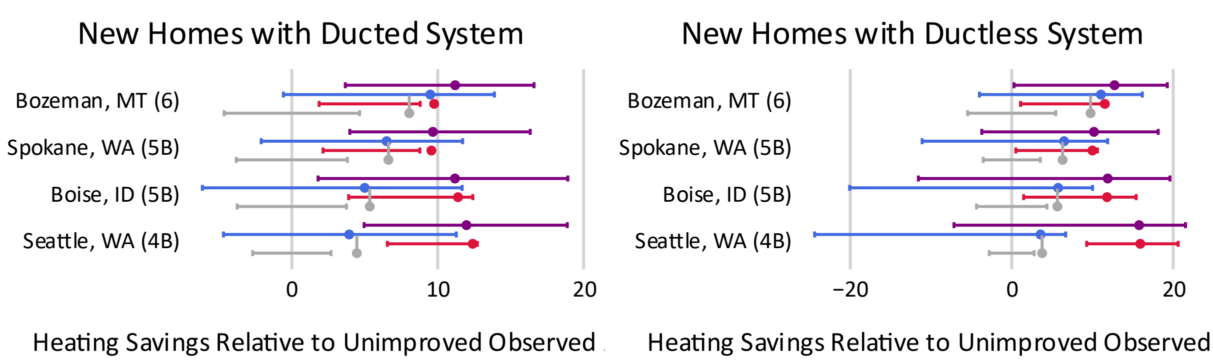

Figure 8. Relative % savings of idealized and observed cohorts by location for new homes by duct type relative to average unimproved *observed top* and *bottom*.

Improvement Savings and Observed Variation

With few exceptions, the expected savings for a particular improvement is outweighed by the observed variations seen in the field. Table 5 illustrates that the modeled savings for an individual improved cohort is significantly less than the overall range of observed performance between *observed top* and *bottom* performance within that cohort. That gap increases with colder climate, ranging from a factor 2.5 times larger variation in improvement to maximum savings in Seattle to a factor of almost 7 in Bozeman.

Table 5. Improvement energy savings versus cohort observed variations for new homes with ductless systems.

Variation	Seattle	Boise	Spokane	Bozeman
Idealized Strategy Savings, kWh	295	229	178	202
Observed Variation in Heating Usage Within a Cohort, kWh	755	1146	1049	1372

Cold Climate Capable Savings

There are often high savings expectations for CCC heat pumps, particularly for installations in cool or cold climate zones (ASHRAE climate zones 5 and 6). Many homes within those climate zones experience limited hours in conditions that result in increased savings for CCC versus the hours that LLE provides savings as shown in Figure 9.

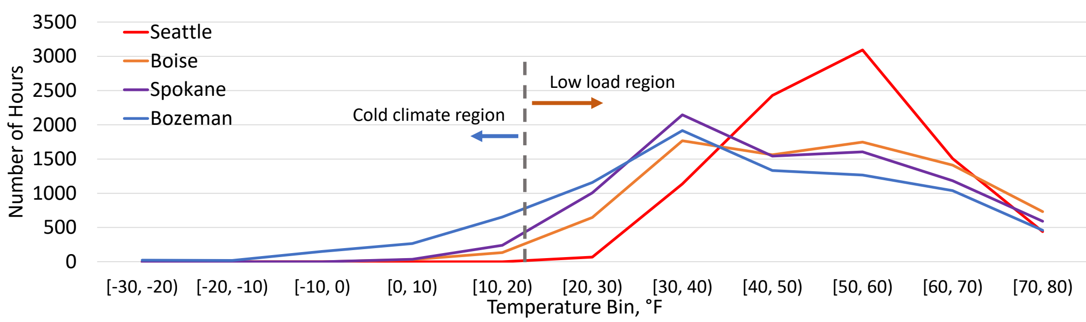

Figure 9. Weather analysis of the modeled locations.

The CCC heat pump provides savings relative to the capacity differential from the unimproved cohort. **Error! Reference source not found.**Figure 10 shows the hourly modeled heat pump supplied thermal load during the simulation, overlayed with the idealized maximum capacity lines for the CCC and unimproved cohorts. The operational points between the two capacity lines indicate the primary opportunity for CCC savings, though prorated savings occur above the CCC savings line. Note that ductless units were sized according to the size calculated for ducted units, as is typical in practice. That resulted in ductless being oversized and showed significantly fewer operational hours in the CCC savings zone. Only Bozeman and ducted units in Spokane showed significant hours of potential CCC savings. Note that the CCC unit modeled in this study represents units that don't utilize enhanced vapor inject (EVI) technology, and this study did not analyze peak impacts, which can be significant regardless of frequency.

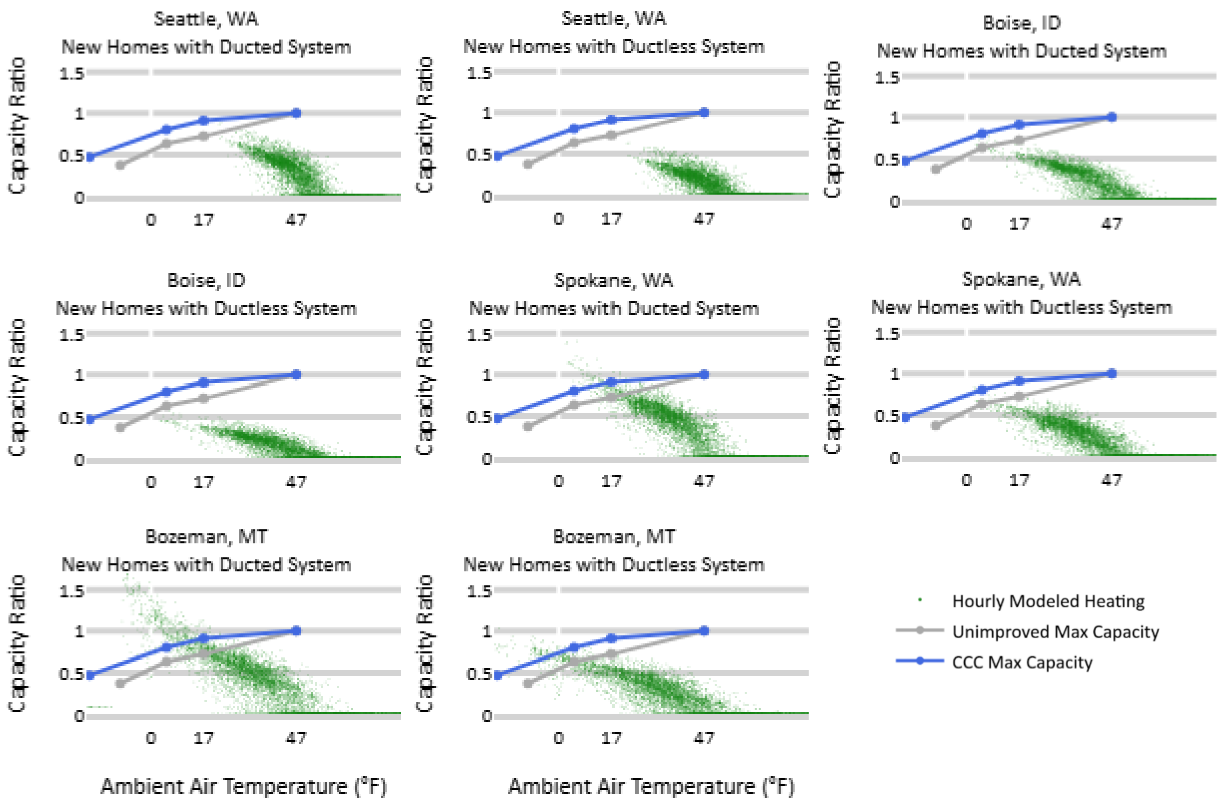

Figure 10. Hourly load with CCC and unimproved maximum capacity overlayed to show the CCC savings zone.

Sizing

Sizing plays an important role in the performance of a heat pump unit and, for the purposes of the model, modifies the operating position on the performance map. Sizing permutations explored under- and over-sizing by one cooling ton.

Figure 11 demonstrates the impact of undersizing, which leads to excessive supplementary electric resistance. LLE units were more impacted, with an average 24% increase in heating electric usage. This is primarily because undersizing shifts low load operating hours into higher relative loads, moving away from the high efficiency, low load range. Under sizing for unimproved and CCC was about a 20% penalty. LLE was also most impacted by oversizing the unit by a ton. It was hypothesized that by oversizing, the shift to even lower part load range might do more harm than good due to cycling losses for loads below the minimum capacity. The simulated results indicate the oversized LLE produces an additional 5% savings, higher than the added 3% savings for unimproved and HPHC.

Seattle showed the largest relative savings for LLE upsizing because of the large distribution of load hours at moderate temperatures. Bozeman, the location with the most hours of low ambient temperature showed the greatest benefit for CCC and HPHC oversizing. Boise showed the most muted modeling response, likely due to having a larger initial capacity unit and a more balanced heating-to-cooling load.

Figure 11 Annual heating consumption by HVAC end-use for sizing permutations in ducted new homes. *Rightsized used loosely to indicate initial sizing methodology.

Duct Distribution Effect

Variable capacity heat pumps modulate air flow rates for precise load matching. This feature introduces a complex interaction when paired with a ducted distribution system where ducts are in an unconditioned space such as an attic or crawlspace.

Often ignored or underestimated in energy modeling, there is efficiency loss from distribution. There are two main mechanisms for heat loss: duct air leakage and duct conduction. Duct air leakage from pressurized supply ducts occur through small, inevitable imperfections in duct installation, resulting in conditioned air escaping into the unconditioned space. This loss is estimated at 5% of supply flow for new build homes and 9% for existing builds with legacy ductwork (NEEA 2024). The heat pump must overproduce thermal energy and the fan must supply additional flow to account for the requirements of the conditioned spaces as well as the lost, but conditioned leakage air. Ducts within an unconditioned space act as a heat exchanger with the surrounding environment. Heat is conducted through the duct surfaces and convected and radiated to the surrounding space.

The models assume the duct surface area is constant regardless of unit capacity or design flow at 640 square feet for the combined supply and return runs based on ASHRAE Standard 152 (ASHRAE 2014). When a variable capacity system runs at a lower flow rate, it results in a lower velocity through the ducts. As with all heat exchangers, lower flow rates result in longer residence times of the air flowing through the ducts which enhance heat exchanger effectiveness. For existing homes replacing a constant speed heat pump with a variable speed heat pump with ducting in unconditioned areas, that effectively creates an "attic tax" where the low-flow operation at lower loads leads to decreased distribution efficiency, offsetting some or all savings.

Duct distribution effectiveness is the ratio of the amount of heat delivered to the conditioned space compared to the amount of heat added by the heat pump over the course of a heating season and is a calculated output of each simulation. Note that ductless systems have a distribution effectiveness of 100% since there is no air distribution system. From Table 6, the model predicted on average a distribution effectiveness for a constant volume system of between 71% and 85% for existing homes and 75% for new homes. The constant volume system is introduced here as a reference point for a traditional system, particularly in an existing home. This amounts to 15-29% loss via the attic.

The distribution effectiveness is sensitive to many factors including the duct leakage, duct insulation, and air volume. Existing homes with legacy ducts have a similar effectiveness to new homes, even though existing homes experience higher duct air leakage and lower duct insulation levels due to impact of increased air volume from the larger capacity units required with the larger loads in existing homes. When a variable capacity heat pump is utilized, the distribution effectiveness was between 63% and 79% for existing homes and 66% for new homes, respectively, a marked decrease from a constant volume system.

In existing homes, two distribution conditions were evaluated: legacy ducts (9% air leakage and R-2 duct insulation) and duct replacement (5% air leakage and R-8 duct insulation). The results in Table 6 indicate existing home duct replacement can increase distribution effectiveness by approximately 15%.

Table 6. Average system distribution effectiveness by fan control, home, and duct vintage.

Case	Average Calculated Distribution System Effectiveness
New Home, Constant Speed	75%
New Home, Variable Speed	66%
Existing Home, Constant Speed, Legacy Ducts	71%
Existing Home, Variable Speed, Legacy Ducts	63%
Existing Home, Constant Speed, Replaced Ducts	85%
Existing Home, Variable Speed, Replaced Ducts	79%

A mixed vintage (drop-in replacement of the heat pump system with the legacy ductwork) solution results in a distribution effectiveness of 63%, or over a third of energy being lost to the unconditioned space. As delivery effectiveness is highly sensitive to the total duct surface area of the distribution system in an unconditioned space, the thermal impacts may be less pronounced in homes with smaller duct footprints.

Conclusions

This study concluded that idealized performance savings estimates can vary widely when considering what occurs in the field. There are modeled savings upwards of 20% for top performing improved units compared to the average field unimproved heat pump, but there are also variations up to 20% within some cohorts between the top and bottom performers.

Overall, LLE units provided the most reliable positive savings, while CCC heat pumps had a higher likelihood of underperforming, sometimes performing worse than the *observed bottom* unimproved units. Installed heat pump units should have true variable capacity performance (e.g. Controls Verification Procedure in AHRI Standard 210/240 (2024)) and units should have evidence of field-validated performance to the published data.

LLE units provide excellent savings in mild climates, such as ASHRAE climate zone 4, while mixed or cold climates, ASHRAE climate zones 5+, should consider HPHC as these climates benefit from low load efficiency as well as cold climate capability. CCC capacity maintenance did not provide significant modeled savings, and the authors suggest further investigation into thresholds to ensure consumption and peak savings goals result from the specification.

While undersizing a heat pump can result in a large increase in supplemental energy use, oversizing by a ton moderately increased savings estimates. Oversizing any cohort of heat pump provided pseudo-CCC behavior.

For existing homes, the decision to replace an existing HVAC system with a variable capacity, variable flow heat pump should consider the condition of the existing duct system if it's located in an unconditioned space, evaluating the length and thermal performance for suitability. Implementing a mixed-vintage system solution can easily eliminate any savings associated with the variable capacity heat pump, regardless of climate or heat pump cohort. While ductless systems avoid distribution losses, efficiency losses attributed to multi-head systems have been reported, though not explored with this study.

The authors encourage usage of the *observed top* and *bottom* performance maps to better estimate the range of expected savings for residential heat pump programs. It is also the hope of the authors that the industry work toward a more consistent performance of heat pumps in the field to that which is published through manufacturer and installer engagement.

References

AHRI (Air-Conditioning, Heating, and Refrigeration Institute). 2024. *Standard 210/240-2024: Performance Rating of Unitary Air-conditioning & Air-source Heat Pump Equipment*. Arlington, VA: AHRI. www.ahrinet.org.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). 2014. *Standard 152-2014: Method of Test for Determining the Design and Seasonal Efficiencies of Residential Thermal Distribution Systems*. Atlanta: ASHRAE.

Big Ladder Software. 2026. "ResDX: Residential HVAC Performance Curves." Accessed March 16, 2026. www.bigladdersoftware.com/projects/resdx/.

BPA (Bonneville Power Administration). 2026. "Measured Field Performance and Power Consumption Data for Residential Variable-Speed Heat Pumps." Unpublished raw data.

CEC (California Energy Commission). 2025. *2025 Single-Family Residential Alternative Calculation Method Reference Manual*. Publication Number CEC-400-2025-006. Sacramento: California Energy Commission.

Efficiency Maine. 2026. "Measured Field Performance Data for Residential Air-Source Heat Pumps." Unpublished raw data.

Krishnamoorthy, S.. 2018. *System Efficiency Improvement for a Variable-Capacity/Variable-Fan-Speed Residential Heat Pump System with Ductwork*. PhD diss., University of California, Davis. Ann Arbor: ProQuest. Publication No. 10929755.

NEEA (Northwest Energy Efficiency Alliance). 2024. 2022 Residential Building Stock Assessment Findings Report: A Characterization of the Existing Residential Single-Family and Multi-Family Building Stock in the Northwest. Prepared by Evergreen Economics, Inc. Published April 2024. neea.org.

NEEP (Northeast Energy Efficiency Partnerships). 2026. "Cold Climate Air Source Heat Pump Product List." Accessed March 16. ashp.neep.org.

The CSE Authors. 2026. "California Simulation Engine v0.927.0." Accessed March 16, 2026. https://github.com/cse-sim/cse.

Wilcox, B., and A. Conant. 2018. *Laboratory Testing of Three Variable Capacity Heat Pumps*. Prepared for Pacific Gas and Electric Company (PG&E) Applied Technology Services. San Ramon, CA: PG&E.

Winkler, J. and S. Ramaraj. 2023. *Field Validation of Air-Source Heat Pumps for Cold Climates*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5500-84745. https://www.osti.gov/biblio/1973090.

AI-assisted tools used: Gemini.
Used for: Creating Figure 1, reference formatting, and light editing.
Human review: The authors reviewed and verified all content, calculations, and citations.