4E Energy Efficient End-Use Equipment
International Energy Agency

AC/HP Test Methods 2.0: Phase 4 Findings Summary

DECEMBER 2024

Technology Collaboration Programme
by iea

About 4E

The Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP), has been supporting governments to co-ordinate effective energy efficiency policies since 2008.

Fourteen countries and one region have joined together under the 4E TCP platform to exchange technical and policy information focused on increasing the production and trade in efficient end-use equipment. However, the 4E TCP is more than a forum for sharing information: it pools resources and expertise on a wide range of projects designed to meet the policy needs of participating governments. Members of 4E find this an efficient use of scarce funds, which results in outcomes that are far more comprehensive and authoritative than can be achieved by individual jurisdictions.

The 4E TCP is established under the auspices of the International Energy Agency (IEA) as a functionally and legally autonomous body.

Current members of 4E TCP are: Australia, Austria, Canada, China, Denmark, European Commission, France, Japan, Korea, Netherlands, New Zealand, Switzerland, Sweden, United Kingdom and the United States.

Further information on the 4E TCP is available from: www.iea-4e.org

Acknowledgements

Prepared for the Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP).

Developed by:

Rick Huddle, Cadeo Group
Mike Stem, Stem Integration
Carl Cochran, Stem Integration
Michael Brychta, Stem Integration
Marcus Dimeo, Cadeo Group
Sarah Widder, Cadeo Group

Executive Summary

Introduction

The introduction of variable capacity, or "inverter-driven," air conditioners (AC) and heat pumps (HP) represents a significant advancement in residential heating, ventilation, and air conditioning (HVAC) technology. These units, promoted as the most efficient equipment available, utilize modulating controls that allow for variable compressor and fan speeds, enhancing their performance and energy efficiency. Despite their benefits, current regulatory test procedures fail to capture the impact of these modulating controls because they typically lock the compressor and fans at fixed speeds, ensuring reproducibility but not always representing real-world performance.

The Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP) hired Cadeo Group to develop and write a new test method for residential ACs and HPs that will better evaluate performance under native controls. The research for this effort was conducted in four phases:

1. Investigate Innovative Test Methods: Researching various AC and HP test methods to identify key issues related to load-based testing.
2. Investigative Testing of Key Issues: Conducting laboratory trials to understand how to test these issues and understand the feasibility of implementing load-based test concepts.
3. Development of Load-based Test Methodology: Developing the 4E Test Method, which uses unlocked controls and a target compensation load to more accurately measure performance.
4. Round Robin Trial of Test Procedure: Conducting round-robin testing across multiple international laboratories to evaluate the reproducibility and representativeness of the 4E Test Method.

In this final phase of the project, Cadeo performed round-robin testing of the test method at four laboratories in three different 4E member countries. The primary goals were to determine reproducibility of the 4E Test Method, assess its representativeness in highlighting differences from existing regulatory procedures, and gather insights into the test burden.

The team provide these key takeaways:

- **The 4E Test Method has high reproducibility.** Comparing results from lab to lab, the coefficients of variance were below 10% in all but one test condition for one unit.
- **The 4E Test Method is more representative than locked controls tests.** The 4E Test Method highlights differences in efficiency when compared to locked-controls testing.
- **The 4E Method has a higher test burden than locked controls tests.** While test burden was high for all labs due to labs' difficulty maintaining stability, a virtual building load can help improve the stability of units, which will likely also reduce test time and further improve reproducibility.

The 4E Test Method is already influencing test procedures. The United States regulatory test procedure for air conditioners and heat pumps now includes a controls verification procedure based on the 4E method. By incorporating a virtual building load, the research team expects the 4E Test Method will improve the representativeness of air conditioner and heat pump testing in other jurisdictions as well.

Contents

List of Tables

List of Figures

List of Equations

Abbreviated Terms

AC	Air Conditioner
AHJ	Authority Having Jurisdiction
ASHP	Air-source Heat Pump
Btu/h	British Thermal Units per Hour
°C	Degrees Celsius
COP	Coefficient of Performance (Watt/Watt)
CV	Coefficient of Variance (%)
CVP	Controls Verification Procedure
EER	Energy Efficiency Ratio (Btu/Watt)
EEV	Electronic Expansion Valve
°F	Degrees Fahrenheit
HP	Heat Pump
kW	Kilowatt
SD	Standard Deviation
T-stat	Thermostat
UUT	Unit Under Test
W	Watt

1 Background and Introduction

Residential air conditioner (AC) and heat pump (HP) manufacturers have increasingly introduced variable capacity, or "inverter-driven," units to the market, promoting them as the most efficient equipment available. However, the current AC and HP test procedures used for regulatory purposes fail to capture the impact of the modulating controls, which have a critical impact on the performance of these variable capacity units.[1] Existing regulatory test procedures lock the compressor and fans at fixed speeds. While this locked approach ensures repeatable results, it fails to capture the effect of the "native" controls.[2] Multiple organisations[3] are developing new load-based methods for testing these products to ensure test procedures and metrics are representative of field performance.

A recent examination of current international test procedures and metrics[4] identified recommendations to improve international alignment and better understand the issues and challenges surrounding new test methods for variable capacity ACs and HPs. Consistent, coordinated test procedures are crucial to providing reliable performance metrics to consumers, meaningful drivers for product developers, and decreasing the test burden on manufacturers attempting to comply with various regulatory schemes. This research follows those recommendations and focuses on resolving issues with load-based testing of variable capacity ACs and HPs.

Research Overview and Goals

This research aims to develop an internationally applicable load-based test method for variable capacity ACs and HPs. The work has four phases:

1. **Investigate Innovative Test Methods**
2. **Investigative Testing of Key Issues**
3. **Development of Load-based Test Methodology**
4. **Round Robin Trial of Test Procedure**

Phase 1 of this research[5] explored various AC and HP test methods and identified key issues related to load-based testing of variable capacity ACs. Phase 2[6] included laboratory testing of variable speed ACs and HPs to further investigate the identified key issues, uncover differences between load-based test concepts, understand the feasibility of implementing these different concepts, and evaluate any potential increased test burden. In Phase 3, the research team developed and wrote a method, called the 4E Test Method, addressing the key issues uncovered in previous phases.

This report documents the results of Phase 4: Round Robin Trial of Test Procedure, where the 4E Test Method is performed in labs worldwide. The research team first discusses the test plan, then presents the test results, and finally provides recommendations for next steps.

---

[1] "AC/HP Test Methods 2.0: Phase 1 Findings Summary," International Energy Agency/4E, https://www.iea-4e.org/wp-content/uploads/2021/08/AC-HP-Test-Methods-Phase-1-Key-Findings_Revised.pdf

[2] "Native" controls mean the programmed control logic from the manufacturer.

[3] CSA SPE-07:23 "Load-based and climate-specific testing and rating procedures for heat pumps and air conditioners," CSA Group, https://www.csagroup.org/store/product/CSA%20SPE-07:23/; ISO/TC 86/SC 6 "Testing and rating of air-conditioners and heat pumps," International Organization for Standardization, https://www.iso.org/committee/50376.html; AHRI 1600 (I-P) "Performance Rating of Unitary Air-conditioning and Air-source Heat Pump Equipment," Air-Conditioning, Heating, and Refrigeration Institute, https://www.ahrinet.org/search-standards/ahri-1600-i-p-performance-rating-unitary-air-conditioning-and-air-source-heat-pump-equipment

[4] "Domestic Air Conditioner Test Standards and Harmonization," International Energy Agency/4E, https://www.iea-4e.org/publications/?_sf_s=domestic%20air%20conditioner%20test%20standards%20and%20harmonization

[5] "AC/HP Test Methods 2.0: Phase 1 Findings Summary," International Energy Agency/4E, https://www.iea-4e.org/wp-content/uploads/2021/08/AC-HP-Test-Methods-Phase-1-Key-Findings_Revised.pdf

[6] "AC/HP Test Method 2.0: Phase 2 Findings Summary," International Energy Agency/4E, https://www.iea-4e.org/wp-content/uploads/2021/08/AC-HP-Test-Methods-Phase-2-key-Findings-2021-08-06-CLEAN.pdf

2 Round Robin Test Plan

Current regulatory tests measure performance under static conditions with overridden controls. Units are typically put in a “test mode” that locks compressor and fan speeds. This provides a repeatable[7] and a reproducible test.[8] For single-speed ACs/HPs, this approach is arguably representative of how the unit will perform in the field. However, for variable-speed units, the manufacturer-installed control algorithms may alter performance, especially compared to the test mode.

Previous phases of this project identified and performed investigative testing on key issues (equipment setup, lab setup, and test concept) related to load-based testing.[9] Based on the findings of the previous phases, the research team developed the 4E Test Method, which tests variable-speed ACs and HPs using unlocked controls while imposing a target compensation load.[10] The purpose of this round robin testing is to determine the difference in results using locked controls versus results using the 4E Test Method.

The team recruited laboratories from 4E member nations to participate in the round robin testing. The primary goals of this testing were to:

1. Determine reproducibility (i.e., consistency of results from lab to lab) of the test method.
2. Determine representativeness of the test method by showing any differences from regulatory test procedures.
3. Gain insights into the burden of the test.

The following sections summarize the insights gained from the investigative testing on key issues and the subsequent test method used for round robin testing.

Load-Based Test Concept

In ongoing discussions with other test researchers, the research team found that the load-based tests can be divided into two testing concepts: pre-defined loads and adaptive loads:

- **Pre-defined load testing** imposes predetermined sensible and latent loads on the unit under test (UUT). These predetermined loads can be either constant or variable and could be based on building models, temperature conditions, tested capacities, or other factors that influence the operation or the performance of the equipment.
- **Adaptive load testing** also imposes sensible and latent loads on the UUT, but these adaptive loads are updated periodically[11] based on the response of the UUT. Adaptive load testing often uses a “virtual load”—software that mimics the response of a typical home.

Both load-based concepts are performed under native controls and thus distinguish themselves from fixed parameter tests that lock the compressor speed, fan speeds, and/or valve position. Table 1 provides comparison between fixed parameter tests and the two load-based test concepts.

[7] Repeatable means every time a given lab performs the test the results are the same.

[8] Reproducible means that different labs will produce the same results.

[9] https://www.iea-4e.org/wp-content/uploads/2021/08/AC-HP-Test-Methods-Phase-2-key-Findings-2021-08-06-CLEAN.pdf

[10] In target compensation load testing, outdoor conditions are maintained (similar to fixed-speed testing), but the indoor side of the test chamber has a constant (heating or cooling) load imposed on it.

[11] Typically using a virtual building emulator.

Table 1: Test Concepts

	Fixed Parameter Test	Pre-Defined Load Test	Adaptive Load Test
Load	Balances with AC/HP capacity	Not affected by AC/HP performance	Adjusts in response to AC/HP performance
Controls	Locked	Native	Native
Example	United States DOE Regulatory Test	4E Test Method, Energy Star CCHP CVP, DOE CCHP Challenge CVP	CSA SPE-07, Waseda University Method, AHRI 210/240 CVP

Ideally, either of the load-based test concepts could be implemented directly to determine equipment coefficient of performance (COPs) and ratings in lieu of traditional locked controls testing. However, with multiple variables to control, adaptive load tests may not be as repeatable or reproducible. To increase the chances of getting repeatable and reproducible tests, the research team adopted a native pre-defined load test as the controls validation procedure (CVP) to ensure the performance of the test unit aligns reasonably well when tested with fixed or overridden settings.

Overview of Test Method

Labs were instructed to test units in accordance with the 4E Test Method. This standard specifies the methodology for validation of control settings used when determining capacity and efficiency ratings of electrically driven, vapor compression, air-source or air-cooled systems rated at or below 19 kW capacity.

Using the pre-defined load test concept, the 4E Test Method replicates the imposed load measured for the regulatory tests and measures the electrical power output of the heat pump (W_{out}) and the electrical power required (W_{in}) at selected test point(s). Applying these measurements, the laboratory can then calculate

Equation 1: Coefficient of Performance

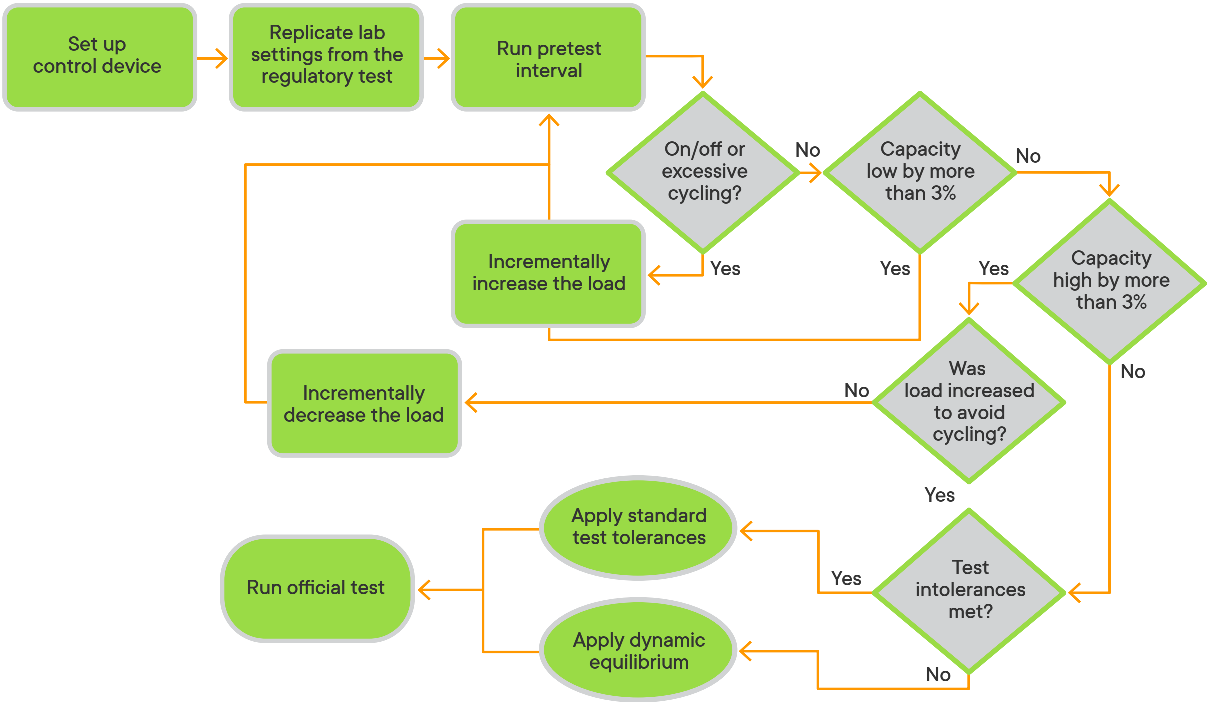
$$COP = \frac{W_{out}}{W_{in}}$$

Summarized descriptions of the 4E Test Method and procedure requirements are provided in this section.

Controls Validation Procedure

The 4E Test Method uses a controls validation test procedure, which complements regulatory tests. This procedure consists of applying a pre-defined load measured during regulatory test conditions. The UUT operates under its own native controls to validate operation and system performance characteristics. An overview of the test procedure process is shown in Figure 1.

Figure 1: Illustrative 4E Test Method Flow Chart



The indoor room/side of the test chamber subjects the unit under test to a target compensation load corresponding to the regulatory test measured capacity. The UUT responds accordingly to maintain the indoor test conditions. The outdoor room maintains constant conditions corresponding to the appropriate regulatory test conditions. If the UUT cycles excessively or the capacity produced is outside of the 3% tolerance range, lab technicians adjust the target compensation load. If the UUT continues to be out of stability specifications, technicians calculate performance through a dynamic equilibrium criterion.

Laboratory Setup

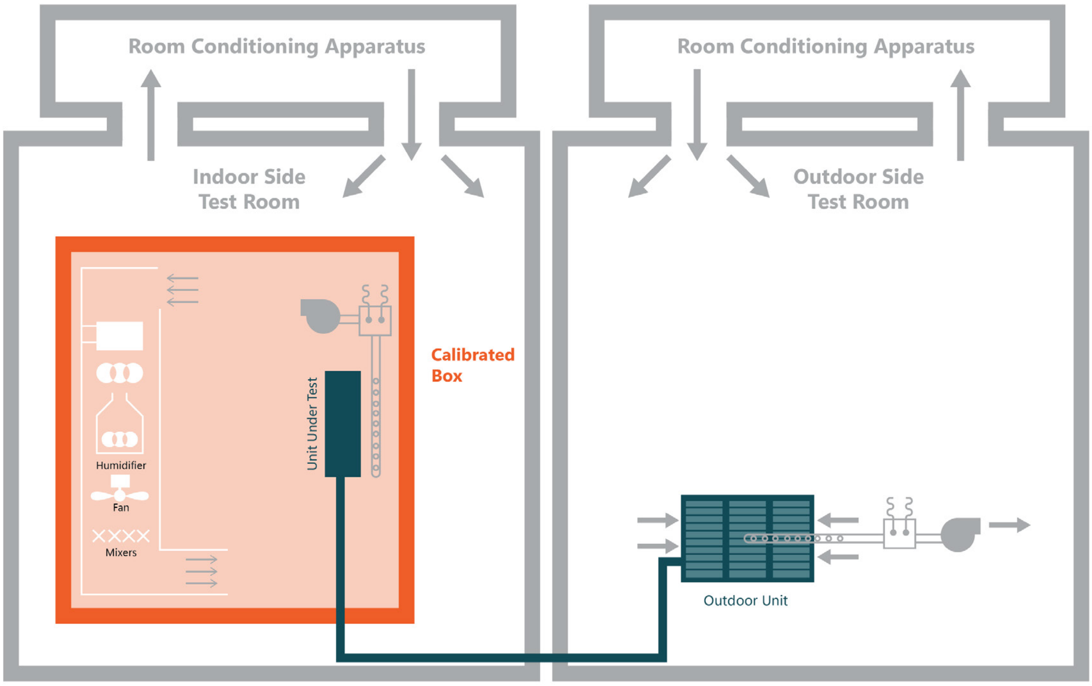
Across Europe and Asia, non-ducted units are more common and are typically tested in a calorimetric chamber. [12]However, in North America, where ducted units are more common, both types of units are tested in psychrometric chambers.[13] In Phase 2 of this project, researchers found that for load-based testing, non-ducted units should be tested either in a calorimeter or in a calibrated box inside a psychrometric room (hybrid method). The 4E Test Method employs the hybrid method for testing non-ducted units and a standard psychrometric chamber for testing ducted units.

The non-ducted hybrid method lab setup is shown in Figure 2. This hybrid method is used to replicate a room calorimetry approach and eliminate the potential issues that occur when attaching ductwork and airflow measurement apparatus to non-ducted units.

[12] Calorimetric chambers are highly insulated rooms that measure the amount of heat generated or absorbed by the UUT. Capacity of the UUT is quantified accurately by measuring the heat input or output from the reconditioning equipment. Smaller, non-ducted heat pumps easily fit in these rooms, which are typically smaller than psychrometric rooms.

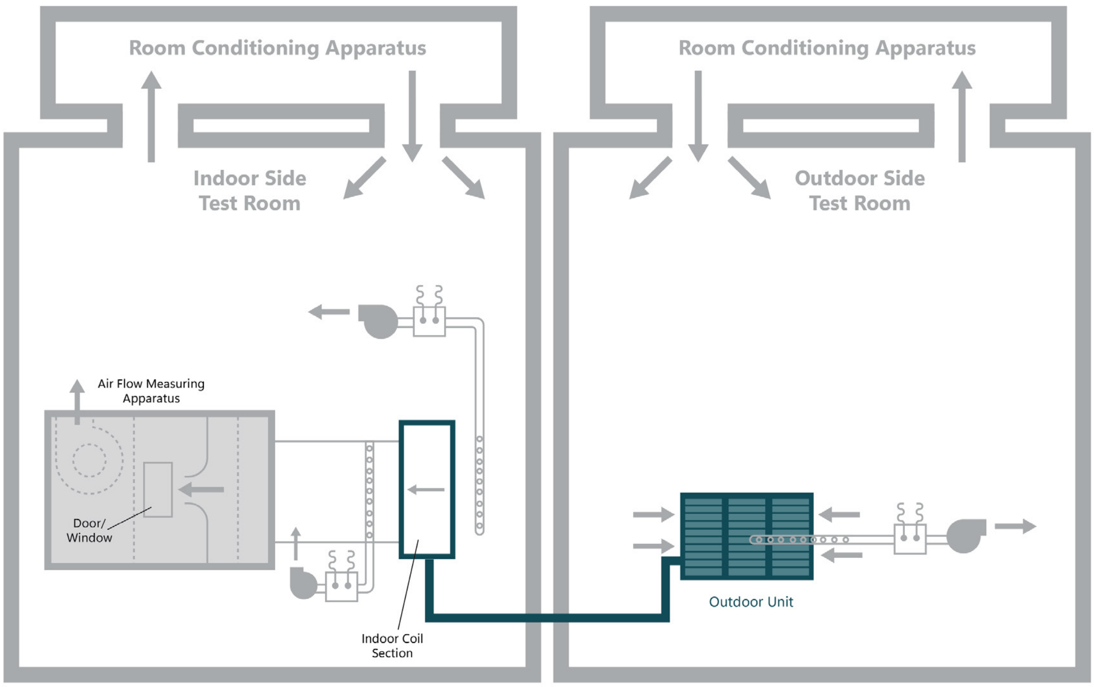
[13] Psychrometric chambers (Figure 3) utilize enthalpy measurements from a captured air stream exiting the UUT. Capacity is calculated from enthalpy difference between the inlet and outlet air streams and mass airflow. These rooms are typically larger than calorimetric chambers.

Figure 2: Non-ducted Lab Setup Showing Indoor Calibrated Box within Psychrometric Room



Ducted systems are tested in a standard psychrometric facility with modified parameters to allow for manual control of sensible and latent loads. This manual load control differs from current locked controls test procedures (such as ISO 5151[14]) but is necessary to achieve the changing loads needed for load-based testing. This ducted lab setup is shown in Figure 3.

Figure 3: Ducted Unit Lab Setup Showing Psychrometric Rooms[15]



[14] ISO 5151:2017 "Non-ducted air conditioners and heat pumps – Testing and rating for performance," International Organization for Standardization, https://www.iso.org/standard/63409.html

[15] ASHRAE 37-2009, Figure 1

Test Conditions

For the round robin testing, lab technicians established target compensation loads by baseline testing in accordance with the United States federally regulated test procedure: Appendix M1 to Subpart B of Part 430 (Appendix M1). [16]Table 8[17] and Table 14[18] in Appendix M1 establish heating mode and cooling mode conditions. These test conditions are summarized in Table 2 below. Test conditions are separated by mode (heating or cooling), compressor speed, and air volume rate. In cooling, all labs tested at several outdoor air temperatures. For heating, two of the labs tested at different outdoor air temperatures, but two of the labs did not have the capability to test at -8°C.

Table 2: Test Conditions

Condition Name	Conditions – Air Entering Indoor Unit		Conditions – Air Entering Outdoor Unit	
	Dry Bulb (°C)	Wet Bulb (°C)	Dry Bulb (°C)	Wet Bulb (°C)
Cooling Full Speed, High Temperature	27	19	35	24
Cooling Min Speed, Medium Temperature	27	19	28	18
Cooling Min Speed, Low Temperature	27	19	19	12
Heating Full Speed, High Temperature	21	16	8	6
Heating Min Speed, High Temperature	21	16	8	6
Heating Full Speed, Low Temperature	21	16	-8	-9

Pretest Interval

At each test condition, lab technicians establish a steady state during the pretest interval. Technicians set the chamber conditioning equipment at the conditioning levels used for the corresponding regulatory test or an advanced control feedback loop from the measured load. The UUT maintains the required outdoor test conditions within the tolerances listed in Table 3. Once the UUT and chamber conditioning system are running, the UUT can adjust to the load that is being applied to the indoor section and achieve equilibrium. Equilibrium is determined in the pretest interval when the UUT is no longer cycling its compressor on and off. Additional adjustment to the thermostat setpoint to account for offset/bias may be needed to achieve equilibrium at the proper indoor temperature.

[16] Appendix M1 to Subpart B of Part 430 – Uniform Test Method for Measuring the Energy Consumption of Central Air Conditioners and Heat Pumps, Code of Federal Regulations, https://www.ecfr.gov/current/title-10/chapter-II/subchapter-D/part-430/subpart-B/appendix-Appendix%20M1%20to%20Subpart%20B%20of%20Part%20430

[17] Appendix M1 to Subpart B of Part 430, Section 3.2.4: Tests for a Unit Having a Variable-Speed Compressor

[18] Appendix M1 to Subpart B of Part 430, Section 3.6.4: Tests for a Heat Pump Having a Variable-Speed Compressor

Table 3: Operating and Condition Tolerances for Target Compensation Load

	Test operating tolerance[19]	Test condition tolerance[20]
Indoor/Sampler dry-bulb:		
Entering temperature	0.56°C (1.0°F)	1.11°C (2.0°F)
Leaving temperature	2.22°C (4.0°F)	-
Indoor/Sampler wet-bulb:		
Entering temperature	0.56°C (1.0°F)	-
Outdoor dry-bulb:		
Entering temperature	0.56°C (1.0°F)	1.11°C (2.0°F)
Leaving temperature	-	-
Outdoor wet-bulb:		
Entering temperature	0.56°C (1.0°F)	0.56°C (1.0°F)
Leaving temperature	-	-
Air temperature surrounding calorimeter:		
Dry-bulb	1.11°C (2.0°F)	0.56°C (1.0°F)
Wet-bulb	0.56°C (1.0°F)	0.28°C (0.5°F)
External resistance to airflow:	12.44 Pa (0.05” H2O)	
Electrical voltage, % of reading:	2.0	1.5
Nozzle pressure drop, % of reading:	8.0	

Locked controls determination during pretest interval[1920]

A locked controls determination is met when the test-operating and test-condition tolerances are met for at least 30 minutes and the average capacity is achieved within +/-3% of the corresponding regulatory test capacity.

System that cycles in cooling mode

For cooling tests, if the system is unable to operate within +/-3% of the corresponding regulatory test capacity with no compressor on/off cycling, the total heating compensation load to the indoor room is incrementally increased (sensible and latent to maintain the sensible to total cooling capacity ratio within 3% of the regulatory test) until there is no on/off compressor cycling and locked controls requirements are achieved. Steady state is met when the test-operating and test-condition tolerances are met for at least 30 minutes. If cooling tests require increased loads to achieve steady state conditions, the 3% tolerance of the corresponding regulatory test capacity does not apply.

System that cycles in heating mode

For heating tests, if the system is unable to operate within +/-3% of the corresponding regulatory test capacity with no compressor on/off cycling, the cooling compensation load to the indoor room is increased until there is no on/off compressor cycling and locked controls requirements are achieved. Steady state

[19] Test operating tolerance is the maximum permissible variation of the observed range.
[20] Test condition tolerance is the maximum permissible variation of the mean average from the specified test condition.

is met when the test-operating and test-condition tolerances are met for at least 30 minutes. If heating tests require increased loads to achieve steady state conditions, the 3% tolerance of the corresponding regulatory test capacity does not apply.

Dynamic Equilibrium

If the system is unable to achieve locked controls requirements within 4 hours after the last incremental adjustment of the load addition to the indoor room, the pretest interval is complete and dynamic equilibrium criteria apply during the official test period.

Official Test Period

The official test period is 1 hour for systems that attain locked controls determination during the pretest interval and can maintain tolerances throughout the official test period.

Dynamic Equilibrium Criteria

Dynamic equilibrium is attained when both average capacity and average system power input measured in successive test period intervals are within 2 percent of each other. For systems that were unable to meet the locked controls determination in the pretest interval and utilize the dynamic equilibrium criteria, the official test period duration is described below:

- If regular cycling occurs, the test period intervals are at least 30 minutes in duration and comprise a whole number of system cycles.
- If regular cycling does not occur, the intervals are 30 minutes in duration.
- The official test measurements for dynamic equilibrium are the average values measured during these two successive intervals.

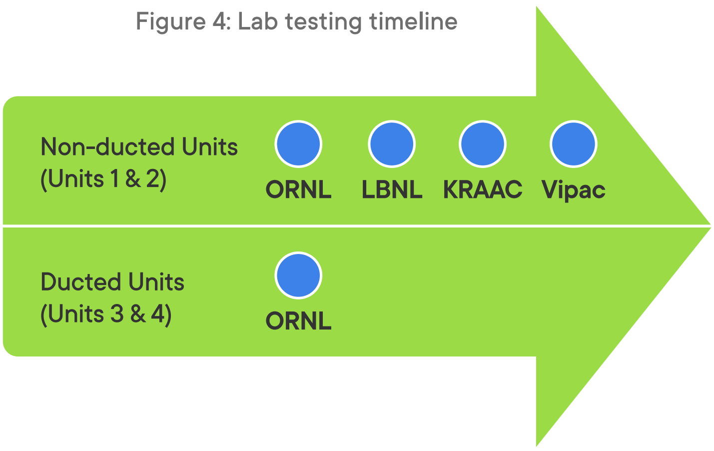
Test Laboratories

Four laboratories participated in the round robin testing. Labs included two national laboratories in the United States (Lawrence Berkeley National Laboratory (LBNL) and Oak Ridge National Laboratory (ORNL)) and independent laboratories in South Korea (Korea Refrigeration and Air Conditioning Assessment Center (KRAAC)), and Australia (Vipac). Because non-ducted and ducted units required different test measurement methods, they were shipped separately to labs with psychrometric or calorimetric capability.

Initially, all units were intended to test at five labs, but several factors limited testing of the ducted units. Ducted heat pumps are not as common outside of North America, and none of the labs outside this region were able to test them. A lab in Denmark (DTI) was not physically able to fit them in their test chamber. LBNL, KRAAC, and Vipac had limited funds and were only able to test the non-ducted units. Several labs had on-site equipment malfunctions that caused delays to the testing schedules.

Round robin testing was conducted from November 2021 to May 2024. Figure 4 illustrates the AC/HP round robin testing timeline by non-ducted and ducted units.

Figure 4: Lab testing timeline



Units Under Test

The units selected for this round robin testing are all variable speed, single zone split system heat pumps. Two units are non-ducted (high-wall mount) and two units are ducted (conventional static). Heat pumps were selected for testing instead of AC units to allow for both heating and cooling modes of operation.[21] UUTs were selected to cover a range of capacities, but also to allow for testing using the calorimetric room method, which is typically limited to 36,000 Btu/h (SI capacity). UUTs were also selected from four different manufacturers to represent different manufacturer control schemes. The units selected are summarized in Table 4.

Table 4: Units Under Test

Test Unit	Nominal Capacity	Configuration/ Indoor Arrangement
1	15,000 Btu/h / 4.5 kW	Non-ducted ASHP Single split Wall mount blower coil w/ remote Variable speed compressor
2	24,000 Btu/h / 7.0 kW	Non-ducted ASHP Single split Wall mount blower coil w/ remote Variable speed compressor
3	36,000 Btu/h / 10.5 kW	Ducted ASHP Single split Proprietary t-stat Variable speed compressor
4	36,000 Btu/h / 10.5 kW	Ducted ASHP Single split Proprietary t-stat Variable speed compressor

[21] The original scope of work emphasized research air conditioners only. The scope was expanded to include heating mode (but excluded transient operation such as defrost and oil return) during Phase 1.

3 Round Robin Test Results

As mentioned before, the primary goals of this round robin testing were to:

1. Determine reproducibility (i.e., consistency of results from lab to lab) of the test method.
2. Determine representativeness of the test method by showing any differences from regulatory test procedures.
3. Gain insights into the burden of the test.

This section reviews the results of the round robin testing and provides analyses for each of these aspects. There are a few limitations of the analyses. Not all units were tested in all labs, and due to lab capabilities, not all labs tested at Heating Low Temp Full Speed (-15°C). In addition, technicians sometimes did not follow the written test procedure and results were then deemed invalid. Invalid results are not included in the analysis in this section, but all results are included in the appendix. Table 5 provides a summary of each UUT's testing success rate as well as descriptions of how many tests were valid at each lab.

Table 5: UUT Testing Success Rate and Summaries

Unit	Testing Success Rate	ORNL	LBNL	KRAAC	Vipac
1	82%	5/6 tests valid	5/6 tests valid	3/5 tests valid	5/5 tests valid
2	91%	4/6 tests valid	6/6 tests valid	5/5 tests valid	5/5 tests valid
3	100%	6/6 tests valid	Did not test	Did not test	Did not test
4	100%	6/6 tests valid	Did not test	Did not test	Did not test

The first analysis is about reproducibility.

Reproducibility

One of the primary goals of the round robin testing was to determine the feasibility of conducting load-based testing at multiple labs. In order for a test method to be adopted by jurisdictions, it needs to be reproducible: labs in different locations must give similar results. Because the 4E Test Method is intended to be a validation method versus a direct rating method, results can be slightly different but must be within a variance that is acceptable to the authority having jurisdiction.

The research team chose coefficient of variance (CV) to examine reproducibility. CV is a statistical measure that describes the dispersion of a data set relative to its mean. CV, shown in Equation 2, is defined as the ratio of the standard deviation to the mean. A higher CV indicates greater variability, while a lower CV suggests more consistency. The research team considered a CV of 10% or lower as an indication of good reproducibility.

Equation 2: Coefficient of Variation

$$CV=\frac{\sigma}{\mu}$$

Where

σ = standard deviation of the 4E Test Method results.

μ = mean of the 4E Test Method results.

The test data was analyzed to determine the lab-to-lab reproducibility by using the CV values for capacity and COP from the two ductless systems. Because only one lab tested the ducted units, units 3 and 4 are not included in the reproducibility comparisons.

Unit 1 Reproducibility

Capacity testing of Unit 1 showed very good reproducibility. Table 6 shows the capacity that each lab measured at each test condition as well as the CV for these capacities. The labs had very good agreement: CVs were all 5% or below.

Table 6. Unit 1 Capacity Comparison and Coefficient of Variance

Test Condition	ORNL (W)	LBNL (W)	KRACC (W)	VIPAC (W)	CV
Cooling Hi Temp Full Speed	4,337	4,387	-	4,375	0%
Cooling Med Temp Lo Speed	1,709	1,614	1,566	1,623	3%
Cooling Min Temp Lo Speed	NR	1,948	1,897	1,891	1%
Heating Hi Temp Full Speed	5,221	5,281	-	5,382	1%
Heating Hi Temp Lo Speed	1,481	-	1,307	1,390	5%
Heating Lo Temp Full Speed	6,011	5,517	NR	NR	4%

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

As shown in Table 7, Unit 1's COP values were consistent but with slightly higher variation than the capacities. CVs for COP were generally below 10% except for the Cooling Min Temp Lo Speed test. LBNL showed a much higher COP than the other two labs. Both KRAAC and Vipac showed higher power consumption (176 W) for this test condition.

The research team has two theories that could account for this higher power consumption: modulation of components and auxiliary components. Many variable speed units, when allowed to control natively, stabilize capacity by varying combinations of modulating components. Compressor speed, indoor fan speed, outdoor fan speed, and expansion valve positions can stabilize at differing spots resulting in variations in efficiency—even at similar capacities. The second theory is that this higher consumption was due to some combination of auxiliary components: the base pan heater, the crankcase heater, and/or the built-in condensate pump. The labs did not sub-meter individual components, but these features could individually or cumulatively increase the power input enough to alter the results. As described in the Changes to Method of Test section, test procedure improvements that standardize the approach direction, rate of temperature change, and instrumentation response time would increase reproducibility by ensuring all labs approach the target compensation load in a well-defined, consistent manner.

Table 7. Unit 1 COP Comparison and Coefficient of Variance

Test Condition	ORNL COP	LBNL COP	KRACC COP	VIPAC COP	CV
Cooling Hi Temp Full Speed	4.2	4.0	-	4.3	4%
Cooling Med Temp Lo Speed	7.8	8.0	7.3	6.5	8%
Cooling Min Temp Lo Speed	NR	15.3	10.7	11.2	16%
Heating Hi Temp Full Speed	4.2	4.0	-	3.9	4%
Heating Hi Temp Lo Speed	5.2	-	4.4	4.9	7%
Heating Lo Temp Full Speed	2.2	2.1	NR	NR	3%

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

Unit 2 Reproducibility

Capacity testing of Unit 2 also showed very good reproducibility. Table 8 shows the capacity that each lab measured at each test condition as well as the CV for these capacities. The labs had very good agreement: CVs were all 5% or below, indicating good reproducibility.

Table 8. Unit 2 Capacity Comparison and Coefficient of Variance

Test Condition	ORNL (W)	LBNL (W)	KRACC (W)	VIPAC (W)	CV
Cooling Hi Temp Full Speed	-	6,511	6,581	6,300	2%
Cooling Med Temp Lo Speed	1,986	1,872	1,935	1,755	5%
Cooling Min Temp Lo Speed	2,057	2,045	1,962	1,917	3%
Heating Hi Temp Full Speed	7,291	7,023	7,802	8,000	5%
Heating Hi Temp Lo Speed	-	2,934	2,989	3,000	1%
Heating Lo Temp Full Speed	5,525	5,536	NR	NR	0%

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

COP reproducibility for Unit 2 is shown in Table 9. CVs were all below 10%, again showing good reproducibility.

Table 9. Unit 2 COP Comparison and Coefficient of Variance

Test Condition	ORNL COP	LBNL COP	KRACC COP	VIPAC COP	CV
Cooling Hi Temp Full Speed	-	3.8	3.9	3.7	2%
Cooling Med Temp Lo Speed	6.5	6.8	5.4	5.1	2%
Cooling Min Temp Lo Speed	10.6	10.8	11.2	9.0	8%
Heating Hi Temp Full Speed	3.3	3.3	2.7	3.0	8%
Heating Hi Temp Lo Speed	-	4.1	3.9	3.9	2%
Heating Lo Temp Full Speed	2.0	2.0	NR	NR	0%

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

Because Units 3 and 4 were only tested at ORNL, the research team was not able to include those units in reproducibility comparisons.

Generally, the reproducibility was very good in Units 1 and 2. CVs were below 10% in all but one test condition of one unit. The research team believes that a virtual building load (VBL[22]) would likely increase the reproducibility. In fact, Vipac initially had difficulty with unit stability, but after implementing a modified VBL they had much better success.

Next, we look at how the representativeness of the 4E Test Method by comparing 4E Test Method results to the results from locked controls tests.

[22] Virtual Building Load is a load-based or native controls test procedure during which the software that controls the indoor test room conditions (i.e., operates the indoor room reconditioning system) is programmed to mimic the response of building heating or cooling in real time by monitoring the capacity of the unit under test and adjusting the indoor room conditions according to the virtual building model. The virtual building model defines the time-dependent rate of change of the indoor room temperature and humidity conditions as a function of the target building load and the measured capacity of the tested system.

Representativeness

The underlying reason to create the 4E Test Method is to more accurately represent the performance of air conditioners and heat pumps. While representativeness traditionally means that the test is more indicative of how a unit behaves in homes, this research did not include comparisons to field performance. The research team here uses representativeness to show potential differences of the UUT under native control compared to the locked-controls testing. The research team expects to see different results from the 4E Test Method when compared to locked controls testing for some systems. To determine the representativeness of the 4E Test Method, the research team compares the capacity and COP as determined in the 4E results with results from locked controls tests. For the comparison, the team used the United States federal regulatory test[23], commonly called Appendix M1.

Appendix M1 has incorporated a controls verification procedure based on the 4E Test Method, and with this incorporation developed a set of tolerances to indicate acceptable compliance.[24] The following analysis uses these same tolerances: 6% for capacity and 10% for efficiency. Efficiency tolerance is one-sided—efficiencies above the values determined from locked controls testing are allowed.

The color scheme in Figure 7 indicates whether test results were in or out of tolerance in the following subsections.

Figure 5: Test Outcomes

Test results are within tolerance	(green)
Test results are out of tolerance	(red)

Unit 1 Representativeness

First, we compare the 4E capacity results to locked controls test results. For Unit 1, the capacities measured by the 4E Test Method at each lab were generally consistent with the M1 results (Table 10). Of the 18 test results, only 3 of the capacities were out of tolerance, and by a small margin. Additional provisions in the test procedure to help standardize testing ramps and thermal mass response along with increased familiarity with the procedure itself may bring these differences within tolerance.

Table 10. Unit 1 Capacity Comparison to Locked Controls

	COP % Difference from M1 (Locked controls)			
Test Condition	**ORNL**	**LBNL**	**KRACC**	**VIPAC**
Cooling Hi Temp Full Speed	3%	2%	-	2%
Cooling Med Temp Lo Speed	7%	5%	4%	5%
Cooling Min Temp Lo Speed	-	0%	-1%	-2%
Heating Hi Temp Full Speed	3%	1%	-	0%
Heating Hi Temp Lo Speed	7%	-	4%	5%
Heating Lo Temp Full Speed	0%	8%	-	-

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

[23] As described in the Test Conditions section of this report

[24] Section III.k.2 in https://www.federalregister.gov/documents/2024/04/05/2024-04784/energy-conservation-program-test-procedure-for-central-air-conditioners-and-heat-pumps

The efficiencies as measured by the 4E test show quite different results. Of the 18 valid test results, 10 of them were out of tolerance. This suggests that Unit 1 has controls that will ensure that the capacity is met, but can do so at varying efficiency levels. It also shows that the locked controls test is not always representative of actual performance when the system is allowed to control natively.

Table 11. Unit 1 COP Comparison to Locked controls

	Capacity % Difference from M1 (Locked controls)			
Test Condition	**ORNL**	**LBNL**	**KRACC**	**VIPAC**
Cooling Hi Temp Full Speed	-2%	4%	-	-5%
Cooling Med Temp Lo Speed	12%	10%	18%	27%
Cooling Min Temp Lo Speed	-	-2%	29%	25%
Heating Hi Temp Full Speed	5%	11%	-	13%
Heating Hi Temp Lo Speed	15%	-	28%	19%
Heating Lo Temp Full Speed	0%	7%	-	-

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

Unit 2 Representativeness

Unit 2's capacity results showed good alignment with M1 in cooling mode. However, the first two labs were out of tolerance in the Heating High Temp Full Speed tests. This difference could be attributed to a leak in the refrigerant system at the charging port flare nut. This leak was discovered after the system was shipped to the third lab. After discovery, the system was recharged, and the flare nut was replaced. This likely resulted in the system being slightly undercharged for the first and second labs. Because less refrigerant is required in cooling mode due to the charge compensator, this difference in charge would not necessarily be apparent in cooling results. Other than the tests performed with uncertain refrigerant, all of the tests were within tolerance.

Table 12. Unit 2 Capacity Comparison to Locked controls

	Capacity % Difference from M1 (Locked controls)			
Test Condition	**ORNL**	**LBNL**	**KRACC**	**VIPAC**
Cooling Hi Temp Full Speed	-	0.3%	-0.8%	3.5%
Cooling Med Temp Lo Speed	1.5%	-0.2%	0.7%	-2.0%
Cooling Min Temp Lo Speed	0.1%	-0.1%	-1.4%	-2.1%
Heating Hi Temp Full Speed	6.3%	9.7%	-0.3%	-2.9%
Heating Hi Temp Lo Speed	-	0.2%	1.2%	1.4%
Heating Lo Temp Full Speed	-1.0%	-1.2%	NR	NR

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

The Unit 2 efficiency results show a similar result to Unit 1. Of the 23 valid tests, 7 of them show COP that were out of tolerance, again indicating that the locked controls test results are not always indicative of field performance.

Table 13. Unit 2 COP Comparison to Locked controls

	COP % Difference from M1 (Locked controls)			
Test Condition	**ORNL**	**LBNL**	**KRACC**	**VIPAC**
Cooling Hi Temp Full Speed	-	-0.1%	-3.2%	3.6%
Cooling Med Temp Lo Speed	4.8%	0.6%	21.3%	25.2%
Cooling Min Temp Lo Speed	-3.7%	-5.4%	-8.7%	12.8%
Heating Hi Temp Full Speed	2.6%	4.1%	20.3%	11.8%
Heating Hi Temp Lo Speed	-	6.7%	10.6%	11.3%
Heating Lo Temp Full Speed	-0.5%	0.0%	-	-

Note: "-" indicates invalid test results and "NR" indicates No Results when a lab did not perform the test.

Unit 3 Representativeness

Unit 3 tested successfully and was within capacity tolerances for all tests, though this unit was tested at ORNL only. As shown in Table 14, the load-based testing results align closely with the M1 test results for capacity.

Table 14: Unit 3 Load-based and M1 Test Results, Capacity

Capacity % Difference from M1 (Locked controls)	
Test Condition	**ORNL**
Cooling Hi Temp Full Speed	+0.5%
Cooling Med Temp Lo Speed	-3.5%
Cooling Min Temp Lo Speed	-3.2%
Heating Hi Temp Full Speed	+1.0%
Heating Hi Temp Lo Speed	-0.2%
Heating Lo Temp Full Speed	+1.2%

Unit 3 tested successfully and was within efficiency tolerances for all tests, though this unit was tested at ORNL only. As shown in Table 15, the load-based testing results align closely with the M1 test results for COP.

Table 15: Unit 3 Load-based and M1 Test Results, COP

COP % Difference from M1 (Locked controls)	
Test Condition	**ORNL**
Cooling Hi Temp Full Speed	+0.2%
Cooling Med Temp Lo Speed	+2.3%
Cooling Min Temp Lo Speed	-1.9%
Heating Hi Temp Full Speed	+3.2%
Heating Hi Temp Lo Speed	-1.0%
Heating Lo Temp Full Speed	-0.5%

Unit 4 Representativeness

Unit 4 tested successfully and was within capacity tolerances for five out of six tests, though this unit was tested at ORNL only. For the Cooling Hi Temp Full Speed test, the lab mistakenly set the thermostat to the lowest achievable setpoint, causing the compressor to overspeed and the unit to exceed the capacity tolerance.

Other than the Cooling Full Speed, High Temperature test, where capacity exceeds tolerance, the load-based test results align closely with the M1 test results for capacity, shown in Table 16.

Table 16: Unit 4 Load-based and M1 Test Results, Capacity

	Capacity % Difference from M1 (Locked controls)
Test Condition	**ORNL**
Cooling Hi Temp Full Speed	-9.8%
Cooling Med Temp Lo Speed	-1.6%
Cooling Min Temp Lo Speed	+0.2%
Heating Hi Temp Full Speed	0.0%
Heating Hi Temp Lo Speed	-0.7%
Heating Lo Temp Full Speed	-4.6%

Unit 4 tested successfully and was within efficiency tolerances for all tests, though this unit was tested at ORNL only. As shown in Table 17, the load-based testing results align closely with the M1 test results for COP. As described in the paragraph above, instead of imposing the target load, the lab used the “buried thermostat” technique and the compressor over sped. This unit uses vapor injection not only at low ambient, but also at high ambient, which cools the compressor and boosts the capacity. The vapor injection kept the power from following the normal power curve and allowed for increased capacity with only marginal efficiency loss.

Table 17: Unit 4 Load-Based and M1 Test Results, COP

	COP % Difference from M1 (Locked controls)
Test Condition	**ORNL**
-9.8%	+2.6%
-1.6%	+3.7%
+0.2%	-3.6%
0.0%	0.0%
-0.7%	+0.7%
-4.6%	-3.8%

Overall, the comparisons to steady-state testing show that units would match locked controls capacity but not efficiencies. This suggests that variable speed heat pump controls are designed to meet temperature requirements but may sacrifice efficiency to do so. It also shows that locked controls testing does not always accurately represent variable speed performance due to the controls and that the 4E method will often “catch” this difference.

Test Burden

One important factor when developing new test methods is to understand burden: how difficult and time-intensive the test is to perform. The concept of load-based testing was new to most labs as locked controls testing has been the norm in most test standards throughout the world. The learning curve differed from lab to lab and the research team had to introduce lab technicians to the method as the testing was being conducted.

Several labs had issues with their on-site equipment, which made it difficult to determine a quantitative estimate of burden. However, the team did gather qualitative information on the difficulties labs experienced when performing the test method. This section discusses the difficulties experienced by each lab.

General Test Burden

Testing HVAC units in both psychrometric and calorimetric rooms presents unique challenges. In a psychrometric room, the primary difficulty lies in getting the chamber controls to adjust the room temperature slowly enough for the unit to stabilize. On the other hand, in a calorimetric room the thermal mass of the chamber is relatively small, which can result in the unit "overpowering" the space and causing fluctuations in the room temperature. These fluctuations make the unit ramp up and down.

Regardless of the test room used, achieving a steady state for the compressor takes time, as each unit's control logic varies. Lab technicians described a learning curve with each unit because their unique control systems dictate how and when the compressor speed fluctuates. The thermal mass of lab apparatus also resulted in extended test times due to continued unit cycling.

Finally, labs reported that some units cycled their control algorithm when the overridden controls were released.

Specific Lab Feedback

KRAAC reported that getting the units to stabilize was very time intensive. They offered that stabilization took a half day, about twice as long as during locked controls testing. Overall testing was 8 hours/unit.

Vipac reported setup time of a full day, but again that was twice as long as their standard setup time.

Vipac had an extraordinarily difficult time running the tests. This was mainly due to their lack of automated controls logic for their reconditioning equipment. The research team assisted Vipac in getting a virtual building load software adjustment and the tests were much more successful after that. Vipac's balanced ambient calorimeter required the addition of a heat rejection controller to allow the unit under test to be subjected to a constant load.

Vipac also offered, "We believe this test method is a true indication of a unit's performance in the real world. We believe it would promote further development into unit logic, improving real world efficiency and performance."

The labs universally requested more detailed instructions for selecting settings on the unit and more insight on how to set up the test chambers for target compensation load control as opposed to unit load matching.

4 Conclusion and Implementation

This section provides suggestions for further developing and adopting the 4E Test Method.

Changes to Method of Test

The 4E Test Method provided a solid foundation for a load-based test method of variable speed systems. However, the research team has identified some areas of improvement throughout the round robin test process.

The 4E method of test utilizes a compensation load-based test procedure where the indoor room is subjected to a load and the AC or HP to be tested responds accordingly as it tries to maintain the desired indoor conditions. This method of test generally works well at higher loads. At lower loads, however, the impact of the laboratory's thermal mass and the reconditioning systems response can cause varying unit responses. These issues can also cause unstable test conditions as the unit under test tries to respond to a lower load in laboratories with a large thermal mass.

One method to correct for this is a virtual building load (VBL). A VBL is software that controls the indoor reconditioning system to mimic the response of a building in real time. The VBL monitors the capacity of the unit under test and adjusts the indoor room conditions according to the virtual building model. The virtual building model defines the time-dependent rate of change of the indoor room temperature and humidity conditions as a function of the target building load and the measured capacity of the tested system. This is used to standardize and overcome the lab thermal mass and interaction with differing UUT control approaches. An example of how the virtual load can be adjusted is given in equations I1 through I6 in AHRI Standard 210/240-2024 Appendix I.

Another issue discovered during the round robin testing is that units, when released from locked controls testing, would cycle their controls. This resulted in the test room starting conditions significantly varying when the unit's native controls began to engage. Separating the locked testing from the native test intervals would correct for this. Again, Standard 210/240-2024 Appendix I describes this process of separating the tests.

Adoption/Use of Test Method

The framework of the 4E load-based Test Method for variable speed systems has already been adopted in recent standards development. AHRI Standard 210/240-2024 Appendix I is a Controls Verification Procedure (CVP) based on the 4E Test Method, with improvements as noted in the section above. Additionally, the United States' federal test procedure for residential air conditioners and heat pumps references AHRI 210/240 and incorporates the same CVP.

To advance the adoption of cold-climate heat pumps, the United States Department of Energy launched a cold-climate heat pump challenge. Eight manufacturers are participating in the development of high-performance heat pumps. Part of this challenge is that the units must pass a CVP based on the 4E Test Method.[25]

The 4E Test Method has already started to have an impact in the United States. By incorporating a virtual building load to increase the reproducibility and lower the test burden, the research team expects the 4E Test Method will improve the representativeness of air conditioner and heat pump testing in other jurisdictions as well.

[25] https://www.energy.gov/sites/default/files/2021-10/bto-cchp-tech-challenge-spec-102521.pdf

Appendix 1: Detailed Round Robin Results

This section lists the detailed results from Phase 4 testing.

Unit 1 Detailed Results

Strikethrough indicates the lab did not follow the test procedure. “NR” means no results—the lab did not perform a test at that condition.

Table 18: Unit 1 Capacities at Each Test Condition

Test Condition	M1 (W)	ORNL (W)	LBNL (W)	KRACC (W)	VIPAC (W)
Cooling Hi Temp Full Speed	4,459	4,337	4,387	~~4,933~~	4,375
Cooling Med Temp Lo Speed	1,393	1,709	1,614	1,566	1,623
Cooling Min Temp Lo Speed	1,964	NR	1,948	1,897	1,891
Heating Hi Temp Full Speed	5,357	5,221	5,281	~~6,387~~	5,382
Heating Hi Temp Lo Speed	1,068	1,481	~~1,465~~	1,307	1,390
Heating Lo Temp Full Speed	6,011	6,011	5,517	NR	NR

Table 19: Unit 1 COP at Each Test Condition

Test Condition	M1	ORN	LBNL	KRACC	VIPAC
Cooling Hi Temp Full Speed	4.1	4.2	4.0	3.5	4.3
Cooling Med Temp Lo Speed	8.9	7.8	8.0	7.3	6.5
Cooling Min Temp Lo Speed	15.0	NR	15.3	10.7	11.2
Heating Hi Temp Full Speed	4.5	4.2	4.0	3.4	3.9
Heating Hi Temp Lo Speed	6.1	5.2	5.1	4.4	4.9
Heating Lo Temp Full Speed	2.2	2.2	2.1	NR	NR

- For the Cooling and Heating Hi Temp Full Speed test, KRACC did not fully follow the test procedure. The load was applied above the target and the unit over sped the compressor above the “nominal” speed, resulting in higher capacity and lower efficiency. The “full” speed does not necessarily mean the compressor is operating at its maximum level in all cases.
- The Cooling Min Temp Lo Speed test had repeatable results on capacity. However, KRAAC and Vipac showed higher power consumption (176 W) for this test condition. The research team has two working theories. It has been observed that systems, when allowed to control natively, may stabilize around a similar capacity with varying combinations of modulating component parameters. Compressor speed, indoor fan speed, outdoor fan speed, and LEV positions can stabilize at differing operating levels resulting in variations in efficiency. The second theory is that this higher consumption was due to some combination of auxiliary components: the base pan heater, the crankcase heater, and/or the built-in condensate pump. The labs did not sub-meter individual components, but these features could individually or cumulatively increase the power input enough to alter the results.
- The Heating Hi Temp Lo Speed tests showed that this system could not turn down to the level measured during the controls override test. All four labs measured capacities between 4.0% and 6.9% higher than the M1 results.

Unit 2 Detailed Results

Table 20: Unit 2 Capacities at Each Test Condition

Test Condition	M1 (W)	ORNL (W)	LBNL (W)	KRACC (W)	VIPAC (W)
Cooling Hi Temp Full Speed	6,532	~~7,320~~	6,511	6,581	6,300
Cooling Med Temp Lo Speed	1,886	1,986	1,872	1,935	1,755
Cooling Min Temp Lo Speed	2,052	2,057	2,045	1,962	1,917
Heating Hi Temp Full Speed	7,778	7,291	7,023	7,802	8,000
Heating Hi Temp Lo Speed	2,922	~~4,314~~	2,934	2,989	3,000
Heating Lo Temp Full Speed	5,469	5,525	5,536	NR	NR

Table 21: Unit 2 COP at Each Test Condition

Test Condition	M1	ORNL	LBNL	KRACC	VIPAC
Cooling Hi Temp Full Speed	3.8	3.4	3.8	3.9	3.7
Cooling Med Temp Lo Speed	6.8	6.5	6.8	5.4	5.1
Cooling Min Temp Lo Speed	10.3	10.6	10.8	11.2	9.0
Heating Hi Temp Full Speed	3.4	3.3	3.3	2.7	3.0
Heating Hi Temp Lo Speed	4.4	3.8	4.1	3.9	3.9
Heating Lo Temp Full Speed	2.0	2.0	2.0	NR	NR

- Cooling Hi Temp Full Speed—ORNL did not fully follow the test procedure for this condition. The load was applied above the target and the unit over sped the compressor above the "nominal" speed resulting in higher capacity and lower efficiency. The "full" speed does not necessarily mean the compressor is operating at its maximum level in all cases.
- Cooling Med Temp Lo Speed—KRAAC and Vipac were unable to stop the unit from cycling due to thermal mass issues and aggressive controls. These labs applied the dynamic equilibrium criteria, matching the time averaged integrated capacity within the tolerance of the locked controls. The on/off cycling of the compressor resulted in increased power and decreased efficiency.
- Cooling Min Temp Lo Speed—Vipac was unable to stop the unit from cycling due to thermal mass issues and aggressive controls. They applied the dynamic equilibrium criteria matching the time averaged integrated capacity within the tolerance of the locked controls. The on/off cycling of the compressor resulted in increased power and decreased efficiency.
- Heating Hi Temp Full Speed—The low capacities at ORNL and LBNL could be attributed to a leak in the refrigerant system at the charging port flare nut, which was discovered after the system was shipped to the third lab. The system has a charge compensator. However, the results showed a consistent drop from the M1 test until the leak was discovered and repaired. The system was recharged and the flare nut was replaced. This potentially resulted in the system having the optimal charge with which it was tested during the original M1 test for the third and fourth labs while being slightly undercharged at ORNL and LBNL.
- Heating Hi Temp Lo Speed—Three of the four labs were able to test at the target capacity levels. ORNL failed to reduce the target compensation load to the appropriate level. However, all labs were unable to replicate the power consumption from the M1 test and had lower native efficiencies. It appears

the LEV's, compressor speeds, and possibly the indoor and outdoor fan speeds were not adequately represented in the M1 tests.

Unit 3 Detailed Results

Table 22: Unit 3 Capacities at Each Test Condition

Test Condition	M1 (W)	ORNL (W)
Cooling Hi Temp Full Speed	9,810	9,757
Cooling Med Temp Lo Speed	3,664	4,004
Cooling Min Temp Lo Speed	4,053	4,367
Heating Hi Temp Full Speed	9,663	9,564
Heating Hi Temp Lo Speed	2,893	2,902
Heating Lo Temp Full Speed	5,108	5,046

Table 23: Unit 3 COP at Each Test Condition

Test Condition	M1	ORNL
Cooling Hi Temp Full Speed	3.7	3.7
Cooling Med Temp Lo Speed	7.0	7.2
Cooling Min Temp Lo Speed	11.5	11.3
Heating Hi Temp Full Speed	3.2	3.3
Heating Hi Temp Lo Speed	4.4	4.4
Heating Lo Temp Full Speed	2.0	2.0

- As shown in the tables above, only one lab was able to conduct testing on the ducted system. However, that lab showed each test within capacity and efficiency tolerances for all tests. The load-based testing results for this unit align closely with the M1 test results.

Unit 4 Detailed Results

Table 24: Unit 4 Capacities at Each Test Condition

Test Condition	M1 (W)	ORNL (W)
Cooling Hi Temp Full Speed	10,353	11,372
Cooling Med Temp Lo Speed	5,936	6,103
Cooling Min Temp Lo Speed	6,484	6,463
Heating Hi Temp Full Speed	9,277	9,277
Heating Hi Temp Lo Speed	5,899	5,960
Heating Lo Temp Full Speed	8,828	9,230

Table 25: Unit 4 COP at Each Test Condition

Test Condition	M1	ORNL
Cooling Hi Temp Full Speed	4.0	3.9
Cooling Med Temp Lo Speed	5.8	5.6
Cooling Min Temp Lo Speed	9.0	8.7
Heating Hi Temp Full Speed	4.1	4.1
Heating Hi Temp Lo Speed	4.5	4.4
Heating Lo Temp Full Speed	2.1	2.2

- Similar to Unit 3, only one lab was able to conduct testing on the ducted system. Again, that lab showed each test within capacity and efficiency tolerances for all tests. The load-based testing results for this unit align closely with the M1 test results.