REPORT # MR25-006

Room Heat Pump Lab Testing Report

December 10, 2025

Prepared by:
Brian Meinrath, Consultant
Kevin McFadden, Consultant

Prepared for:
Elaine Miller, Strategy Manager
Alex Wurzel, Program Manager

Resource Innovations
719 Main Street, Suite A
Half Moon Bay, CA, 94019
(888) 217-0217
info@calmta.org

Acknowledgements

CalMTA would like to thank Midea, Gradient, Gree, and GE Appliances, a Haier Company, for their involvement and support in room heat pumps testing discussed in this report.

The manufacturers listed above provided the room heat pumps used in the lab tests; however, they did not fund or sponsor this study. All findings and conclusions are independent.

Contents

List of Abbreviations

Abbreviation	Definition
AHRI	The Air-Conditioning, Heating, and Refrigeration Institute
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BL	Building Load
CEER	Combined Energy Efficiency Ratio
COP	Coefficient of Performance
CPUC	California Public Utilities Commission
CSA	Canadian Standards Association
EPA	Environmental Protection Agency
ER	Electric Resistance
HEER	Heating Energy Efficiency Ratio
MT	Market Transformation
MTI	Market Transformation Initiative
NRTL	Nationally Recognized Testing Laboratory
RAC	Room Air Conditioner
RH_CVP	Resistance Heat Controls Verification Procedure
RHP	Room Heat Pump
RTD	Resistance Temperature Detector
VL	Virtual Load
WHP	Window Heat Pump

Executive Summary

Room heat pumps (RHPs) are a technology poised for commercialization and are the focus of a CalMTA Market Transformation Initiative (MTI). To determine the efficacy of these products, CalMTA pursued testing of five RHP units made by four manufacturers at a nationally recognized testing laboratory (NRTL) using a psychrometric chamber to evaluate heating performance and gather data to inform the market transformation (MT) strategy. Because they are an emerging technology and because they are categorized as room air conditioners under federal appliance regulations, RHPs are not subject to mandatory test procedures for heating. The purpose of this report is to share observations and takeaways from the lab testing that will inform the MTI's product plan and manufacturer engagement interventions. Through this research, CalMTA has corroborated manufacturers' ENERGY STAR® ratings, confirmed that RHPs significantly exceed the performance of portable units with electric resistance (ER) heat, and noted observations that may improve future load-based testing of RHPs.

CalMTA tested units to verify the heating energy efficiency ratio (HEER) ratings specified by the manufacturers.[1] Three of the units tested meet the ENERGY STAR classification for Type 4 room air conditioners with active defrost and operation below 5˚F, and one meets the classification for Type 3 with active defrost and operation down to 17˚F. The fifth unit is a 120V plug-in heat pump product that can be installed through a window opening in the same manner as a room heat pump but does not meet the Environmental Protection Agency (EPA) definition of a room air conditioner. For the five RHP units we tested, we observed HEER values ranging from 7.5 to 10.8, which all exceed the ENERGY STAR draft version 6 minimum HEER requirement of 6.8, and three units demonstrated values that exceed the ENERGY STAR draft version 7 requirement of 8.3. The results for all units fell within 4% of the HEER values reported by their manufacturers. All five units were able to achieve a coefficient of performance (COP) higher than 2 at 17°F, indicating strong low temperature performance for these heat pumps.[2] In addition, two of the three Type 4 units achieved a COP of better than 2 at 5°F.[3]

CalMTA also conducted load-based testing, a broad term for testing that evaluates the performance of units operating under their native controls. Load-based testing assesses the

[1] In 2024, the EPA released an ENERGY STAR voluntary test procedure for evaluating the heating performance of room air conditioners using fixed-speed heating tests. The test procedure defines a heating efficiency metric called the heating energy efficiency ratio (HEER).

[2] Based upon the H3, full fixed speed test from the ENERGY STAR voluntary test procedure for evaluating the heating performance of room air conditioners.

[3] Based upon the H4, full fixed speed test from the ENERGY STAR voluntary test procedure for evaluating the heating performance of room air conditioners.

influence of controls algorithms on product performance and produces efficiency results that can be compared to fixed-speed test results. During load-based testing, the RHPs exhibited a general increase in COP with higher outdoor air temperature, but the COPs observed do not increase uniformly at all temperatures, suggesting that load-based testing results in greater uncertainty at lower part-load conditions than near full-load conditions. There was no observable correlation between the target building load and the RHP COP.

At and near full-load operation, the efficiencies observed during load-based testing align closely with the efficiencies observed during ENERGY STAR fixed-speed testing. At part-load conditions, however, efficiencies increasingly deviate from the ENERGY STAR fixed-speed efficiencies, with ENERGY STAR fixed-speed results appearing significantly more efficient.

COPs observed during load-based testing for Type 3 and Type 4 units were similar across the two types. This suggests that Type 3 units may have similar seasonal efficiencies to Type 4 units in mild winter climates like those throughout California.

Some tests did not achieve convergence in the time allotted in this testing effort. Convergence refers to units reaching a steady state or predictable pattern of operation during testing, so that data can be reported reliably. Time limitations and interactions between RHP controls and chamber controls may have prohibited convergence from occurring. Furthermore, the load-based test method used for this effort defines a single value for room capacitance, which influences the rate at which the test room changes temperature in response to heat delivered by the RHP. Creating more accurate thermal capacitance values for single zone HVAC products may lead to more stable and repeatable test results under native controls; this report includes recommendations to improve test procedures for RHPs.

Overall, the testing confirmed manufacturer reported heating performance based on the new ENERGY STAR voluntary test procedure and demonstrated highly efficient heating down to 5°F and below for Type 4 RHPs, indicating RHPs have the capability to efficiently heat rooms across a broad range of temperatures.

1 Introduction

1.1 Overview

CalMTA's RHP MTI includes a range of product form factors under its product definition. RHPs currently fall into two different categories under federal appliance standards, both of which only have efficiency standards for the cooling cycle and not for heating. Window Heat Pumps (WHP) fall under the category of Room Air Conditioners (RAC) with Reverse Cycle. RAC form factors include saddlebag (Figure 1), U-shape (Figure 2), and traditional window units (Figure 3). These products can be self-installed with their outdoor and indoor components straddling the window sash.

From March through May 2025, CalMTA completed testing of five RHP units at an NRTL in Plano, Texas. CalMTA conducted testing to verify whether RHPs can meet the heating needs of California residents, understand how RHPs perform under part-load conditions, validate energy model assumptions, encourage manufacturers to test and publish heating performance data, and provide insights into controls logic.

Four of the products included in this testing fall under the category of RAC, while the fifth unit falls under the category of split system. The RHP MTI also includes portable heat pumps which fall under the category of Portable Air Conditioners and describes moveable products that sit on the floor and connect to outdoor air via ducts running to a window. This testing effort did not include portable heat pumps.

Figure 1. Saddlebag RAC

Figure 2. U-shape RAC

Figure 3. Traditional window unit

1.2 Background

1.2.1 Product performance

The primary performance metrics for a heat pump are efficiency and capacity for both heating and cooling modes. Unlike larger residential and commercial heat pumps, which have standard tests for both heating and cooling, RHPs are categorized as air conditioners under federal test procedures. As a result, RHPs are only tested and regulated for cooling performance. As of September 2025, there was no mandatory test procedure for RHP heating performance and relatively little testing data are available for the heating performance of RHPs when compared to more established technologies.

The Northeast and Northwest regions of the United States have generated the most interest and activity on RHPs. One challenge this presents is the information generated from these regions include heating-season design conditions that are colder than those experienced in most of

California.[4,5] As a result, manufacturers have primarily innovated and developed products that target colder climates. CalMTA performed laboratory testing to understand how available RHPs perform in California's specific climate zones and how performance compares to manufacturers' claims.

1.2.2 Cooling performance

The cooling efficiency of RACs (and heat pumps) is measured by combined energy efficiency ratio (CEER), measured in Btu/Wh, with a higher number indicating more efficient cooling. In 2026, there will be a major increase in minimum required efficiency for RACs, enacted in 2014.[6] Current minimum CEER requirements for RACs with reverse cycle range from 8.7 to 9.8, depending on unit capacity and whether the unit has louvered sides, and will increase in 2026 to range from 12.8 to 14.4.

1.2.3 Heating performance

In July 2024, the EPA released an ENERGY STAR voluntary test procedure for heating performance of room air conditioners.[7] This test defines a heating efficiency metric, HEER, also measured in Btu/Wh, and defines four categories of RHP products depending upon the operation temperature and defrost strategy:

- **Type 1:** does not have active defrost or specified compressor cut-in and cut-out temperatures are not both less than 40˚F.
- **Type 2:** has active defrost and specified compressor cut-in and cut-out temperatures are both less than 40˚F but not both less than 17˚F.
- **Type 3:** has active defrost and specified compressor cut-in and cut-out temperatures are both less than 17˚F but not both less than 5˚F.

---

[4] Consortium for Energy Efficiency. 2024. *Report 2023 Residential Sector Report*. November 5. Accessed September 30, 2025. https://cee1.org/AnnualReport2023/#slide2.

[5] Northwest Energy Efficiency Alliance. 2023. *Micro Heat Pump Field Study Results*. July 20. Accessed September 30, 2025. https://neea.org/resource/micro-heat-pump-field-study-results-product-council/.

[6] "Energy Conservation Program: Energy Conservation Standards for Room Air Conditioners" Fed. Reg. 88, no. 102 (May 26, 2023): 34298.

[7] United States Environmental Protection Agency. 2024. *ENERGY STAR Final Heating Mode Test Procedure for Room Air Conditioners (Rev. November 2024).* November 26. Accessed September 30, 2025. https://www.energystar.gov/sites/default/files/2024-11/ENERGY%20STAR%20Version%201%20Test%20Method%20to%20Determine%20Room%20Air%20Conditioner%20Heating%20Mode%20Performance%20%28Rev.%20November%202024%29.pdf.

- **Type 4:** has active defrost and specified compressor cut-in and cut-out temperatures are both less than 5˚F.

Most RHPs available today are Type 1 and do not have active defrost. Newly released saddlebag models fall into the Type 4 category with active defrost and operation below 5˚F. The development of cold climate RHPs is a major advancement for the product category and introduces products that can compete with ductless mini-splits on heating performance in small spaces.

The ENERGY STAR program released draft Version 6.0 and Version 7.0 ENERGY STAR specifications for room air conditioners in November 2024. Version 6.0 is intended to address heating mode efficiency by establishing HEER requirements in the specification for the first time.[8] As Table 1 shows, Version 6.0 requires a HEER rating of at least 5.1 for RHP Types 1 and 2 and a rating of 6.8 for RHP Types 3 and 4. In addition to increasing cooling efficiency requirements, Version 7.0 will increase heating efficiency requirements to 5.8 HEER for Types 1 and 2 and 8.3 for Types 3 and 4.[9]

Table 1. Minimum ENERGY STAR HEER rating for RHPs

	Version 6.0	Version 7.0
Type 1 and 2	5.1	5.8
Type 3 and 4	6.8	8.3

1.3 Objectives

CalMTA completed laboratory testing of RHP products to gather data and knowledge to inform program strategy and real-world operational efficiency. Aims of the testing included:

- Verify product HEER ratings and associated fixed-speed efficiencies
- Explore product performance and efficiency under native controls operation[10]

---

[8] Holly Tapani to ENERGY STAR Room Air Conditioners Brand Owner Partners and Other Interested Stakeholders. 14 January 2025. Available: https://www.energystar.gov/sites/default/files/2025-01/ENERGY%20STAR%20Version%206.0%20and%207.0%20Room%20Air%20Conditioner%20Draft%201%20Comment%20Extension%20Cover%20Letter.pdf.

[9] Holly Tapani to ENERGY STAR Room Air Conditioners Brand Owner Partners and Other Interested Stakeholders. 26 November 2024. Available: https://www.energystar.gov/sites/default/files/2024-11/ENERGY%20STAR%20Version%206.0%20and%207.0%20Room%20Air%20Conditioner%20Draft%201%20Specifications%20Cover%20Letter.pdf.

[10] Evaluate performance when operated using the controls built into each device, rather than operating unit compressors at fixed speeds, as required by many test procedures.

- Evaluate whether load-based testing results can be compared against fixed-speed testing results
- Assess whether Type 3 and Type 4 RACs exhibit a significant difference in heating performance when operating under California heating season conditions
- Verify Type 3 and Type 4 RAC capabilities at low ambient temperatures
- Evaluate whether the test conditions in current draft test procedures are suitable for mild climates
- Identify improvements to load-based test methods
- Validate or improve performance assumptions in energy models used to assess RHP products
- Build stronger support for manufacturer publication of performance metrics
- Provide insights into how built-in controls operate RHP units at different conditions
- Provide better understanding of RHP operation under part-load conditions
- Inform future code measure proposals
- Provide data that support RHP measure development in California

1.4 Equipment Tested

We tested five variable-speed RHPs made by four manufacturers. Four meet the Department of Energy definition for a RHP: a RAC that utilizes reverse cycle refrigeration as its prime source for heating the indoor space.[7] Of these four units, three meet the ENERGY STAR classification for Type 4 RAC, and one meets the classification for Type 3.

The fifth unit consists of an outdoor unit, similar to a mini-split heat pump, connected to a semi-portable indoor unit via a thin connection hose. This design forms a closed refrigerant system with no requirement for a licensed installer.

The four manufacturers represented in this report all graciously supported the testing effort through provision of target airflows, capacities, and efficiencies; specification of setting instructions for fixed-speed and load-based testing; and, in multiple cases, onsite support at the testing lab.

2 Testing Methods

2.1 Preliminary cooling tests and heat balance evaluation

While testing focused on heating, CalMTA first conducted a rated cooling performance test for all units to validate published performance data and establish the building load line for later load-based testing. Table 2 shows the test conditions for the rated cooling performance test, which matches the A_{Full} rating test as described in the Air-Conditioning, Heating, and Refrigeration Institute (AHRI) 210/240.[11] To ensure a proper heat balance between the indoor and outdoor sides of the RHP, the lab conducted the closed-duct test from the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) 37 on two of the units.[12] A heat balance was attempted on a third RHP but was unsuccessful because the unit includes a condensate spray nozzle integral to the RHP on the outdoor side of the unit. The team did not perform additional cooling performance tests because rated cooling performance and associated heat balances were already verified.

2.2 ENERGY STAR heating tests

Following the confirmation of cooling performance, the team conducted heating tests to verify the HEER ratings specified by the manufacturers, as described in the ENERGY STAR RHP test procedure.[13] Since all units tested have variable-speed compressors, manufacturer override settings were required to set the compressor speeds for the different test conditions specified in the test procedure. Table 2 shows the full set of fixed speed tests conducted. Three RHPs tested were Type 4 units, so all tests applied. One RHP tested was a Type 3 unit; for this unit as well as the moveable unit, test #7 was not performed.

---

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

[12] ASHRAE. (2024). *Standard 37: Methods of Testing for Rating Electrically Driven Unitary Air-Conditioning and Heat Pump Equipment.* Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

[13] United States Environmental Protection Agency. 2024. *ENERGY STAR Final Heating Mode Test Procedure for Room Air Conditioners (Rev. November 2024).* November 26. Accessed September 30, 2025. https://www.energystar.gov/sites/default/files/2024-11/ENERGY%20STAR%20Version%201%20Test%20Method%20to%20Determine%20Room%20Air%20Conditioner%20Heating%20Mode%20Performance%20%28Rev.%20November%202024%29.pdf.

Table 2. Fixed speed tests following ENERGY STAR RHP procedure

Category	Test #	Test Name	Indoor Air Dry Bulb [°F]	Indoor Air Wet Bulb [°F]	Outdoor Air Dry Bulb [°F]	Outdoor Air Wet Bulb [°F]	Compressor Speed
ENERGY STAR	1	A_{Full}	80	67	95	75	Full
ENERGY STAR	2	$H0_{Low}$	70	60	62	56.5	Low
ENERGY STAR	3	$H1_{Nom}$	70	60	47	43	Nominal
ENERGY STAR	4	$H1_{Low}$	70	60	47	43	Low
ENERGY STAR	5	$H2_{Int}$	70	60	35	33	Intermediate
ENERGY STAR	6	$H3_{Full}$	70	60	17	15	Full
ENERGY STAR	7	$H4_{Max}$	70	60	5	4	Maximum

Although the ENERGY STAR test procedure includes methods for evaluating low-temperature compressor cut-out and cut-in and resistance heat controls verification, evaluating these functions was outside the scope of this testing.

2.3 Load-based testing

Following completion of the fixed-speed ENERGY STAR RHP tests, the team then tested each WHP unit under its native controls to evaluate the influence of control algorithms and assess efficiency compared to the fixed-speed test results. CalMTA did not perform load-based testing on the moveable unit. Load-based testing is a broad term used to describe various methodologies currently under consideration in the HVAC industry. For this testing effort, CalMTA used a target virtual load (VL) approach, informed by Appendix I of AHRI Standard 210/240-2024 and CSA SPE-07:23. [14,15]

The VL is a fixed heating load, simulated in the test chamber by adjusting the ambient temperature setpoint of the indoor chamber in response to the instantaneous capacity delivered by the unit under test. If the delivered capacity is greater than the target VL during heating operation, the indoor room air temperature rises, simulating a space where temperature warms in response to ample heating provided by the heat pump. Alternatively, if the delivered capacity is

---

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

[15] CSA. (2023). *Special Publication Electrical-07:23: Load-based and climate-specific testing and rating procedures for heat pumps and air conditioners.* Toronto: Canadian Standards Association

less than the target VL, the indoor room air temperature falls, simulating a space where the temperature drops due to insufficient heat pump capacity.

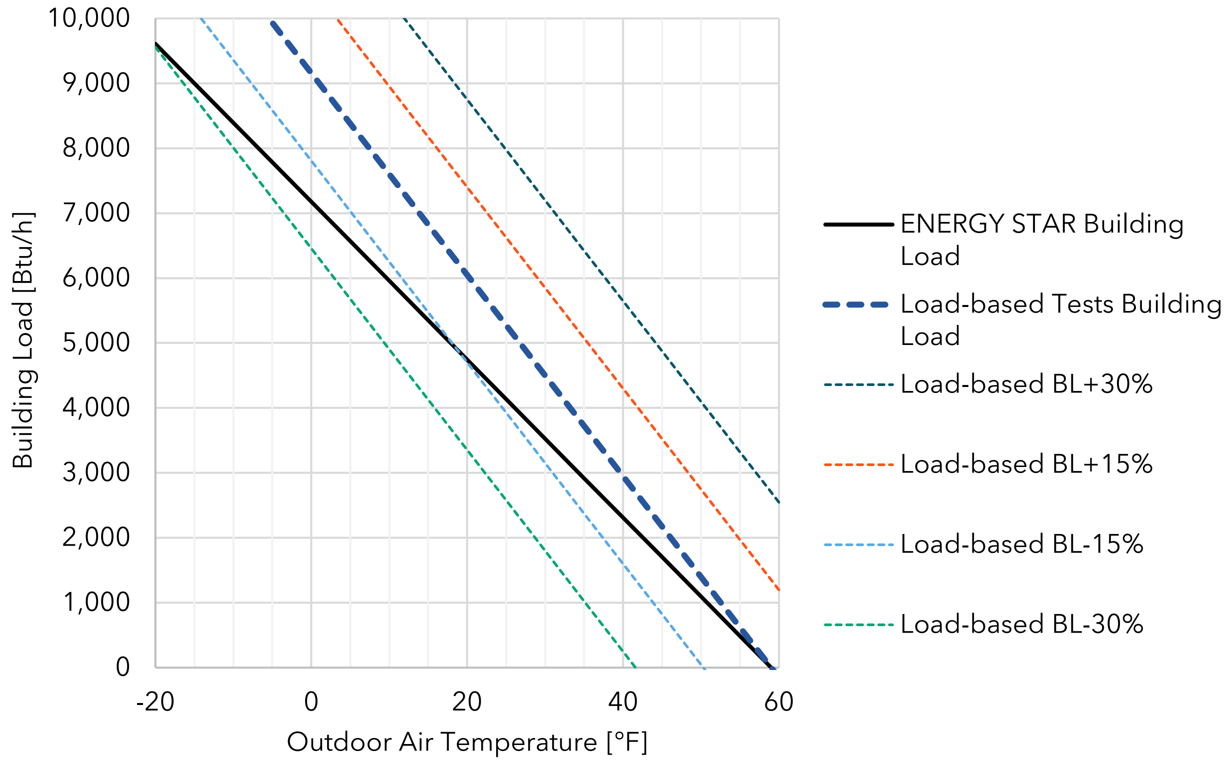
The test chamber room air temperature setpoint updating equation is given in Equation 1 (See Appendix A: Equations.). The thermal capacitance for load-based testing, which aligns with the controls verification procedure in Appendix I of AHRI 210/240, is shown in Equation 2 of Appendix A.***Error! Bookmark not defined.***

For a variable-speed heat pump, compressor speed typically increases to provide more capacity as indoor air temperature falls. If compressor speed exceeds the demand needed to meet the load, indoor air temperature rises. In response to the rising temperature, the compressor speed then modulates back down. Ideally, the capacity delivered by the heat pump converges toward the target VL, leading to a stable indoor room air temperature and directly measurable capacity and energy usage at that specific test condition. The integrated thermostat on each WHP is set to 70°F, so the test chamber is expected to stabilize near 70°F. However, it is anticipated that some units might include internal offsets in display temperature, leading chamber temperatures to deviate slightly from 70°F.

For this testing exercise, the target VL for each test point is based on the building load line shown in Equation 3, Appendix A. This equation deviates from the load line assumptions in the ENERGY STAR test procedure, which assumes an outdoor design temperature of -15°F. By mirroring the load line used for packaged terminal heat pumps, which adopts a design temperature of 1°F, the target loads at milder outdoor temperatures are more suitable for California's climate zones and more attainable by the test lab. At very low load conditions, the test lab and the units under test are unable to stabilize in an acceptable amount of time.

Load-based testing evaluated a range of outdoor ambient temperatures, with the target VL equal to the building load from Equation 3. The team conducted additional tests at 15% above the building load, 30% above the building load, 15% below the building load, and 30% below the building load. These percentages are based on the tested net cooling capacity. Figure 4 shows the difference in the building load lines for a representative 9,000 Btu/h WHP.

Figure 4. Building load lines for ENERGY STAR and load-based tests



Example test targets for a representative 9,000 Btu/h WHP are shown as percentages in Table 3 and as target loads in Table 4.

Table 3. Target percentage of tested cooling capacity, %

Outdoor Temperature [°F]	5	15	25	35	45	55
+30%	123%	106%	89%	71%	54%	37%
+15%	108%	91%	74%	56%	39%	22%
Building Load Line	93%	76%	59%	41%	24%	7%
-15%	78%	61%	44%	26%	9%	
-30%	63%	46%	29%	11%		

Table 3. Target VL in Btu/h

Target VL, Btu/h						
Outdoor Temperature [°F]	5	15	25	35	45	55
+30%	11,079	9,528	7,976	6,424	4,872	3,321
+15%	9,729	8,178	6,626	5,074	3,522	1,971
Building Load Line	8,379	6,828	5,276	3,724	2,172	621
-15%	7,029	5,478	3,926	2,374	822	
-30%	5,679	4,128	2,576	1,024		

When the target VL exceeds the tested cooling capacity, a "runaway condition" is possible. A "runaway condition" is when the indoor room temperature continually reduces and provides no usable data. For tests where the team observes this behavior, the team ends the test, and they conduct a new full-load test. The purpose of the full-load test is to measure the maximum capacity and associated efficiency at the tested outdoor air temperature. During full-load tests, the indoor room temperature is fixed at 70°F, the WHP thermostat is set to a maximum upper value, and the system is permitted to operate under native controls. For units with published performance below 5°F outdoor air temperature, an additional full-load test is conducted to evaluate the capacity and efficiency at a minimum temperature below -5°F.

Due to the limited scope of this testing effort, the team did not conduct the full collection of test points for the load-based portion of the testing for all WHPs. This is discussed further in the Results and Discussion section of this report.

At the beginning of each test, the WHPs are placed in heating mode with their thermostats set to 70°F. No adjustments are made to account for offsets or biases in display temperature versus indoor room temperature as measured by the return air thermocouples or chamber resistance temperature detectors (RTDs). The WHPs begin delivering heating capacity once the indoor room temperature, updating in response to Equation 1, falls sufficiently to trigger a demand for heating based on internal controls logic.

Test duration and the associated reported performance data are dependent of the behavior of the unit. There are four potential scenarios, as follows:

1) WHP converges to steady, stable operation at a single compressor speed and indoor ambient condition. Data are reported over a 30-minute interval during which the tolerances of ASHRAE 37 are maintained.
2) WHP does not converge, but cycles. This could be cycling between two discrete compressor speeds, cycling between one compressor speed and off, or operating in a

recurring pattern of defrost cycles. Data are reported from the end of one cycle to the end of the next cycle.

3) WHP cannot meet the VL target, resulting in runaway indoor air temperatures. In this case the test is ended and replaced with a full-load test. Data are reported from the full-load test over a 30-minute interval during which the tolerances of ASHRAE 37 are maintained.

4) WHP does not converge before a pre-determined maximum test duration. Data are reported over a 1-hour period during which the integrated delivered capacity is equal to the integrated VL. The maximum test duration varies as the test plan is executed, with earlier tests allocated three hours and later tests allocated only 90 minutes due to scheduling constraints.

All four potential scenarios are considered valid load-based tests and are suitable for reporting capacity and efficiency values. The scenarios parallel the criteria adopted the Canadian Standards Association (CSA) SPE-07:23 for achieving convergence and reporting results, aside from the duration of time required for each scenario, for which CSA generally requires longer intervals. [16]

2.4 Test facility

2.4.1 Laboratory and testing type

All testing was conducted in a psychrometric test chamber at an NRTL in Plano, TX. No calorimeter testing was conducted.

Because the focus of this testing effort was on heat pump performance, the team did not perform resistance heat controls verification procedure (RH_CVP) testing. Compressor cut-out and cut-in testing was also excluded from the scope of this testing effort.

2.4.2 Testing conditions

Each unit was installed in accordance with the requirements of ASHRAE 37, as specified in the ENERGY STAR Test Method to Determine Room Air Conditioner Heating Mode Performance.[17] Ductwork was fabricated by the NRTL for each unique unit and affixed to ensure proper supply air flow and unrestricted return air flow.

---

[16] CSA. (2023). *Special Publication Electrical-07:23: Load-based and climate-specific testing and rating procedures for heat pumps and air conditioners*. Toronto: Canadian Standards Association.

[17] EPA. 2024. *ENERGY STAR Final Heating Mode Test Procedure for Room Air Conditioners (Rev. November 2024).* November 26. Accessed September 30, 2025. https://www.energystar.gov/sites/default/files/2024-11/ENERGY%20STAR%20Version%201%20Test%20Method%20to%20Determine%20Room%20Air%20Conditioner%20Heating%20Mode%20Performance%20%28Rev.%20November%202024%29.pdf.

3 Results and Discussion

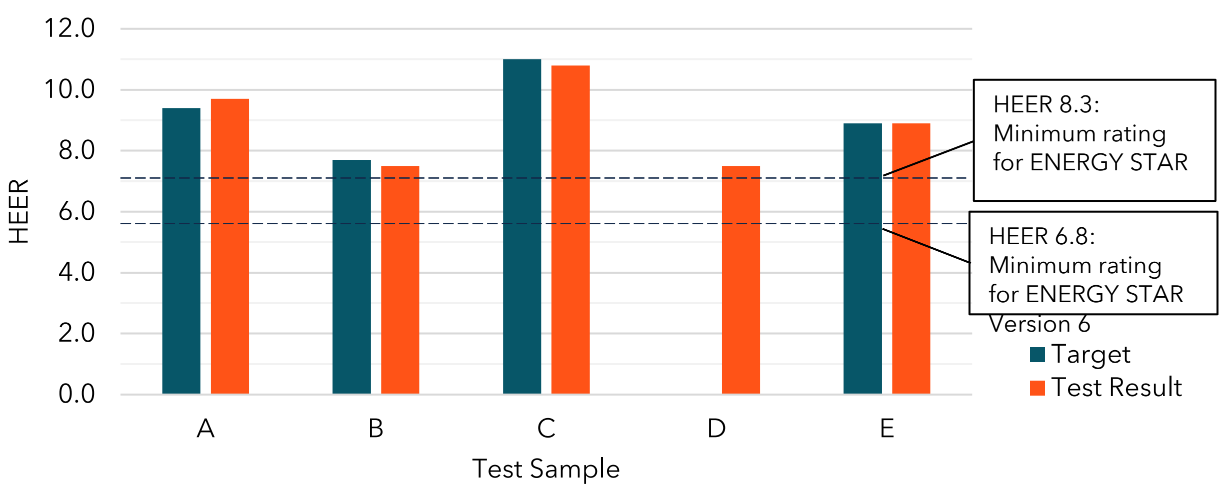
3.1 ENERGY STAR heating tests

The lab testing efforts conducted using the ENERGY STAR test procedure confirmed the HEER ratings claimed by the manufacturers.

3.1.1 HEER equivalence

Figure 5 shows the test result HEER for each RHP alongside the target HEER. Targets are based on manufacturer literature when available and otherwise are based on manufacturer-provided specifications. (One unit, identified as "D" in Figure 5 below, did not have a target HEER available from the manufacturer.) CalMTA observed test results within 4% of the target in all cases, with the minimum being 3% below the target. For those manufacturers that provided individual HEER test targets, the tested capacities exceeded the target specified capacity by about 4% on average. The minimum tested capacity was 5% below the target. Tested COP exceeded the target specified capacity by about 6% on average. The minimum tested COP was 1% below the target.

Figure 5. RHP tested HEER values versus target HEER values



The results shown in Figure 5 demonstrate that all tested units would meet the efficiency requirements to qualify for the draft Version 6.0 ENERGY STAR specification for room air conditioners of 6.8 HEER or above. Only three of the tested units would meet the requirement of the draft Version 7.0 ENERGY STAR specifications of 8.3 HEER or above.

All five units were able to achieve a COP higher than 2 at 17°F (H3,full test point), indicating strong low temperature performance for these heat pumps. In addition, two of the three Type 4 units achieved a COP of better than 2 at 5°F (H4,full test point). The COPs observed at the H0,low test point (62°F) are from 4.8 to 6.1. These observed COPs are very high and far exceed the observed COPs at higher temperatures in load-based testing as discussed in subsequent sections.

3.2 Load-based testing

Several factors influence load-based performance results, including time to achieve convergence, impacts of thermal capacitance (i.e., assumed room size and characteristics), thermostat offset and bias, and user-configurable settings. This section discusses the impacts of these factors.

3.2.1 Performance values observed during load-based testing

Figure 6 shows the heating capacity measured during each load-based test for each WHP unit. (As discussed in the Testing Methods section, CalMTA only performed load-based testing on WHP units; we did not include the moveable RHP unit.) Data are normalized to a 9,000 Btu/h WHP to maintain anonymity. The central green dashed line designates the target building load corresponding to the 9,000 Btu/h RHP. The other dashed lines, from top to bottom, represent the additional targets of 30% above the building load, 15% above the building load, 15% below the building load, and 30% below the building load, matching the lines shown in Figure 6. Each circle in Figure 6 corresponds to a measured test result, indicating close alignment with the target VLs. Each triangle in the figure represents a full-load test in which the initial load-based test was unable to maintain a sufficient indoor air temperature. The two black triangles show a Type 3 unit, and the two red triangles (which nearly overlap at 5°F and around 8,900 Btu/H) show Type 4 units as defined by the ENERGY STAR test procedure.

Figure 6. RHP delivered capacities during load-based test versus target virtual building loads

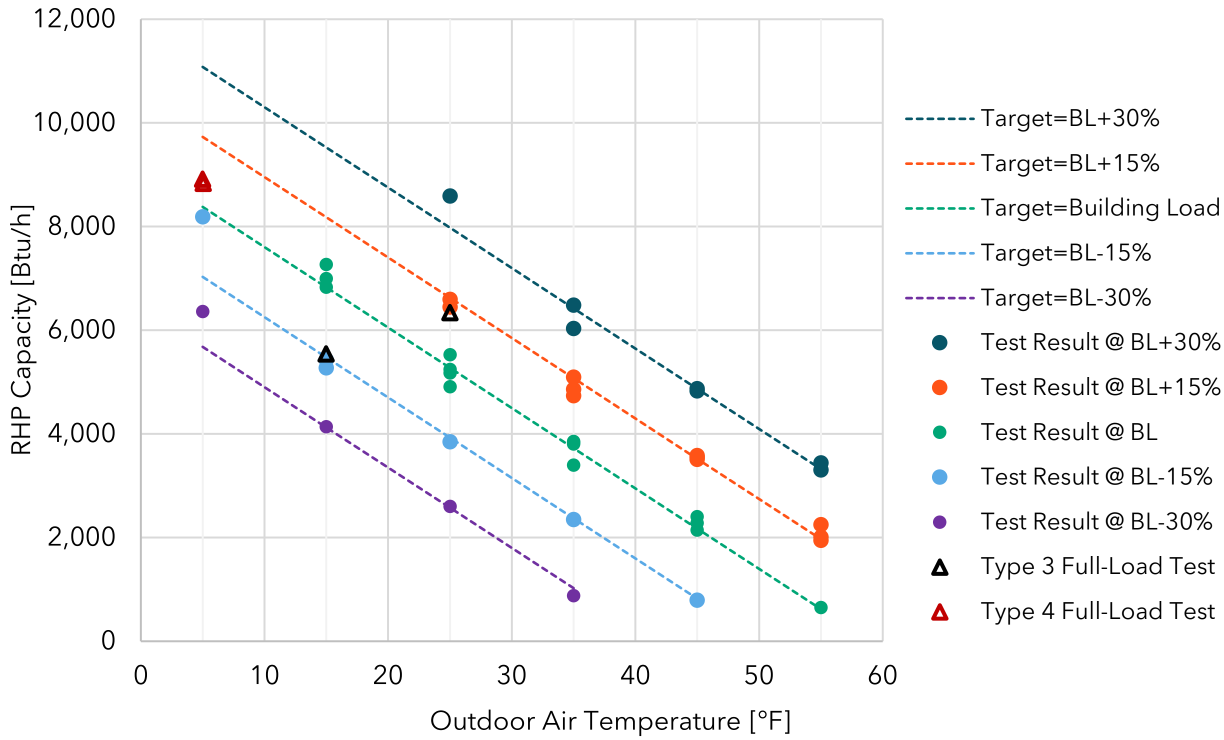
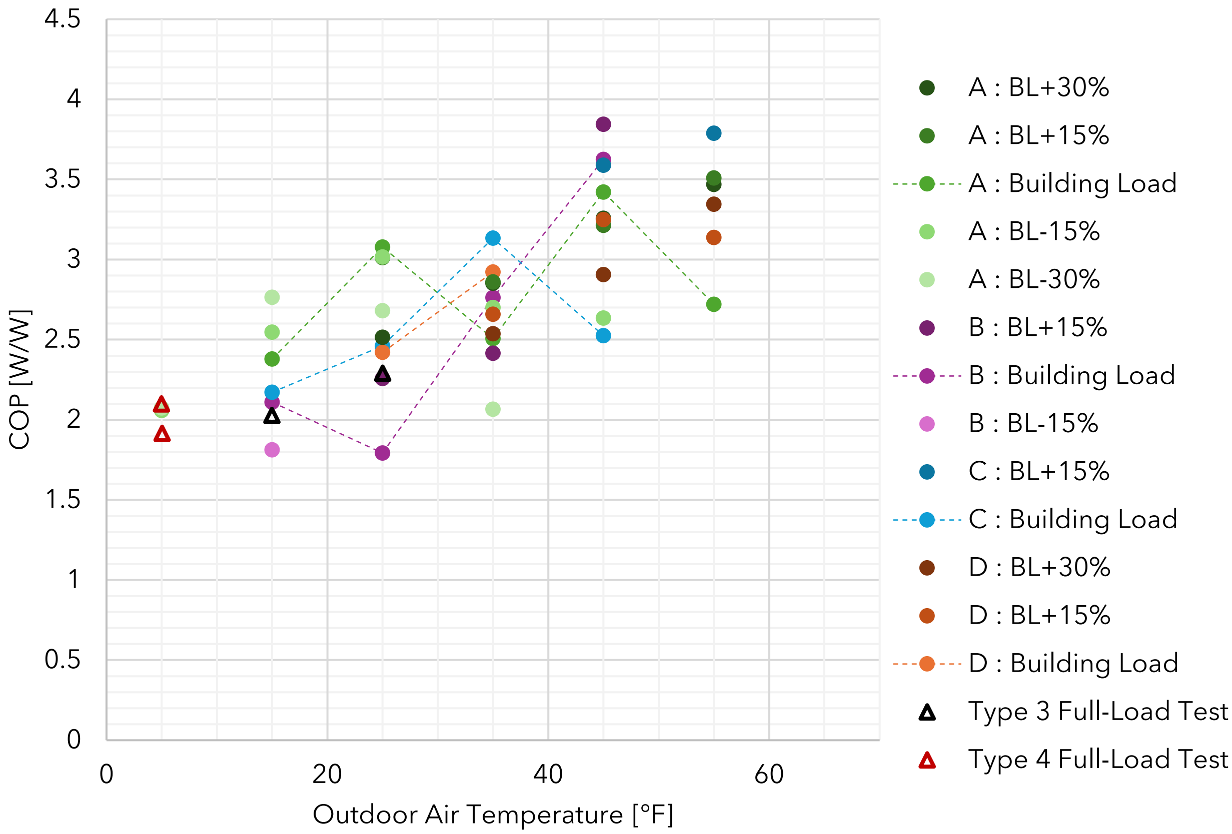


Figure 7 shows the efficiency (COP) measured during each load-based test. In the figure, each WHP is designated by a letter (A-D) and assigned a color. For each color, darker shading signifies higher VL targets: the darkest shading corresponds to 30% above the building load, and the lightest shading corresponds to 30% below the building load (at each respective outdoor air temperature). For each unit, tests conducted with the VL target equal to the building load are connected by dashed lines to illustrate that COP variability does not appear to correlate with the target load at a given outdoor air temperature. Each circle in the figure represents a load-based test. Each triangle in the figure represents a full-load test in which the initial load-based test was unable to maintain a sufficient indoor air temperature. The two black triangles show a Type 3 unit, and the two red triangles show Type 4 units as defined by the ENERGY STAR test procedure.

Figure 7. Load-based testing measured COPs



Test data shown in Figure 7 confirm that COP generally increases as outdoor air temperature increases. Load-based data from this testing effort indicate that COPs for RHPs are in the range of 2.5-4.0 above 40°F; 2.0-3.0 between 20°F and 40°F, and 1.5-2.5 below 20°F.

While COP increases with increasing outdoor air temperature, Figure 7 does not show correlation between the target building load and the WHP COP. The figure depicts this observation because neither the lighter nor darker shades consistently show higher COPs at a given outdoor air temperature. Similarly, the dashed lines do not show a steady increase with increasing outdoor air temperatures. This observation suggests that the variability in load-based testing results obscures potential conclusions around the impact of target building load on WHP COP at a given outdoor air temperature.

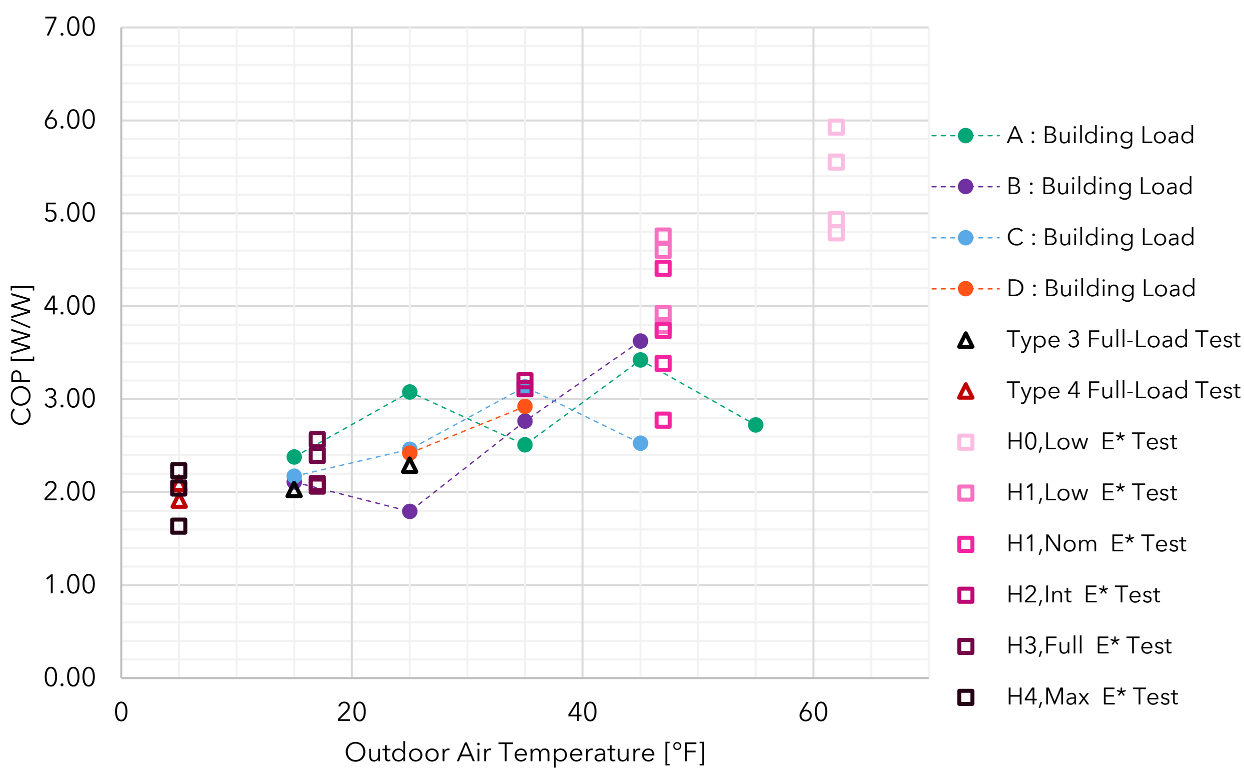
For the WHPs in Figure 7 that have multiple load-based tests at a given outdoor air temperature, the figure indicates that the range from minimum COP to maximum COP is greater at milder outdoor air temperatures. This suggests that load-based testing results in greater uncertainty at lower part-load conditions than near full-load conditions. This also indicates that the efficiency of a given unit may vary more widely at low part-load conditions than near full-load conditions.

While the capacities of the Type 3 units tested were limited at colder temperature when compared to the Type 4 units, as evidenced by the full-load test capacities in Figure 6, the COPs observed during load-based testing for the Type 3 and Type 4 units were similar, based on the data in Figure 7. This suggests that Type 3 units may have similar seasonal efficiencies to Type 4 units in relatively mild climates like California.

In addition to Figure 7 depictions, two of the Type 4 units claimed the ability to deliver heat pump capacity below -5°F. Load-based, full-load testing confirmed that both units maintained capacity at more than 80% of the rated value and exceeded COP of 1.5.

Figure 8 adds unit COP observed during fixed-speed (ENERGY STAR) tests to the data from Figure 7. The squares labeled $H0_{Low}$ through $H4_{Max}$ correspond to the test conditions specified in the ENERGY STAR test procedure listed in Table 2.

Figure 8. Load-based testing measured COPs versus fixed-speed ENERGY STAR (E*)[18] HEER testing measured COPs



At and near full-load operation, Figure 8 indicates that the efficiencies observed during load-based testing align closely with the efficiencies observed during ENERGY STAR fixed-speed

[18] ENERGY STAR has been abbreviated to "E*" in Figure 8 to improve readability.

testing. The figure represents this observation by comparing the 5°F and 17°F ENERGY STAR test results (squares) with the 5°F and 15°F load-based test results (triangles). At milder part-load conditions, such as the 45°F and 55°F load-based tests and the 47°F and 62°F low-speed ENERGY STAR tests ($H1_{Low}$ and $H0_{Low}$), efficiencies observed during load-based testing increasingly deviate from the ENERGY STAR fixed-speed efficiencies. Consequently, the ENERGY STAR fixed-speed results appear significantly more efficient.

While this data set implies that the ENERGY STAR HEER rating may overstate the efficiency of a WHP operating under native controls, this conclusion would be premature. The following sections discuss CalMTA's findings on convergence and room capacitance and describe how additional development of load-based testing procedures and further testing is needed to reach definitive conclusions about the efficiencies of WHPs operating under native controls compared to operating during fixed-speed tests.

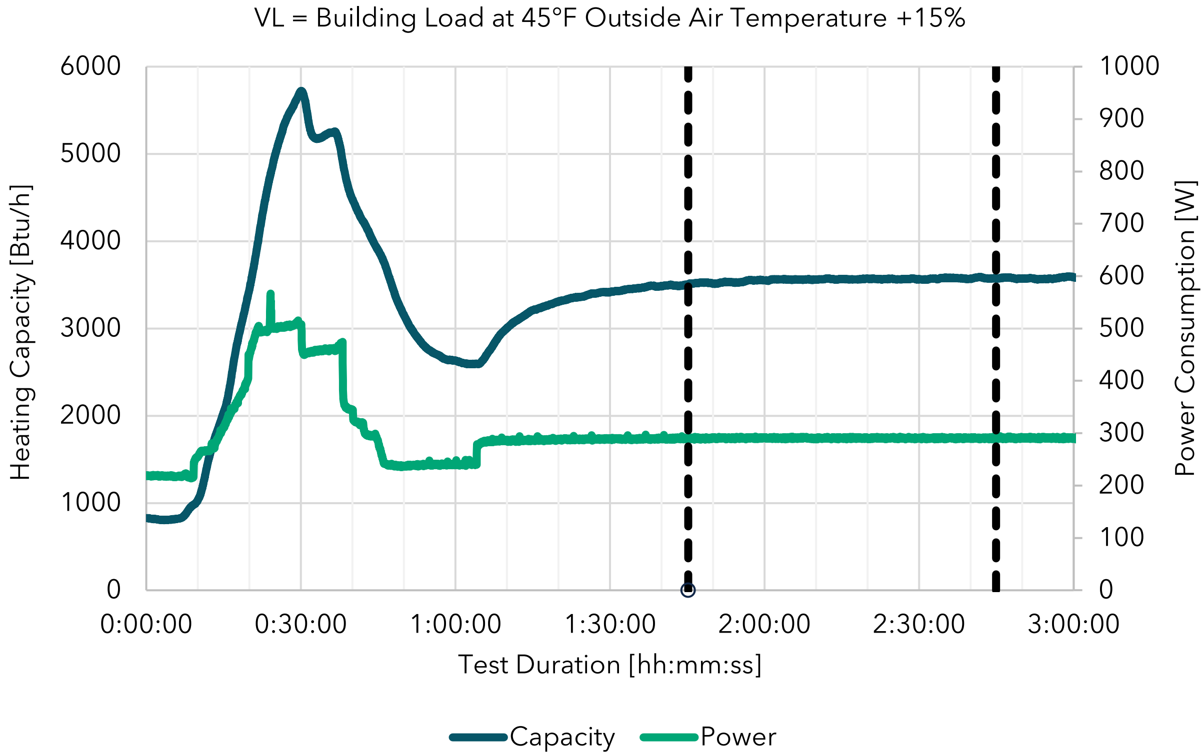
3.3 Convergence conditions during load-based testing

As covered briefly in the Testing Methods section, the data reported for each load-based test vary depending on the behavior of the unit during the test and whether set points and indoor conditions converge. Convergence refers to the occurrence during a test of a sufficient length of time during which data can be reported reliably. Data can be reported reliably because predetermined test tolerances have been met and output parameters such as capacity and power consumption have stabilized into either steady-state operation or a predictable pattern. Convergence criteria are intended to ensure that load-based tests can be conducted repeatably and reproducibly. Figure 9 through Figure 12 illustrate four convergence scenarios from test data in which the area between the vertical dashed lines represents the period from which data are reported.

3.3.1 Convergence to steady-state operation

Figure 9 shows the results of one unit tested at 45°F outside air temperature targeting roughly 3,500 Btu/h. In this example, the time between the vertical dashed lines represents a period during which the WHP's controls converged to stable operation, presumably at a single compressor speed.

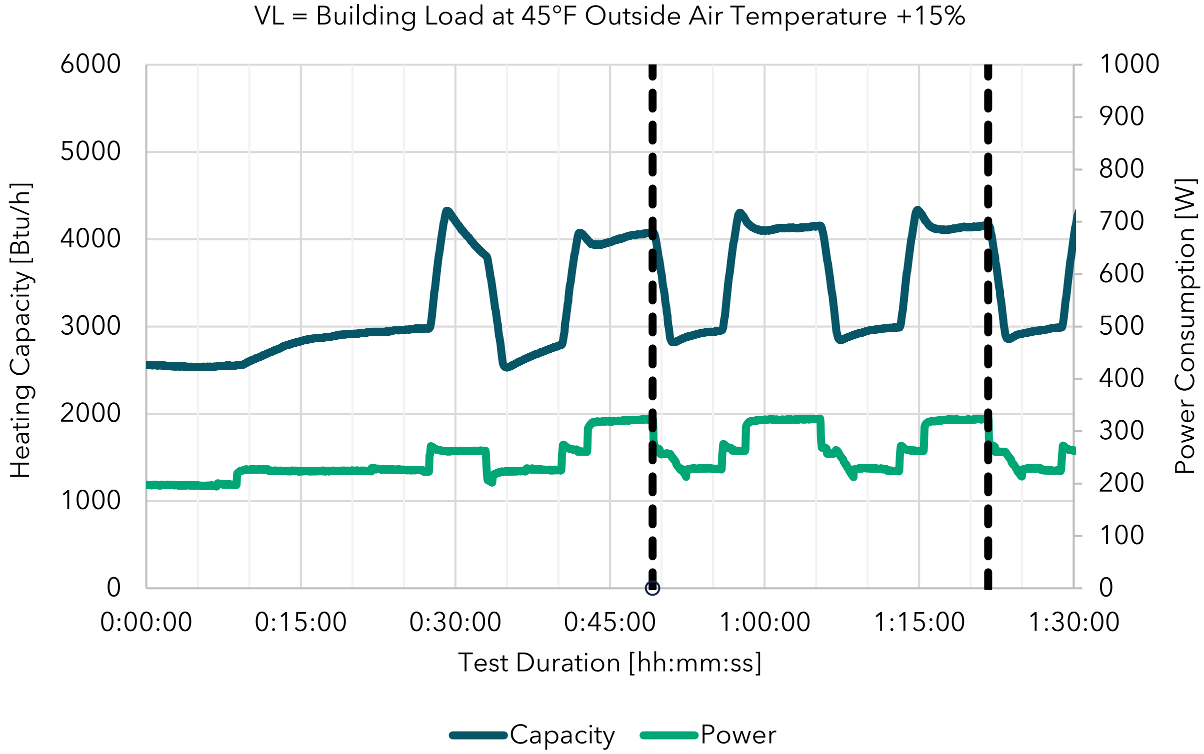
Figure 9. Convergence to steady-state operation



3.3.2 Convergence to a repeating pattern

Figure 10 shows the results of another unit tested at 45°F outside air temperature, again targeting roughly 3,500 Btu/h. In this example, the time between the vertical dashed lines represents a period during which the WHP's controls did not stabilize at a single condition, but the unit did operate in a consistently repeating pattern. The data are considered to have converged and were reported over two consecutive time periods.

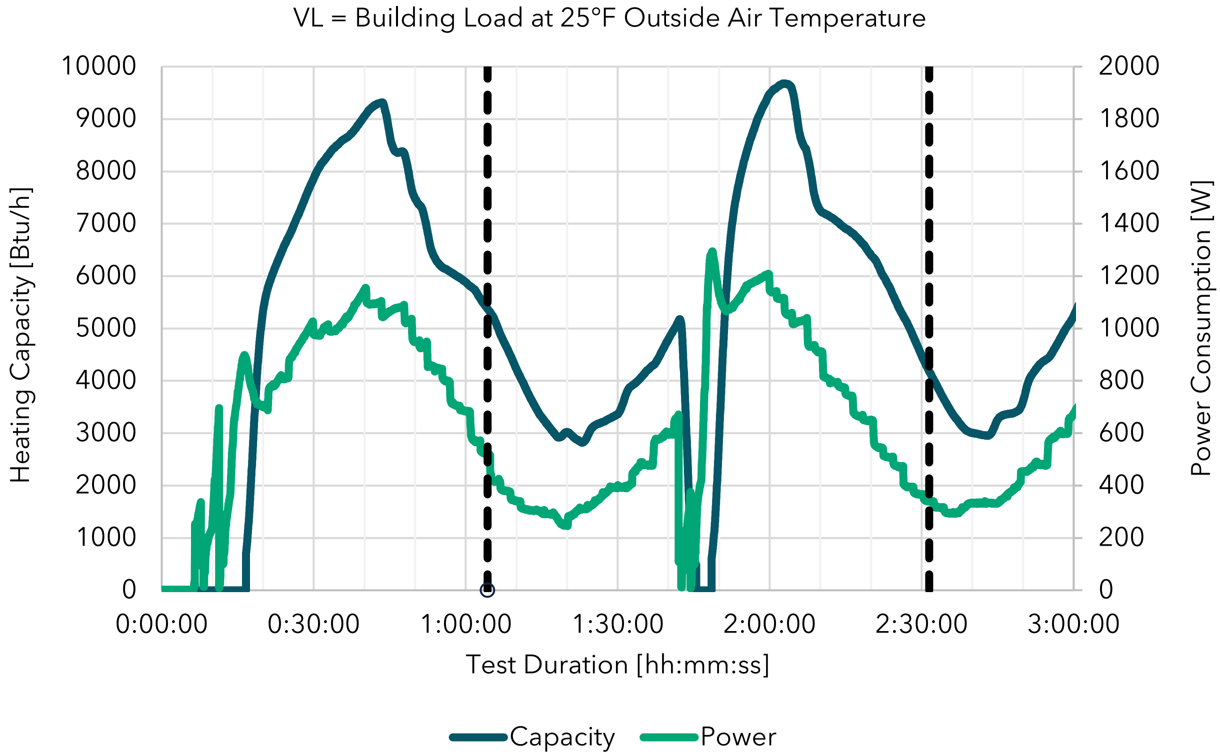
Figure 10. Convergence to a repeating pattern



3.3.3 Test tolerances met but convergence not obtained

Figure 11 shows the results of a unit tested at 25°F outside air temperature, targeting approximately 5,200 Btu/h. In this example, the time between the vertical dashed lines represents a period during which the WHP's controls did not converge to a single condition or to a consistently repeating pattern prior to the conclusion of the test period, although the test room conditions met test condition tolerance requirements. The reported data are from a period during which the average capacity was near the target VL, in this example being 5,200 Btu/h.

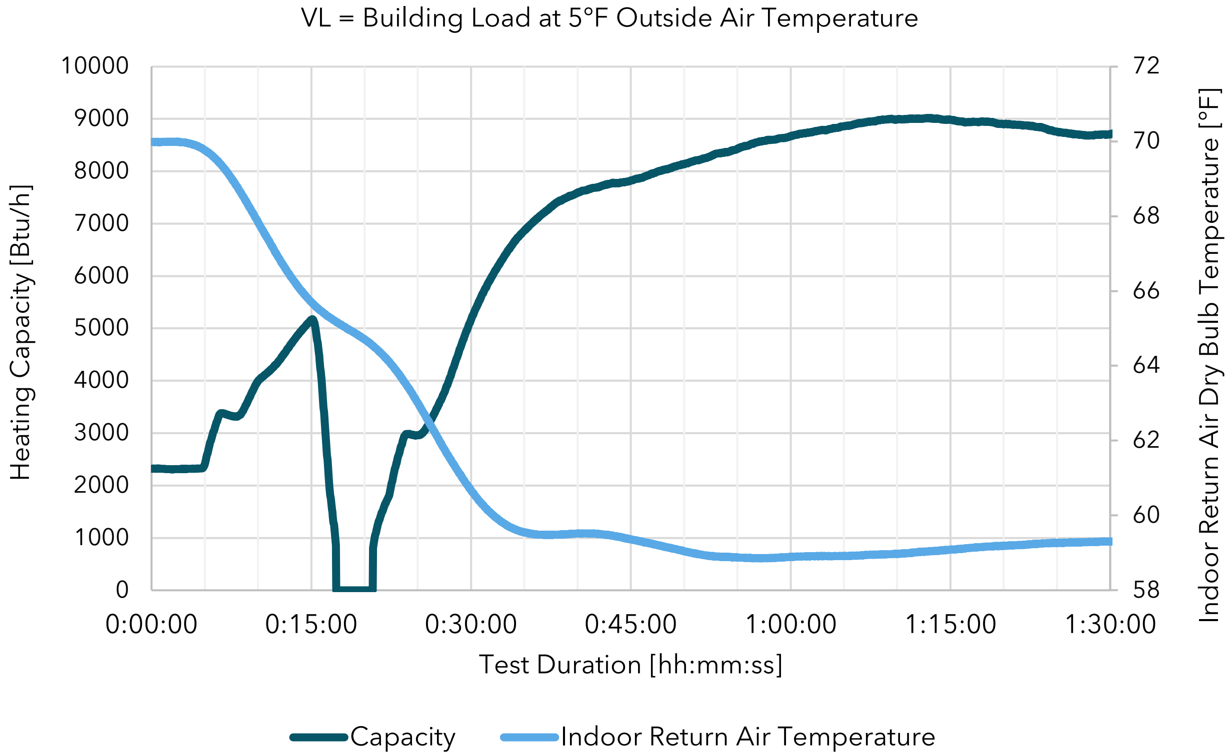
Figure 11. Test tolerances met but convergence not obtained



3.3.4 Runaway indoor air temperature, triggering replacement by full-load test

Figure 12 shows the results of a unit tested at 5°F outside air temperature, targeting approximately 8,400 Btu/h. In this example, the delivered heating capacity suggests convergence, but the WHP was unable to raise the indoor temperature acceptably close to the thermostat setpoint, as shown by the 59°F indoor temperature one hour into the test. In this case, the test data shown are not reported; instead, a separate full-load test was conducted at the same 5°F outside air temperature condition to ensure that the reported capacity and efficiency are sufficiently close to the target indoor air temperature of 70°F.

Figure 12. Runaway indoor air temperature, triggering replacement by full-load test



3.4 Factors introducing uncertainty in load-based testing

These convergence examples highlight the subjective nature of some aspects of test data reported during this testing effort and provide a starting point for discussion of testing uncertainty.

3.4.1 Insufficient time to achieve convergence

Due to limited testing scope, the team predetermined the maximum duration for tests. Duration was generally based on familiarity with previous load-based testing in other HVAC product categories. Without prior data on unit controls behavior and test chamber response characteristics, the time to reach convergence was not known. We limited many tests to 90 or 120 minutes.

Some tests did not achieve convergence in the allotted time. For some cases, typically at the coldest outdoor ambient conditions, the units were unable to meet the target VL (as in Figure 12).

In other cases, data suggest that the units may have eventually converged, but the interactions between WHP controls and chamber controls prohibited convergence from occurring in the time allocated (as in Figure 11). For these test cases, determining usable data was somewhat subjective and more rigorous criteria, such as the AHRI 210/240 controls verification procedure, could not be applied.Error! Bookmark not defined. While factors like test operating tolerances and target capacities were prioritized, CalMTA acknowledges that the load-based test data reported here provide insight into WHP performance but would not be suitable for critical performance evaluations, such as those for product certification or standards enforcement.

3.4.2 Room thermal capacitance assumptions (e.g., assumed room size and construction)

During load-based testing, some units exhibited on-off cycling behavior where the fixed-speed ENERGY STAR tests would suggest that steady-state operation should occur (e.g., operating a low compressor speed under low-load conditions). This cyclic operation drives lower efficiency than anticipated at these conditions and suggests a lower overall seasonal efficiency. However, it is premature to conclude that these units are equipped with sub-optimal controls logic. It may be the case that the controls are simply optimized for a different sized space than what is assumed in the test plan. As discussed in the Testing Methods Section, the load-based test method used for this effort defines a single value for room capacitance (Equation 2) which influences the rate at which the test room changes temperature in response to the heat delivered by the WHP.

The value of 24 °F/hr in the denominator of the equation, used also in AHRI 210/240 and CSA SPE-07, is based on centrally ducted systems. Given that centrally ducted systems condition entire living spaces, these results could indicate that a different value may be more appropriate for RHPs and other HVAC equipment types intended for individual rooms. A lower capacitance would reflect the faster reaction time of smaller spaces; this would be represented by a denominator value higher than 24 °F/hr.

Stated differently, on-off cycling behavior observed during this testing should not be taken as a clear call for manufacturers to update their controls logic. It may be that the manufacturer has optimized their units' controls for spaces with a capacitance that is different from what the load-based test methodology suggests. CalMTA suggests further research to inform a more appropriate capacitance value for HVAC products conditioning individual rooms. This additional research would benefit initiatives that promote RHPs as well as other HVAC types like ductless mini-splits. The recommended research could include lab testing to assess typical thermostat response times as well as field validation of existing RHP controls to verify whether current controls perform efficiently in real-world applications.

3.4.3 Thermostat offset and thermostat bias

The RHP lab testing effort did not include thermostat offset and bias. Thermostat offset refers to an intentional difference that a manufacturer programs into controls software between the return air temperature measured by a sensor and the temperature used internally to dictate controls decisions. The offset temperature may be the value displayed on the RHP, rather than what the RHP itself measures. The offset temperature may also be related to the thermostat deadband.

Thermostat bias, which is a distinct concept from thermostat offset, refers to unintentional differences between the temperature value measured by the RHP's return air sensor and the dry bulb temperature value measured by a lab's test chamber sensor near the same physical location. For example, in an ideal load-based test targeting an indoor return air temperature of 70°F, the reported performance would occur with the unit delivering heating while converging at 70°F. In practice, an initial test will frequently converge at a different temperature, due to thermostat offset and bias.

Rigorous load-based testing should include preliminary tests and subsequent adjustments to thermostat setpoints or internal settings to ensure that official results are reported with convergence at, or very close to, the target indoor return air temperature. This testing effort did not conduct preliminary testing to inform adjustment for thermostat offset or bias due to budget and time constraints. In practice, nearly all results reported here occurred with indoor return air temperatures between 66°F and 69°F. Tests yielding indoor temperatures substantially lower than 70°F were not reported, with full-load test results reported instead, as noted earlier in this report.

3.4.4 User-configurable settings

While some test units appear to offer only one heating mode, others included options such as "eco" modes. All units tested for this effort were set to a standard heating mode for load-based tests, with no further testing to consider alternate heating modes. Units similarly offered multiple settings for cooling mode, suggesting that controls prioritize comfort, efficiency, and/or dehumidification to different degrees. Assessing variations in performance and efficiency under native controls from different settings available to users was outside of the scope of this testing effort.

4 Conclusions

4.1 Summary

- In all test cases and temperature ranges the efficiency of RHPs significantly and consistently exceeds that of ER heat, even at the coldest temperatures tested.
- This study's fixed-speed tests validated manufacturers' efficiency claims under the ENERGY STAR test procedure. For tested RHPs that had manufacturer-provided HEER test targets, all of them came within 5% below their target-specified capacities and within 1% below target COP. Load-based tests resulted in wide efficiency variations.
- This study observed similar efficiencies during fixed-speed and load-based testing at and near full-load operations. At milder part-load conditions, however, fixed-speed tests indicated significantly higher efficiencies than load-based tests.
- Under load-based testing at milder outdoor temperatures, consistent with California's climate, the Type 3 and Type 4 units tested demonstrated similar efficiencies. Given that Type 3 units are less expensive than Type 4 units, the former are likely more cost-effective for California consumers.

- For units capable of operating below -5°F, load-based full-load testing demonstrated COP values exceeding 1.5.

5 Recommendations

5.1 Improvements to load-based testing

CalMTA undertook load-based testing to obtain and share data on RHP energy usage and bill rates when operating under native controls. The goal was to compare results from load-based testing to those from fixed-speed testing and to evaluate the suitability of load-based testing procedures as applied to RHP units. Lack of convergence and variability in load-based testing results suggest that load-based testing is not suitable for compliance or enforcement testing without refinement of the test procedure. In its current state, load-based testing remains useful for validating fixed-speed performance of systems near full-load operation and for identifying potential shortcomings in unit control strategies.

Lower capacitance values may be more appropriate for HVAC equipment intended for individual rooms. CalMTA recommends that industry groups and organizations fund and conduct further research to inform a more appropriate capacitance value for RHPs and other individual room-focused equipment types, such as ductless mini-splits. This research could include further lab testing and field verification. By using appropriate room capacitance values, future lab testing will likely achieve convergence more quickly and frequently, reducing the cost and time required for testing, and facilitating more data availability on RHP performance for RHP programs.

Appendix A: Equations

Equation 1: test chamber room air temperature setpoint updating equation:

$$RAT(t + \Delta t) = RAT(t) - \frac{\Delta t \cdot [VL - \dot{q}_h]}{C}$$

Where:

- $RAT(t)$ = the current indoor dry-bulb temperature setpoint for the indoor room reconditioning system, °F
- $\dot{q}_h$ = the net heating capacity provided by the unit under test in the current time step, as determined by air-side measurements, Btu/h
- Δt = the time interval for updating the indoor room reconditioning system controller setpoint, h
- C = the simulated thermal capacitance of the building interior, Btu/°F

Equation 2: thermal capacitance for load-based testing. (This equation aligns with the controls verification procedure in Appendix I of AHRI *210/240*.):

$$C = \frac{1}{4} \cdot \frac{\dot{q}_{A,Full,sensible}}{N_{max} \cdot \Delta T_{db}} = \frac{SHR_{A,Full} \cdot \dot{q}_{A,Full}}{24}$$

Where:

- $SHR_{A,Full}$ = the sensible heat ratio delivered by the unit under test during the A_{Full} test, dimensionless
- $\dot{q}_{A,Full}$ = the net cooling capacity delivered by the unit under test during the A_{Full} test, Btu/h
- N_{max} = the assumed maximum on/off cycling frequency of the unit under test = 3 cycles/h
- ΔT_{db} = the assumed thermostat deadband of the unit under test = 2°F

Equation 3: building load line as basis for target VL:

$$BL(t_j) = \frac{\{59 - t_j\}}{\{59 - 1\}} \cdot \dot{q}_{A,Full}$$

Where:

- t_j = the outdoor air dry-bulb temperature
- $\dot{q}_{A,Full}$ = the net cooling capacity provided by the unit under test, as measured during the A_{Full} cooling test conducted during this testing exercise